Notes on Model Theory

Gabriel Conant

June 8, 2016

These notes were prepared for the first week of the Notre Dame Center for Mathematics Thematic Program on Model Theory (June 6, 2016 through June 10, 2016). The progression of topics largely follows *Model Theory: An Introduction* by David Marker, and many of the exercises are taken from this text. The material assumes no prior knowledge of model theory or mathematical logic. Many of the examples and exercises require some familiarity with groups and fields.

Contents

By now in your mathematical education, you have studied (or at least heard of) many areas of mathematics which focus on the "theory" of a certain kind of abstract mathematical structure. For example: group theory, ring theory, field theory, or graph theory. These notes will introduce you to *model theory*, which provides a formal unifying framework, with which one can study any of these examples (and more).

1 Languages and Structures

To motivate our main definition, recall some common mathematical structures.

Example 1.1.

1. A *group* is a tuple $(G, *, e)$ where
   - G is a set,
   - $*$ is a binary function on G,
   - e is an element of G,
   - certain axioms are satisfied.
2. An *ordered ring* is a tuple $(R, +, -, \cdot, <, 0, 1)$ where
   - R is a set,
   - $+, -, \cdot$ are binary functions on R,
   - $0, 1$ are elements of R,
   - $<$ is a binary relation on R (i.e. a subset of $R \times R$),
   - certain axioms are satisfied.
3. A *graph* is a tuple (V, E) where
   - V is a nonempty set,
   - E is a binary relation on V,
   - certain axioms are satisfied.

Recall that, given a set X and an integer $n \geq 1$, an *n-ary relation on X* is a subset of X^n.

Definition 1.2.

1. A **structure** is a tuple $\mathcal{M} = \Big(M, (f_i^{\mathcal{M}})_{i \in I}, (R_j^{\mathcal{M}})_{j \in J}, (c_k^{\mathcal{M}})_{k \in K}\Big)$ where
   - M is a nonempty set,
   - each $f_i^{\mathcal{M}}$ is a function on M of arity $n_i \geq 1$,
   - each $R_j^{\mathcal{M}}$ is a relation on M of arity $m_j \geq 1$,
   - each $c_k^{\mathcal{M}}$ is an element of $\mathcal{M}$.
2. A structure $\mathcal{M}$ has an associated **language** of symbols

$$\mathcal{L} = \{f_i : i \in I\} \cup \{R_j : j \in J\} \cup \{c_k : k \in K\},$$

   which are called **function symbols**, **relation symbols**, and **constant symbols**, respectively. Each function and relation symbol has an implicit **arity** $n \geq 1$.

In practice, one often first fixes a language $\mathcal{L}$, and considers different structures in that language (i.e. *$\mathcal{L}$-structures*). The following are some languages that we will use frequently.

Definition 1.3.

1. Let $\mathcal{L}_g = \{*, e\}$ be the *language of groups*, where $*$ is a binary function symbol and e is a constant symbol.

2. Let $\mathcal{L}_r = \{+, -, \cdot, 0, 1\}$ be the *language of rings (with unity)*, where $+, -, \cdot$ are binary function symbols and $0, 1$ are constant symbols.[1]

3. Let $\mathcal{L}_o = \{<\}$ be the *language of orders*, where $<$ is a binary relation symbol. Define the *language of ordered groups* $\mathcal{L}_{og} = \mathcal{L}_g \cup \{<\}$ and the *language of ordered rings* $\mathcal{L}_{or} = \mathcal{L}_r \cup \{<\}$.

4. Let $\mathcal{L}_{gr} = \{E\}$ be the *language of graphs*, where E is a binary relation symbol.

Note that there is no substantive difference between $\mathcal{L}_o$ and $\mathcal{L}_{gr}$. Note also that any group can be interpreted as an $\mathcal{L}_g$-structure, but an $\mathcal{L}_g$-structure *does not necessarily need to be group.* In particular, unlike Example 1.1, Definition 1.2 says nothing about "certain axioms being satisfied" (this comes later in Section 3). For example, we may define an $\mathcal{L}_g$-structure $(\mathbb{N}, *, 472)$, where $x * y = x^y + \lfloor \log(x + y + 1) \rfloor$.

When studying mathematical objects it is useful to work with maps which preserve a certain amount of structure. We can generalize such notions to arbitrary $\mathcal{L}$-structures.

Definition 1.4. Let $\mathcal{L}$ be a language and let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures.

1. A function $\sigma : M \longrightarrow N$ is an **$\mathcal{L}$-embedding** if σ is injective and:

   (i) for any function symbol f in $\mathcal{L}$ of arity n, and $a_1, \ldots, a_n \in M$,
   $$\sigma(f^{\mathcal{M}}(a_1, \ldots, a_n)) = f^{\mathcal{N}}(\sigma(a_1), \ldots, \sigma(a_n));$$

   (ii) for any relation symbol R in $\mathcal{L}$ of arity n, and $a_1, \ldots, a_n \in M$,
   $$(a_1, \ldots, a_n) \in R^{\mathcal{M}} \quad \Leftrightarrow \quad (\sigma(a_1), \ldots, \sigma(a_n)) \in R^{\mathcal{N}};$$

   (iii) for any constant symbol c in $\mathcal{L}$,
   $$\sigma(c^{\mathcal{M}}) = c^{\mathcal{N}}.$$

   In this case, we say σ is an **embedding from $\mathcal{M}$ to $\mathcal{N}$**, and write $\sigma : \mathcal{M} \longrightarrow \mathcal{N}$.

2. An **$\mathcal{L}$-isomorphism from $\mathcal{M}$ to $\mathcal{N}$** is a bijective $\mathcal{L}$-embedding from $\mathcal{M}$ to $\mathcal{N}$.

3. $\mathcal{M}$ and $\mathcal{N}$ are **isomorphic**, written $\mathcal{M} \cong \mathcal{N}$, if there is an $\mathcal{L}$-isomorphism $\sigma : \mathcal{M} \longrightarrow \mathcal{N}$.

4. $\mathcal{M}$ is a **$\mathcal{L}$-substructure** of $\mathcal{N}$, written $\mathcal{M} \subseteq \mathcal{N}$, if $M \subseteq N$ and the inclusion map $\iota : M \longrightarrow N$, such that $\iota(a) = a$ for all $a \in M$, is an $\mathcal{L}$-embedding. In other words $\mathcal{M} \subseteq \mathcal{N}$ if and only if $M \subseteq N$ and:

   (i) for any function symbol f in $\mathcal{L}$ of arity n, $f^{\mathcal{M}} = f^{\mathcal{N}}|_{M^n}$,
   (ii) for any relation symbol R in $\mathcal{L}$ of arity n, $R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^n$,
   (iii) for any constant symbol c in $\mathcal{L}$, $c^{\mathcal{M}} = c^{\mathcal{N}}$.

[1] We are including the symbol "$-$" for convenience (see Example 2.13(3)).

Example 1.5.

1. $(\mathbb{Z}, +, 0)$ is an $\mathcal{L}_g$-substructure of $(\mathbb{R}, +, 0)$.
2. $(\mathbb{N}, +, 0)$ is an $\mathcal{L}_g$-substructure of $(\mathbb{Z}, +, 0)$.
3. The function $x \mapsto e^x$ is an $\mathcal{L}_r$-embedding from $(\mathbb{Z}, +, 0)$ to $(\mathbb{R}^+, \cdot, 1)$.
4. Recall that if (V, E) is a graph, then an *subgraph* of (V, E) is a graph (W, F) where $W \subseteq V$ and $E \subseteq F$. A subgraph (W, F) is an *induced subgraph* if $F = W^2 \cap E$. Now suppose (V, E) is a graph and (W, F) is a subgraph. Then (W, F) is a $\mathcal{L}_{gr}$-substructure of (V, E) if and only if it is an induced subgraph.

2 Formulas and Definable Sets

Our next task is to define a formal syntax for expressing properties of $\mathcal{L}$-structures using the symbols in $\mathcal{L}$. To motivate the definitions, we make the following observations.

Example 2.1. Consider the $\mathcal{L}_{or}$-structure $(\mathbb{R}, +, \cdot, <, 0, 1)$. There are many more functions and relations, which are not in $\mathcal{L}_{or}$, but are still expressible using the symbols in $\mathcal{L}_{or}$. For example:

1. the *unary* function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f(x) = x + 1$;
2. the *ternary* relation $R = \{(x, y, z) \in \mathbb{R}^3 : x < y + z\}$.

To address this issue, we formally define how to build new functions and relations from the symbols in a given language. In particular, we define $\mathcal{L}$*-terms* and $\mathcal{L}$*-formulas*, which will be certain special strings of symbols built from:

- the symbols in $\mathcal{L}$,
- the equality sign $=$ (to be interpreted as equality),
- countably many variable symbols: e.g. u, v, w, x, y, z, or v_i for $i \in \mathbb{N}$, etc...
- the Boolean connectives $\wedge$ and $\neg$ (to be interpreted as "and" and "not", respectively),
- the existential quantifier symbol $\exists$ (to be interpreted as "there exists", respectively),
- parentheses and commas (for parsing and listing).

We will later observe that several other "natural" logical operators are expressible using these symbols (see Remark 2.9).

2.1 Terms (new functions)

Definition 2.2. Let $\mathcal{L}$ be a language. The set of $\mathcal{L}$**-terms** is the smallest set $\mathcal{T}$ satisfying the following properties:

(*i*) $c \in \mathcal{T}$ for any constant symbol c in $\mathcal{L}$,

(*ii*) $v \in \mathcal{T}$ for each variable symbol v,

(*iii*) if f is an n-ary function symbol in $\mathcal{L}$, and $t_1, \ldots, t_n \in \mathcal{T}$, then $f(t_1, \ldots, t_n) \in \mathcal{T}$.

Returning to Example 2.1, we can now express the function $f(x) = x + 1$ as an $\mathcal{L}_{or}$-term. If we pedantically follow the full formality of the definition, then this term would be:

$$+(x, 1).$$

For the sake of better comprehension, we abuse notation and write this term as $x + 1$.

As suggested by Example 2.1, we will interpret $\mathcal{L}$-terms as functions on $\mathcal{L}$-structures.

Convention 2.3. In several places, it will be convenient to think of constant symbols as "function symbols of arity 0". To make sense of this, we use the convention $M^0 = \{\emptyset\}$ for any set M. Given a language $\mathcal{L}$, an $\mathcal{L}$-structure $\mathcal{M}$, and a constant symbol c in $\mathcal{L}$, we identify the interpretation $c^{\mathcal{M}}$ with the 0-ary function $\emptyset \mapsto c^{\mathcal{M}}$ from M^0 to M.

Definition 2.4. Fix a language $\mathcal{L}$. Let t be an $\mathcal{L}$-term and let $\mathcal{M}$ be an $\mathcal{L}$-structure. By induction on the construction of terms, we define a function $t^{\mathcal{M}} : M^n \longrightarrow M$, where n is the number of distinct variable symbols appearing in t.

(i) If t is a constant symbol c, then $t^{\mathcal{M}} : M^0 \longrightarrow M$ such that $t^{\mathcal{M}}(\emptyset) = c^{\mathcal{M}}$.

(ii) If t is a variable symbol, then $t^{\mathcal{M}} : M \longrightarrow M$ is the identity function.

(iii) Suppose f is an m-ary function symbol, and t is the $\mathcal{L}$-term $f(t_1, \ldots, t_m)$, where $t_1, \ldots, t_m$ are $\mathcal{L}$-terms using variables from among $v_1, \ldots, v_n$. Define $t^{\mathcal{M}} : M^n \longrightarrow M$ such that

$$t^{\mathcal{M}}(\bar{a}) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a}), \ldots, t_m^{\mathcal{M}}(\bar{a})),$$

where, for $1 \leq i \leq m$, $t_i^{\mathcal{M}}(\bar{a})$ denotes the function $t_i^{\mathcal{M}}$ evaluated on the subtuple of $\bar{a}$ corresponding to the variables used in t_i (which can be $\emptyset$ if t_i is a constant symbol).

2.2 Formulas (new relations)

Definition 2.5. Let $\mathcal{L}$ be a language.

1. An **atomic $\mathcal{L}$-formula** is a string φ of symbols of one of the following forms:
   - *(i)* $t_1 = t_2$, where t_1, t_2 are $\mathcal{L}$-terms, or
   - *(ii)* $R(t_1, \ldots, t_n)$, where R is an n-ary relation symbol in $\mathcal{L}$ and $t_1, \ldots, t_n$ are $\mathcal{L}$-terms.
2. The set of **$\mathcal{L}$-formulas** is the smallest set $\mathcal{F}$ satisfying the following properties:
   - *(i)* any atomic $\mathcal{L}$-formula is in $\mathcal{F}$,
   - *(ii)* if $\varphi \in \mathcal{F}$ then $\neg\varphi \in \mathcal{F}$,
   - *(iii)* if $\varphi, \psi \in \mathcal{F}$ then $(\varphi \wedge \psi) \in \mathcal{F}$,
   - *(iv)* if $\varphi \in \mathcal{F}$ and v is a variable symbol, then $\exists v(\varphi) \in \mathcal{F}$.

Returning to Example 2.1, we can express the relation R as the atomic $\mathcal{L}_{or}$-formula

$$<(x, +(y, z)).$$

Once again, for the sake of comprehension and readability, we instead write: $x < y + z$.

Definition 2.6. Given $\mathcal{L}$-formula φ, and a variable v used in φ, we say v **occurs freely** if v is does not occur in the scope of $\exists v$. If v does not occur freely in φ then we say v is **bound** in φ. If no variable occurs freely in φ then φ is an **$\mathcal{L}$-sentence**.

Remark 2.7. By renaming bound variables, we may assume that no variable v has both free and bound occurrences in the same formula. For example, if φ is the $\mathcal{L}_{or}$-formula $x < y$ and ψ is the $\mathcal{L}_{or}$-formula $\exists x(x + y = 0)$, we will write the conjunction $\varphi \wedge \psi$ as $(x < y) \wedge \exists z(z + y = 0)$.

We will write $\varphi(v_1, \dots, v_n)$ to emphasize that φ is an $\mathcal{L}$-formula with free variables $v_1, \dots, v_n$. We now define the interpretation $\mathcal{L}$-formulas as relations on $\mathcal{L}$-structures.

Definition 2.8. Let $\varphi(v_1, \dots, v_n)$ be an $\mathcal{L}$-formula.

1. Given $\bar{a} \in M^n$, we inductively define what it means for $\bar{a}$ to **satisfy** $\varphi(\bar{v})$ **in** $\mathcal{M}$, written $\mathcal{M} \models \varphi(\bar{a})$.

   (*i*) If $\varphi(v_1, \dots, v_n)$ is of the form $t_1 = t_2$ where t_1 and t_2 are $\mathcal{L}$-terms using variables among $v_1, \dots, v_n$, then
   $$\mathcal{M} \models \varphi(\bar{a}) \;\Leftrightarrow\; t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}).$$

   (*ii*) If $\varphi(v_1, \dots, v_n)$ is of the form $R(t_1, \dots, t_m)$ where R is an m-ary relation symbol and $t_1, \dots, t_m$ are $\mathcal{L}$-terms with variables among $v_1, \dots, v_n$ then
   $$\mathcal{M} \models \varphi(\bar{a}) \;\Leftrightarrow\; (t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}}.$$

   (*iii*) If $\varphi(v_1, \dots, v_n)$ is an $\mathcal{L}$-formula then
   $$\mathcal{M} \models \neg\varphi(\bar{a}) \;\Leftrightarrow\; \mathcal{M} \not\models \varphi(\bar{a}).$$

   (*iv*) If $\varphi(v_{i_1}, \dots, v_{i_r})$ and $\psi(v_{j_1}, \dots, v_{j_s})$ are $\mathcal{L}$-formulas, with $\{i_1, \dots, i_r, j_1, \dots, j_s\} = \{1, \dots, n\}$, then
   $$\mathcal{M} \models (\varphi \wedge \psi)(\bar{a}) \;\Leftrightarrow\; \mathcal{M} \models \varphi(a_{i_1}, \dots, a_{i_r}) \text{ and } \mathcal{M} \models \psi(a_{j_1}, \dots, a_{j_s}),$$

   (*v*) If $\varphi(v_1, \dots, v_n, w)$ is an $\mathcal{L}$-formula then
   $$\mathcal{M} \models (\exists w\varphi)(\bar{a}) \;\Leftrightarrow\; \text{there exists } b \in M \text{ such that } \mathcal{M} \models \varphi(\bar{a}, b).$$

2. Define the subset $\varphi^{\mathcal{M}} = \{\bar{a} \in M^n : \mathcal{M} \models \varphi(\bar{a})\}$.

The reader should think about the previous construction of $\varphi^{\mathcal{M}}$ in the case that the formula φ is a *sentence* with no free variables. We will discuss this further in Section 3.

Remark 2.9.

1. We will use the following abbreviations for the expression of other "logical notions".

   (*i*) *disjunction*: $\varphi \vee \psi$ ("φ or ψ") is an abbreviation for $\neg(\neg\varphi \wedge \neg\psi)$.

   (*ii*) *implication*: $\varphi \to \psi$ ("φ implies ψ") is an abbreviation for $\neg\varphi \vee \psi$.

   (*iii*) *equivalence*: $\varphi \leftrightarrow \psi$ ("φ if and only if ψ") is an abbreviation for $(\varphi \to \psi) \wedge (\psi \to \varphi)$.

   (*iv*) *universal quantification*: $\forall v(\varphi)$ ("for all v, φ") is an abbreviation for $\neg\exists v(\neg\varphi)$.

   The reader may verify that these abbreviations are coherent (see Exercise 7.2.5).

2. Depending on the particular language $\mathcal{L}$, one can define further abbreviations. For example, consider the language $\mathcal{L}_{or}$. We often drop the multiplication symbol, and write v_1v_2 for $v_1 \cdot v_2$. We can express the squaring function as the $\mathcal{L}_{or}$-term $v \cdot v$, which will be abbreviated as v^2. For example, the following formula expresses that every positive element has a square root:

$$\forall x(x > 0 \to \exists y(x = y^2)).$$

For another example, we can express the *ternary relation* $|x - y| < z$ as

$$(0 \leq x - y < z) \vee (0 \leq y - x < z),$$

where $v_1 \leq v_2 < v_3$ is an abbreviation for: $((v_1 = v_2) \vee (v_1 < v_2)) \wedge (v_2 < v_3)$.

Now consider an expanded language $\mathcal{L} = \mathcal{L}_{or} \cup \{f\}$, where f is a new unary function symbol. The following $\mathcal{L}$-formula expresses that the function f is continuous at x:

$$\forall v_1 \Big(v_1 > 0 \to \exists v_2 \Big(v_2 > 0 \wedge \forall y \big(|x - y| < v_2 \to |f(x) - f(y)| < v_1\big)\Big)\Big).$$

Recall that an $\mathcal{L}$-embedding between two structures is defined to preserve all symbols in $\mathcal{L}$. A natural question is the extent to which $\mathcal{L}$-embeddings preserve more complicated formulas.

Definition 2.10. Given a language $\mathcal{L}$, an $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n)$ is **quantifier-free** if it is constructed from atomic formulas using only iterations of $\neg$ and $\wedge$.

Proposition 2.11. *Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures, and $\sigma : \mathcal{M} \longrightarrow \mathcal{N}$ is an $\mathcal{L}$-embedding. For any quantifier-free formula $\varphi(v_1, \ldots, v_n)$ and $\bar{a} \in M^n$,*

$$\mathcal{M} \models \varphi(a_1, \ldots, a_n) \;\; \Leftrightarrow \;\; \mathcal{N} \models \varphi(\sigma(a_1), \ldots, \sigma(a_n)).$$

Proof. Given a tuple $\bar{a} \in M^n$, let $\sigma(\bar{a}) = (\sigma(a_1), \ldots, \sigma(a_n)) \in N^n$. We must first prove a claim concerning $\mathcal{L}$-terms.
Claim: If $t(\bar{v})$ is a term and $\bar{a} \in M^n$ then $\sigma(t^{\mathcal{M}}(\bar{a})) = t^{\mathcal{N}}(\sigma(\bar{a}))$.
Proof: We proceed by induction on the construction of terms. If t is a constant symbol c then $\sigma(c^{\mathcal{M}}) = c^{\mathcal{N}}$ since σ is an $\mathcal{L}$-embedding. If t is a variable and $a \in M$, then $\sigma(t^{\mathcal{M}}(a)) = \sigma(a) = t^{\mathcal{N}}(\sigma(a))$. Now suppose $t(v_1, \ldots, v_n)$ is of the form $f(t_1, \ldots, t_m)$, where f is an m-ary function symbol and $t_1, \ldots, t_m$ are terms, which satisfy the claim and use variables among $v_1, \ldots, v_n$. Then, for $\bar{a} \in M^n$,

$$\begin{aligned} \sigma(t^{\mathcal{M}}(\bar{a})) &= \sigma(f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a})), \ldots, f^{\mathcal{M}}(t_m^{\mathcal{M}}(\bar{a}))) \\ &= f^{\mathcal{N}}(\sigma(t_1^{\mathcal{M}}(\bar{a})), \ldots, \sigma(t_m^{\mathcal{M}}(\bar{a}))) \quad \text{(since } \sigma \text{ is an embedding)} \\ &= f^{\mathcal{N}}(t_1^{\mathcal{N}}(\sigma(\bar{a})), \ldots, t_m^{\mathcal{N}}(\sigma(\bar{a}))) \quad \text{(by induction)} \\ &= t^{\mathcal{N}}(\sigma(\bar{a})). \end{aligned} \qquad \dashv_{\text{claim}}$$

We now prove the proposition by induction on the construction of formulas. Suppose φ is the formula $R(t_1, \ldots, t_m)$, where R is an m-ary relation symbol and $t_1, \ldots, t_m$ are terms. Then

$$\begin{aligned} \mathcal{M} \models \varphi(\bar{a}) \;\; &\Leftrightarrow \;\; (t_1^{\mathcal{M}}(\bar{a}), \ldots, t_m^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}} \\ &\Leftrightarrow \;\; (\sigma(t_1^{\mathcal{M}}(\bar{a})), \ldots, \sigma(t_m^{\mathcal{M}}(\bar{a}))) \in R^{\mathcal{N}} \quad \text{(since } \sigma \text{ is an embedding)} \\ &\Leftrightarrow \;\; (t_1^{\mathcal{N}}(\sigma(\bar{a})), \ldots, t_m^{\mathcal{N}}(\sigma(\bar{a}))) \in R^{\mathcal{N}} \quad \text{(by the claim)} \\ &\Leftrightarrow \;\; \mathcal{N} \models \varphi(\sigma(\bar{a})). \end{aligned}$$

Viewing equality as a binary relation, the same argument works when φ is the formula $t_1 = t_2$ (this uses injectivity of σ). This proves the result for atomic formulas.

Assume the result for $\varphi(\bar{v})$. Then

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

where the second equivalence is by induction. We leave it to the reader to finish the $\varphi \wedge \psi$ case. $\square$

In general, the quantifier-free assumption in the previous result is necessary (see Exercise 7.2.1). We will consider preservation of arbitrary formulas in Section 5.

Given an $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n)$, the subset $\varphi^{\mathcal{M}} \subseteq M^n$ is a particular case of the more general notion of a *definable set* in the structure $\mathcal{M}$.

Definition 2.12. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Given $n > 0$, a subset $X \subseteq M^n$ is **definable in $\mathcal{M}$** if there is an $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n, w_1, \ldots, w_m)$ and a tuple $\bar{b} \in M^m$ such that

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

In the above definition, the elements in $\bar{b}$ are referred to as *parameters* and one can treat $\varphi(\bar{v}, \bar{b})$ as an *$\mathcal{L}$-formula with parameters from M*. Alternatively, $\varphi(\bar{v}, \bar{b})$ can be viewed as a formula in the language $\mathcal{L}_M$ obtained from $\mathcal{L}$ by adding constant symbols for all elements of M (and interpreting those symbols in the obvious way). For our purposes, we treat these viewpoints as equivalent (although there are areas of model theory where the distinction is crucial).

Let $\mathcal{L}$ be a language and $\mathcal{M}$ an $\mathcal{L}$-structure. Suppose $X \subseteq M^n$ is definable in $\mathcal{M}$. Then it is possible that there is more than one formula which defines X. We give a few examples.

Example 2.13. Consider the $\mathcal{L}_{or}$-structure $(\mathbb{R}, +, -, \cdot, <, 0, 1)$.

1. Let $X = \left\{\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right\}$. Then X is defined by $\varphi\left(x, \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right)$, where $\varphi(x, v_1, v_2)$ is the formula

   $$x = v_1 \vee x = v_2,$$

   and also by the formula: $x^2 - x - 1 = 0$ (with no extra parameters).

2. Let $X \subseteq \mathbb{R}^3$ be the set of triples (a, b, c) such that the quadratic function $ax^2 + bx + c$ has a root in $\mathbb{R}$. Then X is defined by

   $$\exists x(ux^2 + vx + w = 0),$$

   and by

   $$v^2 - 4uw \geq 0 \wedge \neg(u = 0 \wedge v = 0 \wedge w \neq 0).$$

   Note that both formulas are in free variables u, v, w, but the second is quantifier-free while the first is not.

3. We can use definable sets to see that, for this particular structure, some symbols in $\mathcal{L}_{or}$ are redundant. For example the graph of the binary function "$-$" can be defined using "$+$", since $z = x - y$ if and only if $x = y + z$ (see also Exercise 7.2.3). Moreover, as subset of $\mathbb{R}^2$, the binary relation $<$ is definable by the $\mathcal{L}_r$-formula

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

   However, as we will discuss in Section 5, the fact that a quantifier is necessary in the definition of $<$ is extremely significant.

3 Sentences and Theories

Given a formula $\varphi(v_1, \ldots, v_n)$, Definition 2.8 produces a subset $\varphi^{\mathcal{M}} \subseteq M^n$, which contains the tuples $\bar{a} \in M^n$ for which $\mathcal{M} \models \varphi(\bar{a})$. If φ has no free variables, then $\varphi^{\mathcal{M}}$ is a subset of M^0, and is therefore either equal to M^0 or $\emptyset$. Working carefully through Definition 2.8, we see that $\varphi^{\mathcal{M}} = M^0$ if and only if $\mathcal{M} \models \varphi$ (see Exercise 7.3.4). In this case, we think of φ as expressing a "true statement" about the structure $\mathcal{M}$.

Definition 3.1. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Define the **theory of** $\mathcal{M}$ to be

$$\mathrm{Th}(\mathcal{M}) = \{\varphi : \varphi \text{ is an } \mathcal{L}\text{-sentence and } \mathcal{M} \models \varphi\}.$$

There is an extremely important observation to be made at this point having to do with our use of quantifiers in $\mathcal{L}$-sentences. In particular, quantifiers range only over *elements* of structures, and not more complicated objects (e.g. *subsets* of structures). This limitation is specified by saying that $\mathcal{L}$-sentences, as we have defined them, are *first-order*. In fact, one should technically apply the adjective *first-order* to many of the previously defined notions (e.g. first-order $\mathcal{L}$-formulas and first-order definable subsets of $\mathcal{L}$-structures). In general, everything done here is regarded as under the umbrella of *first-order logic*.

For an example to emphasize this distinction, consider $\mathcal{L}_o$-structure $(\mathbb{R}, <)$. A very important feature about this structure is the *least upper bound property*: any nonempty subset of $\mathbb{R}$ with an upper bound in $\mathbb{R}$ contains a least upper bound in $\mathbb{R}$. If we try to express this property as a *first-order* $\mathcal{L}_o$-sentence, we run into trouble because it requires quantification over *subsets* of $\mathbb{R}$. In fact, there is no way to express the least upper bound property as a first-order $\mathcal{L}$-sentence in any language $\mathcal{L}$ (see Exercise 7.5.3).

Suppose now that $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures such that $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N})$. To what extent are $\mathcal{M}$ and $\mathcal{N}$ alike? This question motivates the next definition.

Definition 3.2. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. Then $\mathcal{M}$ and $\mathcal{N}$ are **elementarily equivalent**, written $\mathcal{M} \equiv \mathcal{N}$ if $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N})$.

Now we can rephrase the question above as a problem: given an $\mathcal{L}$-structure $\mathcal{M}$, classify the $\mathcal{L}$-structures elementarily equivalent to $\mathcal{M}$. This question has motivated much of modern model theory, and has led to deep advances in mathematics.

We begin with an unsurprising example of elementary equivalence.

Proposition 3.3. *Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure. For any $\mathcal{L}$-structure $\mathcal{N}$, if $\mathcal{M} \cong \mathcal{N}$ then $\mathcal{M} \equiv \mathcal{N}$.*

Proof. Apply Exercise 7.2.6. □

The converse of this fact only holds if $\mathcal{M}$ is finite (see Exercise 7.3.6). In particular, if $\mathcal{M}$ is an infinite $\mathcal{L}$-structure, then we will see that there are $\mathcal{L}$-structures of arbitrarily large cardinality elementarily equivalent to $\mathcal{M}$ (see Proposition 4.5). We first need to develop tools for working with theories of structures. For these tools to be the most useful, we want to consider theories in greater generality.

Definition 3.4. Let $\mathcal{L}$ be a language.

1. An **$\mathcal{L}$-theory** is a set T of $\mathcal{L}$-sentences.

2. Given an $\mathcal{L}$-theory T and an $\mathcal{L}$-structure $\mathcal{M}$, we say $\mathcal{M}$ is a **model of** T, written $\mathcal{M} \models T$, if $\mathcal{M} \models \varphi$ for all $\mathcal{L}$-sentences φ in T.

3. An $\mathcal{L}$-theory T is **satisfiable** if it has a model.

4. Given an $\mathcal{L}$-theory T, let $\mathrm{Mod}(T)$ be the class of models of T.

5. Given an $\mathcal{L}$-theory T and an $\mathcal{L}$-sentence φ, we say φ is a **logical consequence of** φ, written $T \models \varphi$, if $\mathcal{M} \models \varphi$ for any model $\mathcal{M}$ of T.

We can use theories to describe or "axiomatize" certain classes of structures we want to study.

Definition 3.5. Let $\mathcal{L}$ be a language. A class $\mathcal{K}$ of $\mathcal{L}$-structures is an **elementary class** if there is an $\mathcal{L}$-theory T such that $\mathcal{K} = \mathrm{Mod}(T)$.

Remark 3.6. To avoid inconsequential complications, we often tacitly assume that classes $\mathcal{K}$ of $\mathcal{L}$-structures are closed under isomorphism.

It is worth going through several examples of elementary classes.

Example 3.7.

1. Consider the language $\mathcal{L}_g$ of groups. Let G consist of the following sentences

$$\forall x \forall y \forall z ((x * y) * z = x * (y * z))$$
$$\forall x (x * e = x = e * x)$$
$$\forall x \exists y (x * y = e = y * x)$$

   Then the class of models of G is precisely the class of groups, and so the class of groups is an elementary class. We also say that G *axiomatizes the theory of groups.*

   Let AG be G together with the $\mathcal{L}_g$-sentence

$$\forall x \forall y (x * y = y * x).$$

   Then AG axiomatizes the theory of abelian groups.

   Let DAG be $\mathrm{AG} \cup \{\varphi_n : n > 0\}$, where φ_n is the $\mathcal{L}_g$-sentence

$$\forall x \exists y (y^n = x),$$

   and y^n is an abbreviation for $y * y * \ldots * y$ (n times). Then DAG axiomatizes the theory of *divisible abelian groups.*

   Let $\mathrm{TFDAG} = \mathrm{DAG} \cup \{\exists x (x \neq e)\} \cup \{\psi_n : n > 0\}$, where ψ_n is the $\mathcal{L}_g$-sentence

$$\forall x (x^n = e \rightarrow x = e).$$

   Then TFDAG axiomatizes the theory of *nontrivial torsion-free divisible abelian groups.*

2. Let $\mathcal{L}_{gr} = \{E\}$ be the language of graphs. Then the class of graphs is an elementary class, whose theory is axiomatized by

$$\forall x (\neg E(x, x)) \wedge \forall x \forall y (E(x, y) \rightarrow E(y, x)).$$

3. *Vector spaces over a field.* Fix a field F and define a language $\mathcal{L} = \{+, 0, (\lambda_a)_{a \in F}\}$, where $+$ is a binary relation symbol, 0 is a constant symbol, and, for $a \in F$, λ_a is a unary function symbol. Define the theory T consisting of AG (in the language $\{+, 0\}$), together with

- for every $a, b \in F$,
$$\forall x\big(\lambda_a(\lambda_b(x)) = \lambda_{ab}(x)\big),$$
- for every $a \in F$,
$$\forall x \forall y\big(\lambda_a(x+y) = \lambda_a(x) + \lambda_a(y)\big),$$
- for every $a, b \in F$,
$$\forall x\big(\lambda_{a+b}(x) = \lambda_a(x) + \lambda_b(x)\big),$$
- $\forall x\big(\lambda_1(x) = x\big)$.

Then T axiomatizes vector spaces over F.

4 The Compactness Theorem

In the last section, we considered several examples of elementary classes. A more difficult problem is to show that certain classes of structures are *not* elementary classes. In particular, given a class $\mathcal{K}$ of $\mathcal{L}$-structures, in order to show that $\mathcal{K}$ is not elementary class one needs to show that $\mathcal{K} \neq \mathrm{Mod}(T)$ for any $\mathcal{L}$-theory T. One way to accomplish this would be to isolate a collection of sentences Δ such that no structure in $\mathcal{K}$ is a model of Δ and then show that $T \cup \Delta$ is satisfiable for any $\mathcal{L}$-theory T such that $\mathcal{K} \subseteq \mathrm{Mod}(T)$. To do this, one often takes Δ to be sentences in some larger language (see Proposition 4.4 below). But in any case, we need general tools for proving satisfiability of theories. This brings us to the Compactness Theorem, which is the cornerstone result lying at the foundation of all of first-order model theory.

Definition 4.1. An $\mathcal{L}$-theory T is **finitely satisfiable** if every finite subset of T is satisfiable.

Theorem 4.2 (The Compactness Theorem)**.** *Every finitely satisfiable $\mathcal{L}$-theory is satisfiable.*

A proof of this result is given in Appendix A. The power and use of the Compactness Theorem cannot be understated; it is used in every facet of first-order model theory. As previously discussed, applications of the Compactness Theorem often involve moving to a larger language (e.g. by adding new constant symbols). Therefore we make the following definition.

Definition 4.3. Let $\mathcal{L}$ and $\mathcal{L}^*$ be languages with $\mathcal{L} \subseteq \mathcal{L}^*$, and suppose $\mathcal{M}^*$ is an $\mathcal{L}^*$-structure. We define the **reduct of $\mathcal{M}^*$ to $\mathcal{L}^*$**, denoted $\mathcal{M}^*|_{\mathcal{L}}$, to be the unique $\mathcal{L}$-structure $\mathcal{M}$ satisfying the following properties:

(*i*) the universe of $\mathcal{M}$ is the universe of $\mathcal{M}^*$, and

(*ii*) the interpretation in $\mathcal{M}$ of any symbol in $\mathcal{L}$ is the same as the interpretation in $\mathcal{M}^*$.

In this case, we also call $\mathcal{M}^*$ an **expansion of $\mathcal{M}$ to $\mathcal{L}^*$**.

The following is our first application of the Compactness Theorem.

Proposition 4.4. *Suppose T is an $\mathcal{L}$-theory with arbitrarily large finite models. Then T has an infinite model.*

Proof. First, we expand the language $\mathcal{L}^* = \mathcal{L} \cup \{c_n : n > 0\}$, where each c_n is a new constant symbol. Note that T is still an $\mathcal{L}^*$-theory. Define the set of $\mathcal{L}^*$-sentences

$$\Delta = \{c_m \neq c_n : m, n > 0,\ m \neq n\}.$$

Set $T^* = T \cup \Delta$ and fix a finite subset $T_0 \subseteq T^*$. Then there is an integer $k > 0$ such that

$$T_0 \subseteq T \cup \{c_m \neq c_n : 0 < m, n \leq k,\ m \neq n\}.$$

By assumption, there is an $\mathcal{L}$-structure $\mathcal{M} \models T$ such that $|M| \geq k$. Let $\mathcal{M}^*$ be the $\mathcal{L}^*$-structure, with universe M, such that

- $\mathcal{M}^*|_{\mathcal{L}} = \mathcal{M}$,
- $c_1^{\mathcal{M}^*}, \ldots, c_k^{\mathcal{M}^*}$ are distinct elements of M, and
- $c_n^{\mathcal{M}^*}$, for $n > k$, is any arbitrarily element of M.

Then $\mathcal{M}^* \models T_0$ by construction.

By the Compactness Theorem, T^* is satisfiable and so we may fix an $\mathcal{L}^*$-structure $\mathcal{N}^* \models T^*$. Clearly, the universe N of $\mathcal{N}^*$ must be infinite. Moreover $\mathcal{N} = \mathcal{N}^*|_{\mathcal{L}}$ is a model of T. □

From this result, we see that if $\mathcal{K}$ is a class of finite $\mathcal{L}$-structures, containing elements of arbitrarily large finite size, then $\mathcal{K}$ is not an elementary class. For example, the classes of finite sets, finite groups, finite graphs, and finite fields (etc...) are *not* elementary.

The following application is of a similar flavor, and is proved using a strengthening of the Compactness Theorem (see Exercise 7.4.5).

Proposition 4.5. *If T is an $\mathcal{L}$-theory with infinite models, and $\kappa \geq \max\{|\mathcal{L}|, \aleph_0\}$, then T has a model of cardinality κ.*

Next, we consider the class of torsion groups (i.e groups in which every element has finite order).

Proposition 4.6. *Let T be an $\mathcal{L}$-theory, where $\mathcal{L}$ contains $\mathcal{L}_g$. Then $\mathcal{K} := \{\mathcal{M}|_{\mathcal{L}_g} : \mathcal{M} \models T\}$ is not the class of torsion groups.*

Proof. Suppose, for a contradiction, that $\mathcal{K}$ is the class of torsion groups. Let $\mathcal{L}^* = \mathcal{L} \cup \{c\}$, where c is a new constant symbol. Define the set of $\mathcal{L}^*$-sentences:

$$\Delta = \{c^n \neq e : n > 0\}.$$

Let $T^* = T \cup \Delta$, and suppose $T_0 \subseteq T^*$ is finite. We may fix an integer $k > 0$ such that

$$T_0 \subseteq T \cup \{c^n \neq e : 0 < n < k\}.$$

By assumption there is a model $\mathcal{N} \models T$ such that $\mathcal{N}|_{\mathcal{L}_g} = \left(\mathbb{Z}/k\mathbb{Z}, +_k, 0\right)$, where $+_k$ is addition modulo k. Let $\mathcal{N}^*$ be the expansion of $\mathcal{N}$ to $\mathcal{L}^*$ by interpreting c as 1. Then $\mathcal{N}^*$ models T_0.

By the Compactness Theorem, there is an $\mathcal{L}^*$-structure $\mathcal{M}^* \models T^*$. Then $\mathcal{M} = \mathcal{M}^*|_{\mathcal{L}}$ is a model of T, and so $\mathcal{M}|_{\mathcal{L}_g} \in \mathcal{K}$. But the interpretation of c in $\mathcal{M}^*$ witnesses that $\mathcal{M}|_{\mathcal{L}_g}$ is not a torsion group, which is a contradiction. □

See Exercise 7.4.6 for an interesting refinement of the previous result.

5 Elementary Extensions

Recall that in Section 1 we defined the notion of $\mathcal{L}$-embeddings between $\mathcal{L}$-structures, which are simply injective functions preserving the symbols in $\mathcal{L}$. Using basic logic, this preservation automatically extends to quantifier-free $\mathcal{L}$-formulas (see Proposition 2.11). This motivates the following definition.

Definition 5.1. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures.

1. An $\mathcal{L}$-embedding $\sigma : \mathcal{M} \longrightarrow \mathcal{N}$ is **elementary** if, for any $\mathcal{L}$-formula $\varphi(v_1, \dots, v_n)$ and any $\bar{a} \in M^n$,
$$\mathcal{M} \models \varphi(a_1, \dots, a_n) \;\Leftrightarrow\; \mathcal{N} \models \varphi(\sigma(a_1), \dots, \sigma(a_n)).$$

2. $\mathcal{M}$ is an **elementary substructure** of $\mathcal{N}$, written $\mathcal{M} \prec \mathcal{N}$, if $M \subseteq N$ and the inclusion map from M to N is an elementary $\mathcal{L}$-embedding. In this case, we also say $\mathcal{N}$ is an **elementary extension** of $\mathcal{M}$.

Remark 5.2. In the definition of elementary embeddings, we implicitly allow φ to be an $\mathcal{L}$-sentence, in which case the definition just says $\mathcal{M} \models \varphi$ if and only if $\mathcal{N} \models \varphi$. Therefore, if $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures, and there is an elementary embedding from $\mathcal{M}$ to $\mathcal{N}$, then $\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent.

The distinction between substructures and elementary substructures is important (see Exercises 7.5.5 and 7.5.7). However, we will eventually see examples where the two notions are the same.

Definition 5.3. A theory T is **model complete** if, for any models $\mathcal{M}, \mathcal{N}$ of T, if $\mathcal{M}$ is a substructure of $\mathcal{N}$ then $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$.

In previous sections, we used the Compactness Theorem to build models of theories. By taking a little extra care, we can refine these methods to build elementary extensions of structures.

Definition 5.4. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Let $\mathcal{L}_M = \mathcal{L} \cup \{\tilde{m} : m \in M\}$, where each $\tilde{m}$ is a constant symbol. We interpret $\mathcal{M}$ as an $\mathcal{L}_M$-structure by setting $\tilde{m}^{\mathcal{M}} = m$.

1. Given an $\mathcal{L}$-formula $\varphi(v_1, \dots, v_n)$ and $m_1, \dots, m_n \in M$, we let $\varphi(\tilde{m}_1, \dots, \tilde{m}_n)$ denote the $\mathcal{L}_M$-sentence obtained by replacing each free occurrence of v_i in $\varphi(\bar{v})$ with the constant $\tilde{m}_i$.

2. The **diagram of** $\mathcal{M}$ is the following set of $\mathcal{L}_M$-sentences:
$$\mathrm{Diag}(\mathcal{M}) = \{\varphi(\tilde{m}_1, \dots, \tilde{m}_n) : \varphi(\bar{v}) \text{ is a quantifier-free } \mathcal{L}\text{-formula and } \mathcal{M} \models \varphi(m_1, \dots, m_n)\}.$$

3. The **elementary diagram of** $\mathcal{M}$ is the following set of $\mathcal{L}_M$-sentences:
$$\mathrm{Diag}_{\mathrm{el}}(\mathcal{M}) = \{\varphi(\tilde{m}_1, \dots, \tilde{m}_n) : \varphi(\bar{v}) \text{ is an } \mathcal{L}\text{-formula and } \mathcal{M} \models \varphi(m_1, \dots, m_n)\}.$$

Remark 5.5. Similar to before, we allow φ to be an $\mathcal{L}$-sentence in the definition of $\mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$, and so $\mathrm{Th}(\mathcal{M}) \subseteq \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$.

Proposition 5.6. *Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure and $\mathcal{N}^*$ is an $\mathcal{L}_M$-structure. Let $\mathcal{N} = \mathcal{N}^*|_{\mathcal{L}}$.*

(a) If $\mathcal{N}^* \models \mathrm{Diag}(\mathcal{M})$ then there is an $\mathcal{L}$-embedding from $\mathcal{M}$ to $\mathcal{N}$.

(b) If $\mathcal{N}^* \models \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$ then there is an elementary $\mathcal{L}$-embedding from $\mathcal{M}$ to $\mathcal{N}$.

Proof. Suppose $\mathcal{N}^* \models \mathrm{Diag}(\mathcal{M})$. Define the function $\sigma : M \longrightarrow N$ such that $\sigma(m) = \tilde{m}^{\mathcal{N}^*}$. Then σ is an $\mathcal{L}$-embedding from $\mathcal{M}$ to $\mathcal{N}$. If $\mathcal{N}^* \models \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$ then σ is elementary. Details are left to the reader (see Exercise 7.5.6). □

The primary use of elementary diagrams to build elementary extensions and substructures is summarized by the *Löwenheim-Skolem Theorems.*

Theorem 5.7. *Let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure.*

(a) (Upward Löwenhein-Skolem Theorem) *Given an infinite cardinal κ, with $\kappa \geq \max\{|M|, |\mathcal{L}|\}$, there is an elementary extension $\mathcal{N} \succ \mathcal{M}$ such that $|N| = \kappa$.*

(b) (Downward Löwenheim-Skolem Theorem) *Given $X \subseteq M$, there is an elementary substructure $\mathcal{N} \prec \mathcal{M}$ such that $X \subseteq N$ and $|N| \leq \max\{|X|, |\mathcal{L}|, \aleph_0\}$.*

The proof is given in Section A.2 of Appendix A. Part (a) uses a strengthening of the Compactness Theorem (see Theorem A.1), while part (b) requires a bit more technology.

We can use diagrams to prove the following result in group theory.

Theorem 5.8 (Levi 1942)**.** *Every torsion-free abelian group can be totally ordered.*

Proof. Let $\mathcal{M} = (M, +, 0)$ be a torsion-free abelian group. Let $\mathcal{L} = \mathcal{L}_{og}$, and define

$$T = \mathrm{Diag}(\mathcal{M}) \cup T_0,$$

where T_0 is a set of $\mathcal{L}$-sentences expressing that $<$ is a group ordering.

Fix a finite subset $\Delta \subseteq \mathrm{Diag}(\mathcal{M})$. Let $X \subseteq M$ be the finite subset of M consisting of elements m such that $\tilde{m}$ appears in some $\mathcal{L}_M$-sentence in Δ. Let M_0 be the subgroup of M generated by X, and let $\mathcal{M}_0 = (M_0, +, 0)$. Then $\mathcal{M}_0$ is a substructure of $\mathcal{M}$ and so $\Delta \subseteq \mathrm{Diag}(\mathcal{M}_0)$ by Proposition 2.11. Moreover, $\mathcal{M}_0$ is a finitely generated torsion-free abelian group, and therefore isomorphic to $\mathbb{Z}^n$ for some $n > 0$. Therefore we can expand $\mathcal{M}_0$ to an $\mathcal{L}_M$-structure $\mathcal{M}_0^*$ by interpreting $<$ as the lexicographic order, and we have $\mathcal{M}_0^* \models \Delta \cup T_0$. Altogether, we have shown that T is finitely satisfiable. By the Compactness Theorem, there is an $\mathcal{L}_M$-structure $\mathcal{N}^* \models T$. By Proposition 5.6, there is an $\mathcal{L}$-embedding from $\mathcal{M}$ to $\mathcal{N} = \mathcal{N}^*|_{\mathcal{L}}$. In particular, $\mathcal{M}$ is isomorphic to a subgroup of the ordered group $\mathcal{N}$, and therefore inherits the ordering of $\mathcal{N}$. □

6 Quantifier Elimination

Definition 6.1. An $\mathcal{L}$-theory T has **quantifier elimination** if, for any formula $\varphi(v_1, \ldots, v_n)$ (with $n \geq 1$) there is a quantifier-free $\mathcal{L}$-formula $\psi(v_1, \ldots, v_n)$ such that

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

Remark 6.2. A useful feature of quantifier elimination is that the quantifier-free formula $\psi(\bar{v})$ is assumed to be in the *same free variables* as the formula $\varphi(\bar{v})$. This can become an issue if there are no quantifier-free $\mathcal{L}$-sentences (i.e. if $\mathcal{L}$ has no constant symbols), and it is for this reason that we emphasize $n \geq 1$ in the previous definition.

However, if T has quantifier elimination and φ is a sentence then, applying the definition with the formula $\varphi \wedge (v = v)$, we obtain a quantifier-free formula $\psi(v)$, in one free variable, such that

$$T \models \forall v\big(\varphi \leftrightarrow \psi(v)\big).$$

On the other hand, if $\mathcal{L}$ has at least one constant symbol, then there is in fact a quantifier-free *sentence* ψ such that $T \models \varphi \leftrightarrow \psi$ (see Exercise 7.6.5).

Quantifier elimination can be viewed as a strengthening of model completeness.

Proposition 6.3. *If T has quantifier elimination then it is model complete.*

Proof. See Exercise 7.6.4. $\square$

Example 6.4. In Example 2.13, we saw an instance of eliminating quantifiers in a single formula. In particular, if $T = \mathrm{Th}(\mathbb{R}, +, -, \cdot, <, 0, 1)$ and $\varphi(u, v, w)$ is the $\mathcal{L}$-formula $\exists x(ux^2 + vx + w = 0)$ then

$$T \models \forall u \forall v \forall w \big(\varphi(u, v, w) \leftrightarrow (v^2 - 4uw \geq 0 \wedge \neg(u = 0 \wedge v = 0 \wedge w \neq 0)\big).$$

In fact, T has quantifier elimination (see, e.g., Marker's text). On the other hand, $\mathrm{Th}(\mathbb{R}, +, -, \cdot, 0, 1)$ does not have quantifier elimination, but is model complete.

Definition 6.5. An $\mathcal{L}$-theory T is **complete** if, for any sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$.

Theorem 6.6. *An $\mathcal{L}$-theory T has quantifier elimination if and only if for any $\mathcal{M} \models T$ and any finitely generated $\mathcal{M}_0 \subseteq \mathcal{M}$, $T \cup \mathrm{Diag}(\mathcal{M}_0)$ is a complete $\mathcal{L}_{M_0}$-theory.*

Proof. The left-to-right direction is Exercise 7.6.6. Assume $T \cup \mathrm{Diag}(\mathcal{M}_0)$ is complete for any $\mathcal{M} \models T$ and $\mathcal{M}_0 \subseteq \mathcal{M}$. Fix an $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n)$, with $n \geq 1$. Define $\Gamma(\bar{v})$ to be the collection of quantifier-free $\mathcal{L}$-formulas $\psi(\bar{v})$ such that $T \models \forall \bar{v}(\varphi(\bar{v}) \to \psi(\bar{v}))$. Note that $\Gamma(\bar{v})$ is closed under conjunctions. Let $\mathcal{L}^* = \mathcal{L} \cup \{c_1, \ldots, c_n\}$, where $c_1, \ldots, c_n$ are new constant symbols.
Claim: $T \cup \Gamma(\bar{c}) \models \varphi(\bar{c})$.
Proof: Fix $\mathcal{M} \models T \cup \Gamma(\bar{c})$. Set $\bar{m} = \bar{c}^{\mathcal{M}}$ and let $\mathcal{M}_0$ be the substructure of $\mathcal{M}$ generated by $\bar{m}$. Let $\widetilde{m} = (\widetilde{m}_1, \ldots, \widetilde{m}_n)$. We want to show $\mathcal{M} \models \varphi(\widetilde{m})$. Since $\mathcal{M} \models T \cup \mathrm{Diag}(\mathcal{M}_0)$ (by Proposition 2.11), it suffices to show $T \cup \mathrm{Diag}(\mathcal{M}_0) \models \varphi(\widetilde{m})$. Suppose not. Since $T \cup \mathrm{Diag}(\mathcal{M}_0)$ is complete, we have $T \cup \mathrm{Diag}(\mathcal{M}_0) \models \neg\varphi(\widetilde{m})$. By the Compactness Theorem, we may fix a finite subset $\Delta \subseteq \mathrm{Diag}(\mathcal{M}_0)$ such that $T \cup \Delta \models \neg\varphi(\widetilde{m})$. By Exercise 7.3.3, we may assume that the formulas in Δ only use the extra constants in $\widetilde{m}$. Let $\psi(\bar{v})$ be an $\mathcal{L}$-formula such that $\psi(\widetilde{m})$ is the conjunction of the $\mathcal{L}_{M_0}$-sentences in Δ. Then $T \models \forall \bar{v}(\varphi(\bar{v}) \to \neg\psi(\bar{v}))$, and so $\neg\psi(\bar{v}) \in \Gamma(\bar{v})$. By assumption, $\mathcal{M} \models \neg\psi(\widetilde{m})$. But $\psi(\widetilde{m}) \in \mathrm{Diag}(\mathcal{M}_0) \subseteq \mathrm{Diag}(\mathcal{M})$, which is a contradiction. $\dashv_{\text{claim}}$

By the claim and the Compactness Theorem, there is a finite subset $\Delta \subseteq \Gamma(\bar{c})$ such that $T \cup \Delta \models \varphi(\bar{c})$. Let $\psi(\bar{v})$ be an $\mathcal{L}$-formula such that $\psi(\bar{c})$ is the conjunction of the $\mathcal{L}^*$-sentences in Δ. Then $T \models \forall \bar{v}(\psi(\bar{v}) \to \varphi(\bar{v}))$. Since $\psi(\bar{v}) \in \Gamma(\bar{v})$, we altogether have $T \models \forall \bar{v}(\varphi(\bar{v}) \leftrightarrow \psi(\bar{v}))$. $\square$

Next, we give a standard tool for demonstrating completeness of a theory.

Proposition 6.7 (Vaught's Test)**.** *Let T be an $\mathcal{L}$-theory with no finite models. Suppose there is some $\kappa \geq \max\{|\mathcal{L}|, \aleph_0\}$ such that all models of T of size κ are elementarily equivalent. Then T is complete.*

Proof. For a contradiction, suppose T is not complete. Then there is a sentence φ such that $T_1 = T \cup \{\varphi\}$ and $T_2 = T \cup \{\neg\varphi\}$ are both satisfiable. Since T has no finite models, it follows that T_1 and T_2 have infinite models. By Proposition 4.5, we may fix $\mathcal{M}_i \models T_i$ such that $\mathcal{M}_i$ has cardinality κ. Then $\mathcal{M}_1$ and $\mathcal{M}_2$ are not elementarily equivalent, which is a contradiction. $\square$

The rest of this section focuses on quantifier elimination for the theory of algebraically closed fields, along with several applications.

Definition 6.8. Let ACF be the $\mathcal{L}_r$-theory consisting of axioms for fields along with, for any $n > 0$, the sentence: $\forall v_0 \ldots \forall v_{n-1} \exists x(x^n + v_{n-1}x^{n-1} + \ldots + v_1 x + v_0 = 0)$.

Theorem 6.9. ACF *has quantifier elimination.*

Proof. By Theorem 6.6, it suffices to show that if F is a finitely generated integral domain, then $\mathrm{ACF} \cup \mathrm{Diag}(F)$ is complete. Let $\mathcal{L} = \mathcal{L}_r \cup \{\tilde{c} : c \in F\}$, and fix models K_1 and K_2 of $\mathrm{ACF} \cup \mathrm{Diag}(F)$ of size $\aleph_1$. We show K_1 and K_2 are isomorphic (as $\mathcal{L}$-structures), and hence elementarily equivalent by Proposition 3.3. Completness of $\mathrm{ACF} \cup \mathrm{Diag}(F)$ will then follow from Vaught's Test.

For $i \in \{1, 2\}$, we set $F_i = \{\tilde{c}^{K_i} : c \in F\}$. By (the proof of) Proposition 5.6, F_i is a subring of K_i isomorphic to F and, moreover, the function $\tau : F_1 \longrightarrow F_2$ such that $\tau(\tilde{c}^{K_1}) = \tau(\tilde{c}^{K_2})$ is a ring isomorphism. Let $E_i \subseteq K_i$ be the field of fractions of F_i. Then τ extends to a unique field isomorphism $\sigma : E_1 \longrightarrow E_2$. Given $i \in \{1, 2\}$, E_i has finite transcendence degree (since F_i is finitely generated). If $X \subseteq K_i$ is countable, then $E_i \cup X$ can only generate a countable algebraically closed subfield of K_i. It follows that the transcendence degree of K_i over E_i is $\aleph_1$. By Fact B.51, σ extends to a field isomorphism $\hat{\sigma} : K_1 \longrightarrow K_2$. By construction, $\hat{\sigma}$ is an $\mathcal{L}$-isomorphism. □

Note that ACF is not a complete theory since algebraically closed fields of different characteristics are not elementarily equivalent.

Definition 6.10. Given $n > 0$, let φ_n denote the $\mathcal{L}_r$-sentence $0 = 1 + 1 + \ldots + 1$ (n times).

1. Let $\mathrm{ACF}_0 = \mathrm{ACF} \cup \{\neg\varphi_n : n > 0\}$.
2. Given a prime p, let $\mathrm{ACF}_p = \mathrm{ACF} \cup \{\neg\varphi_n : 0 < n < p\} \cup \{\varphi_p\}$.

Note that $\mathrm{ACF}_p \models \mathrm{ACF}$ for any p (prime or 0), and so ACF_p also has quantifier elimination. Using Vaught's Test (as in the proof of Theorem 6.9), we also see that ACF_p is complete.

Definition 6.11. Let F be a field. Given integers $m, n > 0$, a *polynomial map from F^m to F^n* is a function of the form

$$\Phi(\bar{x}) = (p_1(\bar{x}), \ldots, p_n(\bar{x})),$$

where $p_i(\bar{x}) \in F[\bar{x}]$ and $\bar{x} = (x_1, \ldots, x_m)$.

Theorem 6.12 (Ax's Theorem)**.** *Fix $n > 0$. If $\Phi : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ is an injective polynomial map, then Φ is surjective.*

We first prove the analogous result for the algebraic closure $\mathbb{F}_p^{alg}$ of $\mathbb{F}_p$, where p is a prime.

Lemma 6.13. *Fix a prime p and an integer $n > 0$. If $\Phi : (\mathbb{F}_p^{alg})^n \longrightarrow (\mathbb{F}_p^{alg})^n$ is an injective polynomial map, then Φ is surjective.*

Proof. Let $F = \mathbb{F}_p^{alg}$ and, for $m > 0$, let $F_m = \mathbb{F}_{p^m}$. Recall that $F = \bigcup_{m>0} F_m$ and $F_m \subseteq F_k$ if and only if m divides k. So we may fix some $m > 0$ such that F_m contains the coefficients of the map Φ. It follows that, for any $k > 0$, $\Phi((F_{km})^n) \subseteq (F_{km})^n$, and so $\Phi((F_{km})^n) = (F_{km})^n$ since Φ is injective and $(F_{km})^n$ is finite. Since $F = \bigcup_{k>0} F_{km}$, it follows that Φ is surjective. □

We now prove Ax's Theorem.

Proof of Theorem 6.12. Fix $n > 0$ and $d > 0$. By quantifying over coefficients of polynomials (as in the definition of ACF), we may construct an $\mathcal{L}_r$-sentence $\psi_{n,d}$ such that a field $F \models \psi_{n,d}$ if and only if every injective polynomial map $\Phi : F^n \longrightarrow F^n$, whose coordinates are polynomials over F of degree at most d, is surjective. We want to show $(\mathbb{C}, +, -, \cdot, 0, 1) \models \psi_{n,d}$. It suffices to show $\mathrm{ACF}_0 \models \psi_{n,d}$. Since ACF_0 is complete, it is enough to prove that $\mathrm{ACF}_0 \cup \{\psi_{n,d}\}$ is satisfiable. By the Compactness Theorem, it suffices to fix a finite subset $\Delta \subseteq \mathrm{ACF}_0 \cup \{\psi_{n,d}\}$, and prove that Δ is satisfiable. By definition of ACF_0, there is a sufficiently large prime p such that $\Delta \subseteq \mathrm{ACF}_p \cup \{\psi_{n,d}\}$. By Lemma 6.13, $(\mathbb{F}_p^{alg}, +, -, \cdot, 0, 1) \models \Delta$. □

Note that the statement of Ax's Theorem holds for any algebraically closed field in place of $\mathbb{C}$. Exercise 7.6.8 captures the model theoretic content of Ax's Theorem, commonly known as the *Lefschetz principle*, and is proved using similar techniques.

Definition 6.14. Let F be a field.

1. Given $S \subseteq F[x_1, \ldots, x_n]$, define $V(S) = \{\bar{a} \in F^n : p(\bar{a}) = 0 \text{ for all } p(\bar{x}) \in S\}$.
2. A subset $X \subseteq F^n$ is **Zariski closed** if is of the form $V(S)$ for some finite $S \subseteq F[x_1, \ldots, x_n]$.
3. A subset $X \subseteq F^n$ is **constructible** if it is a finite Boolean combination of Zariski closed sets.

Lemma 6.15. *Let K be an algebraically closed field. A subset $X \subseteq K^n$ is definable if and only if it is constructible.*

Proof. The reverse direction is clear. Suppose X is definable by some formula $\varphi(\bar{x}, \bar{a})$, with $\bar{a} \in K^m$ for some $m > 0$. By quantifier elimination, we may assume $\varphi(\bar{x}, \bar{y})$ is quantifier-free, and therefore a Boolean combination of atomic formulas. So we may assume $\varphi(\bar{x}, \bar{y})$ is atomic, which means it is equivalent to $p(\bar{x}, \bar{y}) = 0$ for some polynomial $p(\bar{x}, \bar{y}) \in \mathbb{Z}[\bar{x}, \bar{y}]$. Then $X = V(p(\bar{x}, \bar{a}))$, which is constructible. □

Theorem 6.16 (Chevalley)**.** *Let K be an algebraically closed field. If $X \subseteq K^n$ is constructible and $\Phi(\bar{x})$ is a polynomial map, then $\Phi(X)$ is constructible.*

Proof. X is definable and Φ is a a definable function. So $\Phi(X)$ is definable, and therefore constructible by Lemma 6.15. (See Exercise 7.2.3.) □

Theorem 6.17 (Hilbert's Nullstellensatz)**.** *Let K be an algebraically closed field and suppose $I, J \subseteq K[\bar{x}]$ are radical ideals. If $V(I) = V(J)$ then $I = J$.*

Proof. Assume $I \neq J$ and suppose, without loss of generality, there is $p \in J \backslash I$. By Fact B.54(b), we may find a prime ideal $P \supseteq I$ such that $p \notin P$. Since P is prime, $K[\bar{X}]/P$ is an integral domain, and so we may define F to be the algebraic closure of its field of fractions. Let $\bar{a} = ([X_1], \ldots [X_n]) \in F^n$. Then $q(\bar{a}) = 0$ for all $q(\bar{x}) \in I$, and $p(\bar{a}) \neq 0$. By Fact B.54(a), we may fix generators $q_1, \ldots, q_m \in I$. Let $\bar{b}$ be the coefficients (in K) of $q_1, \ldots, q_m$ and p. Let $\varphi(\bar{x}, \bar{b})$ be a formula expressing $q_i(\bar{x}) = 0$ for all $1 \leq i \leq m$, and $p(\bar{x}) \neq 0$. We have $F \models \exists \bar{x} \varphi(\bar{x}, \bar{b})$. Since ACF is model complete (by Proposition 6.3) and K is a substructure of F, it follows that $K \models \exists \bar{x} \varphi(\bar{x}, \bar{b})$. A solution in K^n to this formula witnesses $V(I) \neq V(J)$. □

Remark 6.18.

1. Suppose F is a field and $X = V(S) \subseteq F^n$, where $S \subseteq F[x_1, \ldots, x_n]$ (not necessarily finite). Using Hilbert's Basis Theorem (Fact B.54(a)), one can show that $X = V(S_0)$ for some finite $S_0 \subseteq F[x_1, \ldots, x_n]$.
2. Suppose K is an algebraically closed field. Hilbert's Nullstellensatz is used to establish a bijection between radical ideals in $K[x_1, \ldots, x_n]$ and Zariski closed subsets of K^n, given by $I \mapsto V(I)$.

7 Exercises

Exercises marked with an asterisk ($*$) may be more challenging.

7.1 Languages and Structures

Exercise 7.1.1. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure.*

(a) *Show that the $\mathcal{L}$-substructure relation $\subseteq$ is transitive, i.e., if $\mathcal{M}_1 \subseteq \mathcal{M}_2$ and $\mathcal{M}_2 \subseteq \mathcal{M}_3$ then $\mathcal{M}_1 \subseteq \mathcal{M}_3$.*

(b) *Suppose $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq \ldots$ is an infinite chain of substructures of $\mathcal{M}$. Let $N = \bigcup_{n\geq 0} M_n$. Prove that there is a unique substructure $\mathcal{N}$ of $\mathcal{M}$, with universe N, such that $\mathcal{M}_n \subseteq \mathcal{N}$ for all $n \geq 0$.*

Exercise 7.1.2. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure and fix a nonempty subset $A \subseteq M$. Define*

$$N = \{t^{\mathcal{M}}(\bar{a}) : n \geq 0,\ \bar{a} \in A^n,\ t^{\mathcal{M}}(v_1, \ldots, v_n) \text{ is an } \mathcal{L}\text{-term}\}.$$

(a) *Suppose f is an n-ary function symbol in $\mathcal{L}$ (with $n \geq 0$). Prove that, for any $\bar{a} \in N^n$, $f^{\mathcal{M}}(\bar{a}) \in N$.*

(b) *Let $\mathcal{N}$ be the $\mathcal{L}$-structure, with universe $\mathcal{N}$, such that:*

   (i) *given an n-ary function symbol f (with $n \geq 0$), $f^{\mathcal{N}} = f^{\mathcal{M}}|_{N^n}$, and*

   (ii) *given an n-ary relation symbol R, $R^{\mathcal{N}} = N^n \cap R^{\mathcal{M}}$.*

   Prove that $\mathcal{N}$ is a substructure of $\mathcal{M}$ containing A.

(c) *Let $\mathcal{N}$ be as in part (b) and suppose that $\mathcal{M}'$ is a substructure of $\mathcal{M}$ containing A. Prove that $\mathcal{N}$ is a substructure of $\mathcal{M}'$. We call $\mathcal{N}$ the* ***substructure of $\mathcal{M}$ generated by*** *A.*

(d) *Suppose $(K, +, \cdot, -, 0, 1)$ is a field. Given $A \subseteq K$, describe the substructure generated by A.*

Exercise 7.1.3. *Let $\mathcal{L}$ be a language and κ an infinite cardinal. Prove that there are at most 2^κ non-isomorphic $\mathcal{L}$-structures of cardinality κ.*

7.2 Formulas and Definable Sets

Exercise 7.2.1. *Find an example of $\mathcal{L}$-structures $\mathcal{M} \subseteq \mathcal{N}$ and a formula $\varphi(v_1, \ldots, v_n)$ such that, for some tuple $\bar{a} \in M^n$, $\mathcal{M} \models \varphi(a_1, \ldots, a_n)$ and $\mathcal{N} \models \neg\varphi(a_1, \ldots, a_n)$.*

Exercise 7.2.2. *Prove that the even numbers are definable in the structure $(\mathbb{N}, +, 0)$.*

Exercise 7.2.3. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure. We say that a function $f : M^m \longrightarrow M^n$ is* ***definable*** *if*

$$\{(\bar{x}, f(\bar{x})) : \bar{x} \in M^m\} \subseteq M^{m+n}$$

is definable in $\mathcal{M}$.

(a) *Prove that if $f : M^k \longrightarrow M^m$ and $g : M^m \longrightarrow M^n$ are definable functions then $g \circ f : M^k \longrightarrow M^n$ is definable.*

(b) *Suppose that $f : M^m \longrightarrow M^n$ is definable. Prove that the set $f(M^m) \subseteq M^n$ is definable.*

Exercise 7.2.4. *Let $\mathbb{K} = (K, +, \cdot, -, 0, 1)$ be a field of characteristic 0. Given $n > 0$, let $GL_n(K)$ be the set of $n \times n$ matrices with entries in K and nonzero determinant.*

(*a*) *Prove that $GL_n(K)$ is definable in $\mathbb{K}$ (where $GL_n(K)$ is viewed as a subset of K^{n^2}).*

(*b*) *Prove that the subset of $GL_n(K)$ consisting of the diagonalizable matrices is definable.*

Exercise 7.2.5. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure.*

(*a*) *Fix $\mathcal{L}$-formulas $\varphi(v_{i_1}, \ldots, v_{i_r})$ and $\psi(v_{j_1}, \ldots, v_{j_s})$, with $\{i_1, \ldots, i_r, j_1, \ldots, j_s\} = \{1, \ldots, n\}$. Given $\bar{a} \in M^n$, prove the following statements:*

  (*i*) *$\mathcal{M} \models (\varphi \vee \psi)(\bar{a})$ if and only if: $\mathcal{M} \models \varphi(a_{i_1}, \ldots, a_{i_r})$ or $\mathcal{M} \models \psi(a_{j_1}, \ldots, a_{j_s})$.*

  (*ii*) *$\mathcal{M} \models (\varphi \rightarrow \psi)(\bar{a})$ if and only if: $\mathcal{M} \models \varphi(a_{i_1}, \ldots, a_{i_r})$ implies $\mathcal{M} \models \psi(a_{j_1}, \ldots, a_{j_s})$.*

  (*iii*) *$\mathcal{M} \models (\varphi \leftrightarrow \psi)(\bar{a})$ if and only if: $\mathcal{M} \models \varphi(a_{i_1}, \ldots, a_{i_r})$ if and only if $\mathcal{M} \models \psi(a_{j_1}, \ldots, a_{j_s})$.*

(*b*) *Fix an $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n, w)$. Given $\bar{a} \in M^n$, prove*

$$\mathcal{M} \models \forall w \varphi(\bar{a}, w) \;\; \Leftrightarrow \;\; \text{for all } b \in M,\ \mathcal{M} \models \varphi(\bar{a}, b).$$

Exercise 7.2.6. *Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures and $\sigma : \mathcal{M} \longrightarrow \mathcal{N}$ is an isomorphism. Prove that, for any $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n)$ and $\bar{a} \in M^n$,*

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

(Hint: start with the proof of Proposition 2.11.)

Exercise 7.2.7.* *Consider the ring $\mathcal{M} = (\mathbb{Z}, +, -, \cdot, 0, 1)$. Prove that ordering on $\mathbb{Z}$ is definable in $\mathcal{M}$ (as a set $\{(x, y) \in \mathbb{Z}^2 : x < y\}$).*

Definition 7.2.8. Let $\mathcal{M}$ be an $\mathcal{L}$-structure, and fix $A \subseteq M$.

1. A set $X \subseteq M^n$ is **A-definable in $\mathcal{M}$** if X is definable using an $\mathcal{L}$-formula with parameters in A.

2. The **definable closure of A in $\mathcal{M}$** is the set

$$\mathrm{dcl}_{\mathcal{M}}(A) = \{b \in M : \{b\} \text{ is } A\text{-definable in } \mathcal{M}\}.$$

Exercise 7.2.9. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure.*

(*a*) *Prove that for any $n > 0$, if $X \subseteq M^n$ is finite then X is definable in $\mathcal{M}$.*

(*b*) *Prove that if $X \subseteq M^n$ is definable in $\mathcal{M}$ then there is a finite set $A \subseteq M$ such that X is A-definable in $\mathcal{M}$.*

(*c*) *Prove that $\mathrm{dcl}_{\mathcal{M}}$ is a* ***closure operator****, i.e.,*

  (*i*) *for all $A \subseteq M$, $A \subseteq \mathrm{dcl}_{\mathcal{M}}(A)$ and $\mathrm{dcl}_{\mathcal{M}}(\mathrm{dcl}_{\mathcal{M}}(A)) = \mathrm{dcl}_{\mathcal{M}}(A)$,*

  (*ii*) *for all $A, B \subseteq M$, if $A \subseteq B$ then $\mathrm{dcl}_{\mathcal{M}}(A) \subseteq \mathrm{dcl}_{\mathcal{M}}(B)$.*

(*d*) *Prove that the closure operator $\mathrm{dcl}_{\mathcal{M}}$ has* ***finite character****, i.e. for any $A \subseteq M$,*

$$\mathrm{dcl}_{\mathcal{M}}(A) = \bigcup_{A_0 \subseteq A,\ |A_0| < \aleph_0} \mathrm{dcl}_{\mathcal{M}}(A_0).$$

(*e*) *Let $X \subseteq M^n$ be A-definable in $\mathcal{M}$, and suppose $A \subseteq \mathrm{dcl}_{\mathcal{M}}(B)$. Prove that X is B-definable in $\mathcal{M}$.*

(*f*) *Suppose $\mathcal{L}$ contains a binary relation $<$, and $\mathcal{M}$ is an $\mathcal{L}$-structure such that $<^{\mathcal{M}}$ is a linear order on M. Prove that $\mathrm{acl}_{\mathcal{M}}(A) = \mathrm{dcl}_{\mathcal{M}}(A)$ for any $A \subseteq M$.*

(*g*) *Suppose $\mathcal{L} = \mathcal{L}_{or}$ and $\mathcal{M} = (\mathbb{R}, +, -, \cdot, <, 0, 1)$. Prove that $\mathrm{dcl}_{\mathcal{M}}(\emptyset)$ contains all real algebraic numbers.*

Definition 7.2.10. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Given $A \subseteq M$ and $b \in M$, we say b is **algebraic over A in $\mathcal{M}$** if there is a formula $\varphi(x, a_1, \ldots, a_n)$ with parameters $a_1, \ldots, a_n \in A$ such that $\mathcal{M} \models \varphi(b, \bar{a})$ and $\varphi(M, \bar{a})$ is finite. In other words, $b \in M$ is algebraic over A in $\mathcal{M}$ if and only if there is a finite A-definable subset of M containing b.

The **algebraic closure of A in $\mathcal{M}$** is the set

$$\mathrm{acl}_{\mathcal{M}}(A) = \{b \in M : b \text{ is algebraic over } A \text{ in } \mathcal{M}\}.$$

Exercise 7.2.11. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Prove that $\mathrm{acl}_{\mathcal{M}}$ is a closure operator with finite character.*

Definition 7.2.12. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. Given $A \subseteq M$, define $\mathrm{Aut}(\mathcal{M}/A)$ to be the set of $\mathcal{L}$-automorphisms σ of $\mathcal{M}$ such that $\sigma(a) = a$ for all $a \in A$.

Exercise 7.2.13. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure.*

(*a*) *Prove that, for any $A \subseteq M$, $\mathrm{Aut}(\mathcal{M}/A)$ is a group under composition of $\mathcal{L}$-automorphisms.*

(*b*) *Suppose $A \subseteq \mathbb{M}$ and $\sigma \in \mathrm{Aut}(\mathcal{M}/A)$.*

(*i*) *Prove that, for any $\mathcal{L}$-formula $\varphi(v_1, \ldots, v_n, a_1, \ldots, a_m)$, with parameters $a_1, \ldots, a_m \in A$, and $b_1, \ldots, b_n \in M$,*

$$\mathcal{M} \models \varphi(b_1, \ldots, b_n, \bar{a}) \quad \Leftrightarrow \quad \mathcal{M} \models \varphi(\sigma(b_1), \ldots, \sigma(b_n), \bar{a}).$$

(*ii*) *Prove that, for any $b \in M$, $b \in \mathrm{dcl}_{\mathcal{M}}(A)$ if and only if $\sigma(b) \in \mathrm{dcl}_{\mathcal{M}}(A)$, and $b \in \mathrm{acl}_{\mathcal{M}}(A)$ if and only if $\sigma(b) \in \mathrm{acl}_{\mathcal{M}}(A)$.*

(*c*) *Let $\mathcal{L} = \mathcal{L}_r$ be the language of rings and $\mathcal{M} = (\mathbb{C}, +, -, \cdot, 0, 1)$. Prove that, for any $A \subseteq \mathbb{C}$, $\mathrm{dcl}_{\mathcal{M}}(A)$ contains the field generated by A, and $\mathrm{acl}_{\mathcal{M}}(A)$ contains the field-theoretic algebraic closure of the field generated by A. (In fact, $\mathrm{dcl}_{\mathcal{M}}(A)$ and $\mathrm{acl}_{\mathcal{M}}(A)$ are precisely these sets; see Exercise 7.6.7.)*

7.3 Sentences and Theories

Exercise 7.3.1. *For any $n > 0$, $\mathbb{Z}^n$ is a group under coordinate-wise addition of tuples. Prove that if $m \neq n$ then $(\mathbb{Z}^m, +, 0)$ and $(\mathbb{Z}^n, +, 0)$ are not elementarily equivalent.*

Exercise 7.3.2.

(*a*) *Consider $\mathcal{L}_r$. Show that the following are elementary classes and give axiomatizations of their theories:*

(*i*) *the class of rings,*

(ii) *the class of fields,*

(iii) *the class of fields of characteristic* 0,

(iv) *the class of fields of characteristic* p *for some fixed prime* p,

(v) *the class of algebraically closed fields.*

(b) *Consider* $\mathcal{L}_{gr}$. *Show that the following are elementary classes and give axiomatizations of their theories:*

(i) *the class of triangle-free graphs,*

(ii) *the class of graphs where every vertex has infinite degree (the degree of a vertex* v *is the number of vertices adjacent to* v*),*

(iii) *the class of bipartite graphs.*

(c) *Consider* $\mathcal{L}_o$. *Show that the following are elementary classes and give axiomatizations of their theories:*

(i) *the class of dense linear orders,*

(ii) *the class of discrete linear orders (i.e. where every non-maximal element has an immediate successor and every non-minimal element has an immediate predecessor).*

Exercise 7.3.3. *Suppose* $\mathcal{M}$ *is an* $\mathcal{L}$*-structure, which is generated by a subset* $A \subseteq M$. *Let* $\mathcal{L}_M = \mathcal{L} \cup \{\tilde{m} : m \in M\}$, *where each* $\tilde{m}$ *is a new constant symbol. Let* $\mathcal{L}_A = \mathcal{L} \cup \{\tilde{m} : m \in A\}$, *and note that* $\mathcal{L}_A \subseteq \mathcal{L}_M$. *We can view* $\mathcal{M}$ *as an* $\mathcal{L}_M$*-structure by interpreting each constant symbol* $\tilde{m}$ *as the element* m. *Let* T^* *be the* $\mathcal{L}_M$*-theory of* $\mathcal{M}$, *and fix a subset* $T \subseteq T^*$.

Prove that there is an $\mathcal{L}_A$*-theory* T_0 *such that, for any* $\mathcal{L}_A$*-sentence* φ, *if* $T \models \varphi$ *then* $T_0 \models \varphi$.

Exercise 7.3.4. *Let* $\mathcal{L}$ *be a language, and suppose* t *is an* $\mathcal{L}$*-term with no variables (i.e. compositions of function symbols and constant symbols). Then we have the* 0*-ary function* $t^{\mathcal{M}} : M^0 \longrightarrow M$ *as given by Definition 2.4. We identify* $t^{\mathcal{M}}$ *with the element* $t^{\mathcal{M}}(\emptyset) \in M$.

Recall that, for any $\mathcal{L}$*-sentence* φ, $\mathcal{M} \models \varphi$ *if and only if* $\varphi^{\mathcal{M}} = M^0$, *where* $\varphi^{\mathcal{M}}$ *is constructed as in Definition 2.8. Prove the following explicit statements:*

(a) *If* φ *is* $t_1 = t_2$, *where* t_1 *and* t_2 *are terms with no variables, then*

$$\mathcal{M} \models \varphi \ \Leftrightarrow \ t_1^{\mathcal{M}} = t_2^{\mathcal{M}}.$$

(b) *If* φ *is* $R(t_1, \ldots, t_n)$, *where* R *is an* n*-ary relation symbol and* $t_1, \ldots, t_n$ *are terms with no variables, then*

$$\mathcal{M} \models \varphi \ \Leftrightarrow \ (t_1^{\mathcal{M}}, \ldots, t_n^{\mathcal{M}}) \in R^{\mathcal{M}}.$$

(c) *If* φ *and* ψ *are sentences then*

$$\mathcal{M} \models \varphi \wedge \psi \ \Leftrightarrow \ \mathcal{M} \models \varphi \text{ and } \mathcal{M} \models \psi.$$

(d) *If* $\varphi(v)$ *is a formula then*

$$\mathcal{M} \models \exists v \varphi(v) \ \Leftrightarrow \ \text{there exists } a \in M \text{ such that } \mathcal{M} \models \varphi(a).$$

Exercise 7.3.5. *Suppose* T *is an unsatisfiable* $\mathcal{L}$*-theory. Prove that any* $\mathcal{L}$*-sentence is a logical consequence of* T.

Exercise 7.3.6. *Suppose $\mathcal{L}$ is a language and $\mathcal{M}$ is a finite $\mathcal{L}$-structure.*

(a) *Assume $\mathcal{L}$ is finite. Prove that there is an $\mathcal{L}$-sentence φ such that any model of φ is isomorphic to $\mathcal{M}$.*

(b)* *Prove that any model of $\mathrm{Th}(\mathcal{M})$ is isomorphic to $\mathcal{M}$.*

7.4 The Compactness Theorem

Exercise 7.4.1.

(a) *Let T be an $\mathcal{L}$-theory, where $\mathcal{L}$ contains $\mathcal{L}_{gr}$. Prove that $\mathcal{K} := \{\mathcal{M}|_{\mathcal{L}_{gr}} : \mathcal{M} \models T\}$ is not the class of connected graphs.*

(b) *Let T be an $\mathcal{L}$-theory, where $\mathcal{L}$ contains $\mathcal{L}_g$. Prove that $\mathcal{K} := \{\mathcal{M}|_{\mathcal{L}_g} : \mathcal{M} \models T\}$ is not the class of cyclic groups.*

Exercise 7.4.2.

(a) *An ordered abelian group $(G, +, <, 0)$ is* ***archimedean*** *if, for all $x, y > 0$ there is some $n > 0$ such that $x \leq ny$. Prove that there is a non-archimedean ordered abelian group elementarily equivalent to $(\mathbb{R}, +, <, 0)$.*

(b) *A linear order $(X, <)$ is a* ***well-order*** *if it contains no infinite descending chains. Prove that there is a non-well-order $(X, <)$ elementarily equivalent to $(\mathbb{N}, <)$.*

Exercise 7.4.3. *Suppose T is an $\mathcal{L}_r$-theory extending the theory of fields. Prove that if T has models of arbitrarily large characteristic, then T has a model of characteristic 0.*

Exercise 7.4.4. *Let T_1 and T_2 be satisfiable $\mathcal{L}$-theories.*

(a) *Suppose $T_1 \cup T_2$ is unsatisfiable. Prove that there is an L-sentence φ such that $T_1 \models \varphi$ and $T_2 \models \neg\varphi$.*

(b) *Suppose that if $\mathcal{M}$ is an $\mathcal{L}$-structure then $\mathcal{M} \models T_1$ if and only if $\mathcal{M} \not\models T_2$. Prove that T_1 and T_2 are finitely axiomatizable.*

Exercise 7.4.5. *Use Theorem A.1 in Appendix A to prove Proposition 4.5.*

Exercise 7.4.6. *Let $\mathcal{L}$ be a language, with $\mathcal{L}_g \subseteq \mathcal{L}$, and let T be an $\mathcal{L}$-theory extending the theory of groups. Assume that for any $n > 0$ there is a group in $\mathrm{Mod}(T)$ containing a torsion point of order greater than n. Prove that there is no $\mathcal{L}$-formula $\varphi(x)$ such that, for any $\mathcal{M} \models T$, $\varphi^{\mathcal{M}}$ is precisely the set of torsion points in $\mathcal{M}$.*

7.5 Elementary Extensions

Exercise 7.5.1 (Tarski-Vaught Test)**.** *Fix $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ such that $\mathcal{M}$ is a substructure of $\mathcal{N}$. Prove that the following are equivalent.*

(i) *$\mathcal{M}$ is an elementary substructure of $\mathcal{N}$.*

(ii) *For any formula $\varphi(v, w_1, \ldots, w_n)$ and $\bar{a} \in M^n$, if $\mathcal{N} \models \exists v \varphi(v, \bar{a})$ then there is some $b \in M$ such that $\mathcal{N} \models \varphi(b, \bar{a})$.*

Exercise 7.5.2. *Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure and $\mathcal{M}_0 \prec \mathcal{M}_1 \prec \mathcal{M}_2 \prec \ldots$ is an infinite chain of elementary substructures of $\mathcal{M}$. Let $\mathcal{N}$ be the substructure of $\mathcal{M}$ with universe $\bigcup_{n\geq 0} M_n$. Prove that $\mathcal{N} \prec \mathcal{M}$, and that $\mathcal{M}_n \prec \mathcal{N}$ for all $n > 0$.*

Exercise 7.5.3.

(a) *Let $\mathcal{L}$ be a language extending the language of ordered groups, and fix an $\mathcal{L}$-structure $\mathcal{R}$ expanding $(\mathbb{R}, +, <, 0)$. Prove that there is an elementary extension $\mathcal{M}$ of $\mathcal{R}$ and an element $\mu \in M$ such that $0 < \mu < r$ for all real numbers $r > 0$.*

(b) *A linear order $(X, <)$ satisfies the* ***least upper bound property*** *if every nonempty $Y \subseteq X$, with an upper bound in X, has a least upper bound in X. Let $\mathcal{L}$ be a language, extending the language of orders, and suppose $\mathcal{R}$ is an $\mathcal{L}$-structure expanding $(\mathbb{R}, <)$.*

  (i) *Let $\mathcal{M}$ be an elementary extension of $\mathcal{R}$ and suppose $X \subseteq M$ is nonempty and definable in $\mathcal{M}$. Prove that if X has an upper bound in M then it has a least upper bound in M.*

  (ii) *Show that that there is an elementary extension $\mathcal{M}$ of $\mathcal{R}$ such that the underlying order on $\mathcal{M}$ does not satisfy the least upper bound property. (Hint: use part (a).)*

Exercise 7.5.4. *Given $k > 0$, a graph is k-**colorable** if there is a coloring of the vertices, using at most k colors, such that no two adjacent vertices are the same color. Prove that a graph is k-colorable if and only if every finite subgraph is k-colorable.*[2]

Exercise 7.5.5. *Find an example of $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ such that $\mathcal{M}$ is a substructure of $\mathcal{N}$, but not an elementary substructure of $\mathcal{N}$.*

Exercise 7.5.6. *Finish the proof of Proposition 5.6.*

Exercise 7.5.7.

(a) *Suppose $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$ and $A \subseteq M$. Prove that $\mathrm{acl}_{\mathcal{M}}(A) = \mathrm{dcl}_{\mathcal{N}}(A)$ and $\mathrm{acl}_{\mathcal{M}}(A) = \mathrm{acl}_{\mathcal{N}}(A)$. (See Definitions 7.2.8 and 7.2.10).*

(b) *Find examples showing that part (a) can fail if we only assume $\mathcal{M} \equiv \mathcal{N}$ and $\mathcal{M} \subseteq \mathcal{N}$.*

7.6 Quantifier Elimination

Exercise 7.6.1. *Consider the theory $T = \mathrm{TFDAG}$ in the language of groups.*

(a) *Let $(G, +, 0) \models T$. Given $q \in \mathbb{Q}$, write $q = \frac{m}{n}$ in lowest terms and define a function $\lambda_q : G \longrightarrow G$ such that, given $g \in G$, $\lambda_q(g)$ is the unique solution in G of the equation*

$$\underbrace{x + x + \ldots + x}_{n\ times} = \underbrace{g + g + \ldots + g}_{m\ times},$$

Prove that $(G, +, 0, (\lambda_q)_{q\in\mathbb{Q}})$ is a vector space over $\mathbb{Q}$ (see Example 3.7(3)).

(b) *Prove that that if $(V, +, 0, (\lambda_q)_{q\in\mathbb{Q}})$ is a vector space over $\mathbb{Q}$ then $(V, +, 0) \models T$.*

(c) *Suppose $\mathcal{M} \models T$ and $\mathcal{M}_0 \subseteq \mathcal{M}$. Prove that any two models of $T \cup \mathrm{Diag}(\mathcal{M}_0)$, of cardinality $\aleph_1$, are isomorphic.*

[2] In particular, using the Four-Color Theorem for *finite* planar graphs, this can be used to conclude that any *infinite* planar graph is also four colorable.

(d) *Prove that T is complete and has quantifier elimination.*

(e) *Classify the models of T up to isomorphism.*

Exercise 7.6.2. *Consider the language $\mathcal{L}_o$ of orders. A linear order $(M, <)$ is* ***dense*** *if, for all $x, y \in M$ there is some $z \in M$ such that $x < y < z$.*

(a) *Write down a finite $\mathcal{L}_o$-theory T such that $\mathcal{M} \models T$ if and only if $\mathcal{M}$ is a dense linear order with no greatest element or least element.*

(b)* *Suppose $\mathcal{M} \models T$ and $\mathcal{M}_0 \subseteq \mathcal{M}$. Prove that any two countable models of $T \cup \mathrm{Diag}(\mathcal{M}_0)$ are isomorphic. Conclude that any two countable models of T are isomorphic.*

(c) *Prove that T is complete and has quantifier elimination.*

(d) *By part (b), $(\mathbb{Q}, <)$ is the unique countable model of T. Prove that $(\mathbb{Q}, <)$ has the following properties:*

- (i) *(universality) any finite linear order is isomorphic to a substructure of $(\mathbb{Q}, <)$;*
- (ii) *(ultrahomogeneity) any isomorphism between finite substructures of $(\mathbb{Q}, <)$ extends to an automorphism of $(\mathbb{Q}, <)$ (Hint: use part (b)).*

(e) *Prove that any countable linear order is isomorphic to a substructure of $(\mathbb{Q}, <)$.*

Exercise 7.6.3. *Consider the language $\mathcal{L}_{gr}$ of graphs.*

(a) *Write down an $\mathcal{L}_{gr}$-theory T such that $\mathcal{M} \models T$ if and only if $\mathcal{M}$ is a graph such that for any finite disjoint $A, B \subseteq M$ there is a vertex $v \in M$ such that v is connected to every element of A and no element of B.*

(b) *Prove that T is satisfiable.*

Hint: Consider an infinite binary sequence $\sigma = (s_0, s_1, s_2, \ldots)$, with $s_i \in \{0, 1\}$ obtained by concatenating all finite binary sequences in some arbitrary order (e.g. by length, then lexicographically). Now consider the graph $(\mathbb{Z}, E)$ such that $(m, n) \in E$ if and only if $m \neq n$ and $s_{|m-n|} = 1$.

(c)* *Suppose $\mathcal{M} \models T$ and $\mathcal{M}_0 \subseteq \mathcal{M}$. Prove that any two countable models of $T \cup \mathrm{Diag}(\mathcal{M}_0)$ are isomorphic. Conclude that any two countable models of T are isomorphic.*

(d) *Prove that T is complete and has quantifier elimination.*

(e) *The unique countable model of T is called the* ***random graph*** *(or* ***Rado graph****), which we denote $\mathcal{R} = (V(\mathcal{R}), E(\mathcal{R}))$. Prove that $\mathcal{R}$ has the following properties:*

- (i) *(universality) any finite graph is isomorphic to an induced subgraph of $\mathcal{R}$;*
- (ii) *(ultrahomogeneity) any isomorphism between finite subgraphs of $\mathcal{R}$ extends to an automorphism of $\mathcal{R}$ (Hint: use part (c)).*

(f) *Prove that any countable graph is isomorphic to an induced subgraph of $\mathcal{R}$.*

Exercise 7.6.4. *Prove Proposition 6.3 (use Proposition 2.11).*

Exercise 7.6.5.* *Suppose $\mathcal{L}$ contains a constant symbol and T is an $\mathcal{L}$-theory with quantifier elimination. Prove that, for any $\mathcal{L}$-sentence φ there is a quantifier-free $\mathcal{L}$-sentence ψ such that $T \models \varphi \leftrightarrow \psi$. (Hint: adapt the right-to-left direction of Theorem 6.6).*

Exercise 7.6.6. *Prove the left-to-right direction of Theorem 6.6.*

Exercise 7.6.7. *Let K be an algebraically closed field and fix a subset $A \subseteq K$.*

(*a*) *Prove that $\mathrm{dcl}_K(A)$ is the field generated by A.*

(*b*) *Prove that $\mathrm{acl}_K(A)$ is the algebraic closure of the field generated by A.*

Exercise 7.6.8 (The Lefschetz Principle)**.** *Let φ be a sentence in the language of rings. Prove that the following are equivalent.*

(*i*) *φ is true in the complex numbers (i.e. $(\mathbb{C}, +, -, \cdot, 0, 1) \models \varphi$).*

(*ii*) *φ is true in any algebraically closed field of characteristic 0 (i.e. $\mathrm{ACF}_0 \models \varphi$).*

(*iii*) *φ is true in some algebraically closed field of characteristic 0 (i.e. $\mathrm{ACF}_0 \cup \{\varphi\}$ is satisfiable).*

(*iv*) *There are arbitrarily large primes p such that φ is true in some algebraically closed field of characteristic p (i.e. $\mathrm{ACF}_p \cup \{\varphi\}$ is satisfiable for arbitrarily large p).*

(*v*) *There is an integer m such that φ is true in any algebraically closed field of characteristic $p > m$ (i.e. there is an integer m such that $\mathrm{ACF}_p \models \varphi$ for all $p > m$).*

A Compactness and Löwenheim-Skolem

In this appendix, we consider the following strengthening of the Compactness Theorem.

Theorem A.1. *Suppose T is a finitely satisfiable $\mathcal{L}$-theory and $\kappa \geq \max\{|\mathcal{L}|, \aleph_0\}$. Then T has a model of cardinality at most κ.*

We will give two proofs of Theorem A.1. Specifically:

1. In Section A.2 we prove the Downward Löwenheim-Skolem Theorem, which requires Lemma A.3 from Section A.1. We then use this result, together with the Compactness Theorem, to prove Theorem A.1. A proof of the Compactness Theorem using ultraproducts will be given in the next course.
2. In Section A.3, we prove Theorem A.1 directly using a *Henkin construction.* This proof requires Lemma A.3 from Section A.1.

A.1 Skolemization

Definition A.2. An $\mathcal{L}$-theory T has **built-in Skolem functions** if, for all $\mathcal{L}$-formulas $\varphi(v, w_1, \ldots, w_n)$ (with n possibly 0) there is an n-ary function symbol f in $\mathcal{L}$ such that the $\mathcal{L}$ sentence

$$\forall \bar{w}\big(\exists v \varphi(v, \bar{w}) \rightarrow \varphi(f(\bar{w}), \bar{w})\big)$$

is in T.

A theory T with built-in Skolem functions is also called *Skolemized.*

Lemma A.3. *Let T be a finitely satisfiable $\mathcal{L}$-theory. Then there is a language $\mathcal{L}^* \supseteq \mathcal{L}$ and an $\mathcal{L}^*$-theory $T^* \supseteq T$ satisfying the following properties:*

(i) $|\mathcal{L}^*| = \max\{|\mathcal{L}|, \aleph_0\}$;

(ii) *T^* is finitely satisfiable and has built-in Skolem functions;*

(iii) *any model $\mathcal{M}$ of T can be expanded to a model $\mathcal{M}^*$ of T^*.*

Proof. We inductively construct chains $\mathcal{L} = \mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \mathcal{L}_2 \subseteq \ldots$ and $T = T_0 \subseteq T_1 \subseteq T_2 \subseteq \ldots$ such that, for all $m \geq 0$,

(*a*) $|\mathcal{L}_m| = \max\{|\mathcal{L}|, \aleph_0\}$ and T_m is a finitely satisfiable $\mathcal{L}_m$-theory;

(*b*) for any subset $\Delta \subseteq T_{m-1}$, any model $\mathcal{M}$ of Δ can be expanded to a model $\mathcal{M}'$ of $\Delta \cup (T_m \backslash T_{m-1})$.

Given $\mathcal{L}_m$ and T_m, set

$$\mathcal{L}_{m+1} = \{f_\varphi : \varphi(v, w_1, \ldots, w_n) \text{ is an } \mathcal{L}_m\text{-formula and } n \geq 0\},$$

where each f_φ is a new n-ary function symbol. Then $|\mathcal{L}_{m+1}| = \max\{|\mathcal{L}|, \aleph_0\}$ by induction. Given an $\mathcal{L}_m$-formula $\varphi(v, w_1, \ldots, w_n)$, let Ψ_φ denote the $\mathcal{L}_{m+1}$-sentence

$$\forall \bar{w}\big(\exists v \varphi(v, \bar{w}) \rightarrow \varphi(f(\bar{w}), \bar{w})\big).$$

Define $T_{m+1} = T_m \cup \{\Psi_\varphi : \varphi \text{ is an } \mathcal{L}_m\text{-formula}\}$.

We prove (b) for $m+1$. Fix a subset $\Delta \subseteq T_m$ and a model $\mathcal{M}$ of Δ and expand $\mathcal{M}$ to a model $\mathcal{M}'$ of $\Delta \cup (T_{m+1} \backslash T_m)$. Fix an $\mathcal{L}_m$-formula $\varphi(v, w_1, \ldots, w_n)$, where $n \geq 0$. We define $f_\varphi^{\mathcal{M}} : M^n \longrightarrow M$ such that

$$f_\varphi^{\mathcal{M}}(\bar{a}) = \begin{cases} \text{some fixed element of } X_{\bar{a}} := \{b \in M : \mathcal{M} \models \varphi(b, \bar{a})\} & \text{if } X_{\bar{a}} \neq \emptyset, \\ \text{an arbitrary } c \in M & \text{if } X_{\bar{a}} = \emptyset. \end{cases}$$

Let $\mathcal{M}'$ be the expansion of $\mathcal{M}$ by interpreting each f_φ as $f_\varphi^{\mathcal{M}}$. Then $\mathcal{M}' \models \Psi_\varphi$ for all φ, and so $\mathcal{M}' \models \Delta \cup (T_{m+1} \backslash T_m)$.

Finally, we prove T_{m+1} is finitely satisfiable. Any finite subset of T_{m+1} is contained in a set of the form

$$\Delta \cup (T_{m+1} \backslash T_m)$$

where $\Delta \subseteq T_m$ is finite. By induction, there is a model $\mathcal{M}$ of Δ. By the above we may expand $\mathcal{M}$ to a model $\mathcal{M}'$ of $\Delta \cup (T_{m+1} \backslash T_m)$. Therefore T_{m+1} is finitely satisfiable.

Now set $\mathcal{L}^* = \bigcup_{m \geq 0} \mathcal{L}_m$ and $T^* = \bigcup_{m \geq 0} T_m$. Then $|\mathcal{L}^*| = \max\{|\mathcal{L}|, \aleph_0\}$ and T^* is finitely satisfiable by property (a). By iterating property (b) (with $\Delta = T_m$), it follows that any model $\mathcal{M}$ of T can be expanded to a model of T^*. Since any $\mathcal{L}^*$-formula is an $\mathcal{L}_m$-formula for some $m \geq 0$, it follows that T^* has built-in Skolem functions. $\square$

A.2 The Löwenheim-Skolem Theorems

Theorem 5.7. *Let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure.*

(a) (Upward Löwenhein-Skolem Theorem) *Given an infinite cardinal κ, with $\kappa \geq \max\{|M|, |\mathcal{L}|\}$, there is an elementary extension $\mathcal{N} \succ \mathcal{M}$ such that $|N| = \kappa$.*

(b) (Downward Löwenheim-Skolem Theorem) *Given $X \subseteq M$, there is an elementary substructure $\mathcal{N} \prec \mathcal{M}$ such that $X \subseteq N$ and $|N| \leq \max\{|X|, |\mathcal{L}|, \aleph_0\}$.*

Proof. Part (a). Let $\mathcal{L}^* = \mathcal{L}_M \cup \{c_i : i \in \kappa\}$, where each c_i is a new constant symbol. Note that $|\mathcal{L}^*| = \kappa$. Define the $\mathcal{L}^*$-theory

$$T^* = \mathrm{Diag}_{\mathrm{el}}(\mathcal{M}) \cup \{c_i \neq c_j : i, j \in \kappa,\ i \neq j\}.$$

Since $\mathcal{M}$ is infinite, it satisfies any finite subset of T^*. By Theorem A.1, T^* has a model $\mathcal{N}^*$ of cardinality at most κ. By definition of T^*, it follows that $|N| = \kappa$. By Proposition 5.6, $\mathcal{N} = \mathcal{N}^*|_{\mathcal{L}}$ is an elementary extension of $\mathcal{M}$.

Part (b). Note that $\mathrm{Th}(\mathcal{M})$ is a satisfiable $\mathcal{L}$-theory. By Lemma A.3, we may fix a language $\mathcal{L}^* \supseteq \mathcal{L}$ and an $\mathcal{L}^*$-theory $T^* \supseteq \mathrm{Th}(\mathcal{M})$ such that $|\mathcal{L}^*| = \max\{|\mathcal{L}|, \aleph_0\}$, T^* has built-in Skolem functions, and $\mathcal{M}$ can be expanded to a model $\mathcal{M}^*$ of T^*. To simplify notation, we may as well assume that $\mathrm{Th}(\mathcal{M})$ has built-in Skolem functions (with respect to $\mathcal{L}$).

We construct a sequence $X = X_0 \subseteq X_1 \subseteq X_2 \subseteq \ldots$ of subsets of M as follows. Given X_m, define

$$X_{m+1} = X_m \cup \{f^{\mathcal{M}}(\bar{a}) : f \text{ is an } n\text{-ary function symbol for some } n \geq 0, \text{ and } \bar{a} \in X_m^n\}.$$

Note that, $|X_{m+1}| \leq \max\{|X_m|, |\mathcal{L}|, |\aleph_0|\}$.

Let $N = \bigcup_{m \geq 0} X_m$. By induction, $|N| \leq \max\{|X|, |\mathcal{L}|, |\aleph_0|\}$ and $X \subseteq N$. We define an $\mathcal{L}$-structure $\mathcal{N}$ with universe N as follows. Suppose f is an n-ary function symbol, for some $n \geq 0$. For any $\bar{a} \in N^n$, we have $\bar{a} \in X_m^n$ for some $m \geq 0$, and so $f^{\mathcal{M}}(\bar{a}) \in X_{m+1} \subseteq N$. Therefore

$f^{\mathcal{M}}(N^n) \subseteq N$, and so we may interpret $f^{\mathcal{N}} = f^{\mathcal{M}}|_{N^n}$. If R is an n-ary relation symbol, then we interpret $R^{\mathcal{M}} = N^n \cap (R^M)$.

We now have an $\mathcal{L}$-structure $\mathcal{N}$, with universe N. By construction, $\mathcal{N} \subseteq \mathcal{M}$. We use the Tarski-Vaught Test (see Exercise 7.5.1) to prove that $\mathcal{N} \prec \mathcal{M}$. In particular, fix a formula $\varphi(v, w_1, \ldots, w_n)$ and $\bar{a} \in N^n$ such that $\mathcal{M} \models \exists v \varphi(v, \bar{a})$. We want to find $b \in N$ such that $\mathcal{M} \models \varphi(b, \bar{a})$. Since $\mathrm{Th}(\mathcal{M})$ has built-in Skolem functions, we have $\mathcal{M} \models \varphi(f(\bar{a}), \bar{a})$ for some function symbol f. Since $\bar{a} \in N^n$, we have $f^{\mathcal{M}}(\bar{a}) = f^{\mathcal{N}}(\bar{a}) \in N$, and so we may choose $b = f^{\mathcal{N}}(\bar{a})$. □

Note that the proof of the Downward Löwenheim-Skolem Theorem uses only the Tarski-Vaught Test, which can be proved directly from definitions and induction on formulas (see Exercise 7.5.1). Therefore, we can use the Downward Löwenheim-Skolem Theorem and the Compactness Theorem (see Theorem 4.2) to prove Theorem A.1.

Theorem A.1. *Suppose T is a finitely satisfiable $\mathcal{L}$-theory and $\kappa \geq \max\{|\mathcal{L}|, \aleph_0\}$. Then T has a model of cardinality at most κ.*

Proof. First, T is satisfiable by Theorem 4.2. If T has finite models then the result holds trivially. Therefore we may assume T has an infinite model. Let $\mathcal{L}^* = \mathcal{L} \cup \{c_i : i \in \kappa\}$, where each c_i is a new constant symbol and set

$$T^* = T \cup \{c_i \neq c_j : i, j \in \kappa,\ i \neq j\}.$$

Since T has an infinite model, it follows that T^* is finitely satisfiable and therefore has a model $\mathcal{M}$ by Theorem 4.2. By construction $|M| \geq \kappa$, so we may fix a subset $X \subseteq M$, with $|X| = \kappa$. By the Downward Löwenhein-Skolem Theorem there is an elementary substructure $\mathcal{N} \prec \mathcal{M}$ such that $X \subseteq N$ and $|N| \leq \max\{|X|, |\mathcal{L}|, |\aleph_0|\} = \kappa$. Then $\mathcal{N}$ is elementarily equivalent to $\mathcal{M}$, and so $\mathcal{N} \models T$. □

A.3 Proof of the Compactness Theorem via a Henkin construction

In this section, we give a direct proof of Theorem A.1, which yields the Compactness Theorem as an immediate corollary. The method of proof is what is known as a *Henkin construction*. We will need a definition and two lemmas.

Definition A.4. An $\mathcal{L}$-theory T is **maximal** if, for every $\mathcal{L}$-sentence φ, either $\varphi \in T$ or $\neg\varphi \in T$.

Lemma A.5. *Suppose T is a maximal, finitely satisfiable $\mathcal{L}$-theory. For any finite $\Delta \subseteq T$ and $\mathcal{L}$-sentence φ, if $\Delta \models \varphi$ then $\varphi \in T$.*

Proof. If $\varphi \notin T$ then $\neg\varphi \in T$ since T is maximal, and so $\Delta \cup \{\neg\varphi\}$ is a finite subset of T. Since T is finitely satisfiable, there is a model $\mathcal{M} \models \Delta \cup \{\neg\varphi\}$, which contradicts the assumption $\Delta \models \varphi$. □

Lemma A.6. *If T is a finitely satisfiable $\mathcal{L}$-theory then there is a maximal finitely satisfiable $\mathcal{L}$-theory $T' \supseteq T$.*

Proof. Let Σ be the set of finitely satisfiable $\mathcal{L}$-theories extending T. Note that $T \in \Sigma$, and so Σ is nonempty. Suppose $C \subseteq \Sigma$ is linearly ordered by $\subseteq$. Let $T_0 = \bigcup C$. Then any finite subset of T_0 is contained in some element of C, and this therefore satisfiable. So $T_0 \in \Sigma$. By Zorn's Lemma, Σ contains a $\subseteq$-maximal element T'. Therefore, to prove T' is maximal, it suffices to show that, for any $\mathcal{L}$-sentence ϕ, either $T' \cup \{\varphi\}$ or $T' \cup \{\neg\varphi\}$ is finitely satisfiable.

So fix an $\mathcal{L}$-sentence φ. If neither $T' \cup \{\varphi\}$ nor $T' \cup \{\neg\varphi\}$ is finitely satisfiable then there are finite subsets $\Delta_1, \Delta_2 \subseteq T'$ such that $\Delta_1 \cup \{\varphi\}$ and $\Delta_2 \cup \{\neg\varphi\}$ are unsatisfiable. It follows that $\Delta_1 \cup \Delta_2$ must be unsatisfiable. But $\Delta_1 \cup \Delta_2$ is a finite subset of T', which contradicts that T is finitely satisfiable. □

We now give a direct proof Theorem A.1.

Theorem A.1. *Suppose T is a finitely satisfiable $\mathcal{L}$-theory and $\kappa \geq \max\{|\mathcal{L}|, \aleph_0\}$. Then T has a model of cardinality at most κ.*

Proof. First, fix a language $\mathcal{L}^* \supseteq \mathcal{L}$ and a $\mathcal{L}^*$-theory $T^* \supseteq T$ satisfying the conclusions of Lemma A.3. In particular, $|\mathcal{L}^*| \leq \kappa$. By Lemma A.6, there is a maximal, finitely satisfiable $\mathcal{L}^*$-theory $T' \supseteq T^*$. Note that T' still has built-in Skolem functions. To simplify notation, replace $\mathcal{L}$ with $\mathcal{L}^*$ and T with T'.

For any $\mathcal{L}$-formula $\varphi(v)$, in one free variable, we have the 0-ary function symbol f_φ in $\mathcal{L}$, which we will treat as a constant symbol c_φ. Since T has built-in Skolem functions, we have the following property:

$(*)$ for any $\mathcal{L}$-formula $\varphi(v)$, $\exists v\varphi(v) \rightarrow \varphi(c_\varphi)$ is in T.

We now build a model $\mathcal{M} \models T$ of cardinality at most κ. Let $\mathcal{C}$ be the set of constant symbols in $\mathcal{L}$. We define a binary relation $\sim$ on $\mathcal{C}$ such that

$$c \sim d \quad \Leftrightarrow \quad c = d \text{ is in } T.$$

Claim 1: $\sim$ is an equivalence relation on $\mathcal{C}$.
Proof: For any $c \in \mathcal{C}$, since $\emptyset \models \{c = c\}$, we have $c = c$ in T by Lemma A.5. For any $c, d \in \mathcal{C}$, if $c = d$ is in T then $d = c$ is in T by Lemma A.5. For any $c, d, e, \in \mathcal{C}$, if $\{c = d, d = e\} \subseteq T$ then $c = e$ is in T by Lemma A.5. $\dashv_{\text{claim}}$

Now set $M = \mathcal{C}/\sim$ and, for $c \in \mathcal{C}$, let c^* denote $[c]_\sim \in M$. Since $|\mathcal{C}| \leq |\mathcal{L}| \leq \kappa$, we have $|M| \leq \kappa$. We construct an $\mathcal{L}$-structure $\mathcal{M}$, with universe M, such that $\mathcal{M} \models T$. First, we give the interpretation of the symbols in $\mathcal{L}$. Given a constant symbol c in $\mathcal{L}$, we set $c^M = c^*$. Now fix an n-ary relation symbol R in $\mathcal{L}$. Suppose $c_1, \ldots, c_n, d_1, \ldots, d_n \in \mathcal{C}$ are such that $c_i \sim d_i$ for all $1 \leq i \leq n$. Then $\{c_1 = d_1, \ldots, c_n = d_n\} \subseteq T$, and so it follows from Lemma A.5 that $R(\bar{c}) \in T$ if and only if $R(\bar{d}) \in T$. Therefore, we have a well-defined set

$$R^{\mathcal{M}} = \{(c_1^*, \ldots, c_n^*) \in M^n : R(c_1, \ldots, c_n) \in T\}.$$

Finally, fix an n-ary function symbol f in $\mathcal{L}$.

Claim 2: For any $c_1, \ldots, c_n \in \mathcal{C}$, there is $c_{n+1} \in \mathcal{C}$ such that $f(c_1, \ldots, c_n) = c_{n+1}$ is in T.
Proof: Let $\varphi(v)$ denote the $\mathcal{L}$-formula $v = f(c_1, \ldots, c_n)$, and set $c_{n+1} = c_\varphi$. We have $\emptyset \models \{\exists v\varphi(v)\}$, and so $\exists v\varphi(v) \in T$ by Lemma A.5. Combined with $(*)$, we have $\{\exists v\varphi(v) \rightarrow \varphi(c_{n+1}), \exists v\varphi(v)\} \subseteq T$, and so $\varphi(c_{n+1}) \in T$ by Lemma A.5. $\dashv_{\text{claim}}$

For any $c_1, \ldots, c_{n+1}, d_1, \ldots, d_{n+1} \in \mathcal{C}$, if $c_i \sim d_i$ for all $1 \leq i \leq n+1$ and $f(c_1, \ldots, c_n) = c_{n+1}$ is in T, then $f(d_1, \ldots, d_n) = d_{n+1}$ is in T by Lemma A.5. Combined with Claim 2, we have a well-defined function $f^{\mathcal{M}} : M^n \longrightarrow M$ such that

$$f^{\mathcal{M}}(c_1^*, \ldots, c_n^*) = c_{n+1}^* \quad \Leftrightarrow \quad f(c_1, \ldots, c_n) = c_{n+1} \text{ is in } T.$$

This finishes the definition of the $\mathcal{L}$-structure $\mathcal{M}$. It remains to show that $\mathcal{M} \models T$. In particular, we prove the following statement:

(†) for all $\mathcal{L}$-formulas $\varphi(v_1, \ldots, v_n)$ and $c_1, \ldots, c_n \in \mathcal{C}$, $\mathcal{M} \models \varphi(\bar{c}^*)$ if and only if $\varphi(\bar{c}) \in T$.

To do this, we first need a claim about terms.
Claim 3: For any $\mathcal{L}$-term $t(v_1, \ldots, v_n)$ and $c_1, \ldots, c_n, d \in \mathcal{C}$,

$$t^{\mathcal{M}}(\bar{c}^*) = d^* \quad \Leftrightarrow \quad t(\bar{c}) = d \text{ is in } T.$$

Proof: We first prove the forward direction by induction on terms. Suppose t is a constant symbol $c \in \mathcal{C}$ (and so $n = 0$). We want to show $c^* = d^*$ if and only if $c = d$ is in T, which is true by definition of $\sim$. Now suppose t is the variable v_1. We want to show $c_1^* = d^*$ if and only if $c_1 = d$ is in T, which is true for the same reason. Assume the result for terms $t_1, \ldots, t_m$ using free variables from among $v_1, \ldots, v_n$, and suppose f is an m-ary function symbol in $\mathcal{L}$.

Suppose $f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{c}^*), \ldots, t_m^{\mathcal{M}}(\bar{c}^*)) = d^*$. By definition of M, we may set $t_j^{\mathcal{M}}(\bar{c}^*) = d_j^*$, for some $d_1, \ldots, d_m \in \mathcal{C}$. Then $f^{\mathcal{M}}(d_1^*, \ldots, d_m^*) = d^*$, and so $f(d_1, \ldots, d_m) = d$ is in T by definition of $f^{\mathcal{M}}$. By induction $t_j(\bar{c}) = d$ is in T for all $1 \leq j \leq m$. Altogether,

$$\{t_1(\bar{c}) = d_1, \ldots, t_m(\bar{c}) = d_m\} \cup \{f(d_1, \ldots, d_m) = d\} \subseteq T,$$

and so $f(t_1(\bar{c}), \ldots, t_m(\bar{c})) = d$ is in T by Lemma A.5.

For the reverse direction, suppose $t(\bar{c}) = d$ is in T. By definition of M, we may set $t^{\mathcal{M}}(\bar{c}^*) = e^*$ for some $e \in \mathcal{C}$. By the forward direction of this claim, $t(\bar{c}) = e$ is in T, and so $\{t(\bar{c}) = e, t(\bar{c}) = d\} \subseteq T$. By Lemma A.5, $d = e$ is in T, and so $t^{\mathcal{M}}(\bar{c}^*) = e^* = d^*$. $\dashv_{\text{claim}}$

Finally, we prove (†) by induction on formulas. Suppose φ is the formula $t_1 = t_2$, for some $\mathcal{L}$-terms t_1 and t_2. Let $t_i^{\mathcal{M}}(\bar{c}^*) = d_i^*$, for $i \in \{1, 2\}$. Then

$$\begin{aligned} \mathcal{M} \models \varphi(\bar{c}^*) &\Leftrightarrow t_1^{\mathcal{M}}(\bar{c}^*) = t_2^{\mathcal{M}}(\bar{c}^*) \\ &\Leftrightarrow d_1^* = d_2^* \\ &\Leftrightarrow d_1 = d_2 \text{ is in } T \\ &\Leftrightarrow t_1(\bar{c}) = t_2(\bar{c}) \text{ is in } T, \end{aligned}$$

where the final equivalence follows from Lemma A.5, and the fact that $\{t_1(\bar{c}) = d_1, t_2(\bar{c}) = d_2\} \subseteq T$ by Claim 3.

Next, suppose φ is the formula $R(t_1, \ldots, t_m)$. Let $t_i^{\mathcal{M}}(\bar{c}^*) = d_i^*$ for some $d_1, \ldots, d_m \in \mathcal{C}$. Then

$$\begin{aligned} \mathcal{M} \models \varphi(\bar{c}^*) &\Leftrightarrow (d_1^*, \ldots, d_m^*) \in R^{\mathcal{M}} \\ &\Leftrightarrow R(d_1, \ldots, d_m) \in T \\ &\Leftrightarrow \varphi(\bar{c}) \in T, \end{aligned}$$

where the final equivalence follows from Lemma A.5, and the fact that $\{t_1(\bar{c}) = d_1, \ldots, t_1(\bar{c}) = d_m\} \subseteq T$ by Claim 3.

This finishes the verification of (†) for atomic $\mathcal{L}$-formulas. Assume the result for the formula $\varphi(v_1, \ldots, v_n)$ and fix $c_1, \ldots, c_n$ in $\mathcal{C}$. Then

$$\mathcal{M} \models \neg\varphi(\bar{c}^*) \quad \Leftrightarrow \quad \mathcal{M} \not\models \varphi(\bar{c}^*) \quad \Leftrightarrow \quad \varphi(\bar{c}) \notin T \quad \Leftrightarrow \quad \neg\varphi(\bar{c}) \in T,$$

where the second equivalence follows from induction, and the third equivalence follows from the fact that T is maximal and finitely satisfiable.

Next, assume the result for φ and ψ, and fix $c_1, \ldots, c_n \in \mathcal{C}$. Then

$$\mathcal{M} \models (\varphi \wedge \psi)(\bar{c}^*) \quad \Leftrightarrow \quad \mathcal{M} \models \varphi(\bar{c}^*) \text{ and } \mathcal{M} \models \psi(\bar{c}^*) \quad \Leftrightarrow \quad \{\varphi(\bar{c}), \psi(\bar{c})\} \subseteq T \quad \Leftrightarrow \quad (\varphi \wedge \psi)(\bar{c}) \in T,$$

where the second equivalence follows from induction, and the third equivalence from Lemma A.5.

Finally, assume the result for $\varphi(v_1, \ldots, v_n, w)$, and fix $c_1, \ldots, c_n \in \mathcal{C}$. Then

$$\mathcal{M} \models \exists w \varphi(\bar{c}^*, w) \iff \mathcal{M} \models \varphi(\bar{c}^*, d^*) \text{ for some } d^* \in M \iff \varphi(\bar{c}, d) \in T \text{ for some } d \in \mathcal{C},$$

where the second equivalence follows from induction. If $\varphi(\bar{c}, d) \in T$ for some $d \in \mathcal{C}$, then we have $\exists w \varphi(\bar{c}, w) \in T$ by Lemma A.5. On the other hand, if $\exists w \varphi(\bar{c}, w) \in T$ then, setting $d = c_\psi$ where $\psi(w)$ is the $\mathcal{L}$-formula $\varphi(\bar{c}, w)$, we have $\varphi(\bar{c}, d) \in T$ by Lemma A.5 and $(*)$. $\square$

The only added Skolem functions necessary for the previous proof were the constants (i.e. 0-ary function symbols) c_φ for $\varphi(v)$ a formula in one free variable. In other words, we only needed T to satisfy property $(*)$. This property is often called the *witness property*, and theories T satisfying the witness property are called *Henkinized*.

B Review: Sets, Cardinality, Algebra, Graphs

The following is a fairly terse list of definitions and facts that will be helpful for these notes. The material on sets, cardinality, and algebra should be fairly familiar, and the topics discussed in these sections will be used frequently in the summer school. The material on graphs will not be as heavily used, and prior exposure to these topics is not imperative.

B.1 Sets and Cardinality

We work in the ZFC axioms of set theory.[3]

Definition B.1. Fix sets X and Y.

1. A function $f : X \longrightarrow Y$ is **injective** (f is an **injection**) if, for all $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.
2. A function $f : X \longrightarrow Y$ is **surjective** (f is a **surjection**) if, for all $y \in Y$, there is some $x \in X$ such that $f(x) = y$.
3. A function $f : X \longrightarrow Y$ is **bijective** (f is a **bijection**) if it is injective and surjective.

Fact B.2. *The binary relation given by "there is a bijection from X to Y" is an equivalence relation on sets.*

Definition B.3. Given a set X, the **cardinality of** X, denoted $|X|$, is the equivalence class of X with respect to the equivalence relation in the previous fact.

Definition B.4. Fix sets X and Y.

1. $|X| \leq |Y|$ if there is an injection $f : X \longrightarrow Y$.
2. $|X| < |Y|$ if $|X| \leq |Y|$ and $|X| \neq |Y|$.

Fact B.5 (Cantor-Shröder-Bernstein)**.** *Given sets X and Y, $|X| = |Y|$ if and only if $|X| \leq |Y|$ and $|Y| \leq |X|$.*

Definition B.6. A **cardinal** is an equivalence class $|X|$ for some set X.

Let $\emptyset$ denote the set with no elements. Let $\mathbb{N}$ denote the set of natural numbers $\{0, 1, 2, 3, \ldots\}$. We use the symbol $\aleph_0$ to denote the cardinal $|\mathbb{N}|$. Given $n \in \mathbb{N}$, we identify n with the cardinal $|\{0, 1, \ldots, n-1\}|$ (in particular, 0 is identified with $|\emptyset|$). Note that if we restrict the ordering on cardinals to the elements of $\mathbb{N}$, then we recover the usual ordering of $\mathbb{N}$.

Definition B.7. A set X is **finite** if $|X| = n$ for some $n \in \mathbb{N}$. A set is **infinite** if it is not finite.

Fact B.8. *Suppose X and Y are finite sets of the same cardinality. Then a function $f : X \longrightarrow Y$ is injective if and only if it is surjective.*

Definition B.9. A set X is **countable** if $|X| \leq \aleph_0$.

Fact B.10. *If X is infinite then $|X| \geq \aleph_0$.*

Definition B.11. Given a cardinal $\kappa = |X|$, we let 2^κ denote the cardinality of the powerset of X (i.e. the set of all subsets of X).

[3] Ignore this if you don't know what it means.

Fact B.12 (Cantor)**.** *If κ is a cardinal then $\kappa < 2^\kappa$.*

Fact B.13. $2^{\aleph_0} = |\mathbb{R}|$.

Definition B.14. A **partial order** is a pair $(X, <)$ where X is a set and $<$ is irreflexive, antisymmetric, and transitive. A **chain** in $(X, <)$ is a subset $C \subseteq X$ such that $(C, <)$ is totally ordered. An upper bound in $(X, <)$ for a chain C is an element $x \in X$ such that $c \leq x$ for all $c \in C$. A **maximal element** for $(X, <)$ is an element $x \in X$ such that $y \leq x$ for all $y \in Y$.

Fact B.15 (Zorn's Lemma)**.** *Suppose $(X, <)$ is a nonempty partial order. If every chain in $(X, <)$ has an upper bound in X, then $(X, <)$ has a maximal element.*

Fact B.16. *If $\mathcal{C}$ is a nonempty collection of cardinals then $\mathcal{C}$ contains a minimal element, i.e. there is some $\kappa \in \mathcal{C}$ such that $\kappa \leq \lambda$ for all $\lambda \in \mathcal{C}$.*

Definition B.17. $\aleph_1$ is the smallest cardinal strictly greater than $\aleph_0$.

By Cantor's theorem, $\aleph_1 \leq 2^{\aleph_0}$.

Definition B.18. The **Continuum Hypothesis** is the assertion that $\aleph_1 = 2^{\aleph_0}$.

Fact B.19 (Gödel-Cohen)**.** *The Continuum Hypothesis is independent of ZFC.*

B.2 Groups

Definition B.20.

1. A **group** is a set G, together with a binary operation $*$ on G, and a distinguished element $e \in G$ such that:

   (i) $*$ is associative,

   (ii) for all $x \in G$, $e * x = x = x * e$,

   (iii) for all $x \in G$ there is a $y \in G$ such that $x * y = e = y * x$.

2. A group $(G, *, e)$ is **abelian** if $x * y = y * x$ for all $x, y \in G$.

Example B.21. The following are examples of groups.

1. $(\mathbb{Z}, +, 0)$, $(\mathbb{Q}, +, 0)$, and $(\mathbb{R}, +, 0)$. Note that $(\mathbb{N}, +, 0)$ is not a group since it fails axiom (iii).

2. $(\mathbb{R}^+, \cdot, 1)$, $(\mathbb{Q}^+, \cdot, 1)$.

3. Given $n > 0$, $\left(\mathbb{Z}/n\mathbb{Z}, +_n, 0\right)$ where $\mathbb{Z}/n\mathbb{Z} = \{0, 1, \ldots, n-1\}$ and $+_n$ is addition modulo n.

4. Given $n > 0$, $(\mathrm{GL}_n(\mathbb{R}), \cdot, I_n)$, where $\mathrm{GL}_n(\mathbb{R})$ is the set of $n \times n$ square matrices, with real entries and nonzero determinant, and I_n is the $n \times n$ identity matrix.

5. Given a set X, $(S_X, \circ, \mathrm{id}_X)$, where S_X is the set of *permutations of* X (i.e. bijections from X to itself), $\circ$ is composition of functions, and id_X is the identity function on X.

Each group in (1), (2), and (3) is abelian. The groups in (4) and (5) are not necessarily abelian.

Definition B.22. Let $(G, *_G, e_G)$ and $(H, *_H, e_H)$ be groups.

1. A function $f : G \longrightarrow H$ is a **group homomorphism** if $f(e_G) = e_H$ and, for all $x, y \in G$, $f(x *_G y) = f(x) *_H f(y)$.

2. An **isomorphism from** $(G, *_G, e_G)$ **to** $(H, *_H, e_H)$ is a bijective group homomorphism $f : G \longrightarrow H$.

3. $(G, *_G, e_G)$ and $(H, *_H, e_H)$ are **isomorphic** if there is an isomorphism from $(G, *_G, e_G)$ to $(H, *_H, e_H)$.

Definition B.23. Let $(G, *, e)$ be a group.

1. Given $x \in G$, the **order of** x is the minimum integer $n > 0$, if it exists, such that $x^n = e$ (where $x^n = x * x * \ldots * x$, n times). If no such n exists then x has **infinite order**.

2. An element $x \in G$ is a **torsion point** if it has finite order.

3. $(G, *, e)$ is **torsion-free** if e is the only torsion point.

4. $(G, *, e)$ is **cyclic** if there is some $x \in G$ such that, for all $y \in G$ there is $n > 0$ such that $y = x^n$.

We usually write abelian groups using additive notation $(G, +, 0)$. Given $x \in G$, we write nx for $x + x + \ldots + x$, n times.

Definition B.24. An abelian group $(G, +, 0)$ is **divisible** if, for all $x \in G$ and $n > 0$, there is some $y \in G$ such that $x = ny$.

B.3 Rings

Definition B.25. A **ring** is a set R, together with binary operations $+$ and $\cdot$ on R, and distinguished elements $0, 1 \in R$ such that:

(i) $(R, +, 0)$ is an abelian group,

(ii) $(R, \cdot, 1)$ is a semigroup with identity (i.e. satisfies (i) and (ii) of Definition B.20(1)),

(iii) for all $a, b, c \in R$

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$
$$(b + c) \cdot a = (b \cdot a) + (c \cdot a).$$

Definition B.26. Let $(R, +, \cdot, 0, 1)$ be a ring.

1. $(R, +, \cdot, 0, 1)$ is **commutative** if $a \cdot b = b \cdot a$ for all $a, b \in R$.

2. $(R, +, \cdot, 0, 1)$ is an **integral domain** if it is commutative and, for all $a, b \in R$, if $a \cdot b = 0$ then $a = 0$ or $b = 0$.

Example B.27. The following are examples of rings.

1. $(\mathbb{Z}, +, \cdot, 0, 1)$, $(\mathbb{Q}, +, \cdot, 0, 1)$, $(\mathbb{R}, +, \cdot, 0, 1)$, $(\mathbb{C}, +, \cdot, 0, 1)$.

2. $\left(\mathbb{Z}/n\mathbb{Z}, +_n, \cdot_n, 0, 1\right)$, where $n > 0$ and $+_n$, $\cdot_n$ are addition and multiplication modulo n, respectively.

3. $(M_n(\mathbb{R}), +, \cdot, 0_n, I_n)$, where $M_n(\mathbb{R})$ is the set of $n \times n$ square matrices with entries in $\mathbb{R}$, 0_n is the $n \times n$ matrix of all $0's$, and I_n is the $n \times n$ identity matrix.

Each ring in (1) is an integral domain.

Definition B.28. Let R be a commutative ring.

1. A subset $I \subseteq R$ is an **ideal** if it is a subgroup of $(R, +, 0)$, and $a \cdot b \in I$ for any $a \in R$ and $b \in I$.

2. An ideal $I \subseteq R$ is **radical** if, for any $a \in R$, if $a^n \in I$ for some $n > 0$ then $a \in I$.

3. An ideal $I \subseteq R$ is **prime** if, for any $a, b \in R$, if $a \cdot b \in I$ then $a \in I$ or $b \in I$.

Definition B.29. Let R be a commutative ring and $I \subseteq R$ an ideal. Define $R/I = \{[a] : a \in R\}$ where, given $a \in R$, $[a] = \{b \in R : a - b \in I\}$.

Fact B.30. *Let R be a commutative ring and $I \subseteq R$ an ideal.*

(a) $(R/I, +, \cdot, [0], [1])$ *is a commutative ring, where* $[a] + [b] = [a + b]$ *and* $[a] \cdot [b] = [a \cdot b]$.

(b) If I is a prime ideal, then R/I is an integral domain.

B.4 Fields

Definition B.31. A **field** is a set F, together with binary operations $+$ and $\cdot$ on F, and distinguished elements $0, 1 \in F$ such that:

(i) $(F, +, \cdot, 0, 1)$ is a commutative ring,

(ii) for all $a \in F$, if $a \neq 0$ then there is some $b \in F$ such that $a \cdot b = 1$.

Example B.32. The following are examples of fields.

1. $(\mathbb{Q}, +, \cdot, 0, 1)$, $(\mathbb{R}, +, \cdot, 0, 1)$, $(\mathbb{C}, +, \cdot, 0, 1)$.

2. $\mathbb{F}_p := \left({}^{\mathbb{Z}}\!/_{p\mathbb{Z}}, +_p, \cdot_p, 0, 1 \right)$, where p is a prime and $+_p$, $\cdot_p$ are addition and multiplication modulo p, respectively.

When working with fields, we often omit the symbol $\cdot$ and write the multiplicative operation as concatenation (i.e. $ab = a \cdot b$ for $a, b \in F$). We also identify the tuple $(F, +, \cdot, 0, 1)$ with F when there is no possibility for confusion.

Definition B.33. Let E and F be fields.

1. A function $\sigma : E \longrightarrow F$ is an **isomorphism** if σ is a group isomorphism from $(E, +, 0)$ to $(F, +, 0)$ and the restriction of σ to $E \backslash \{0\}$ is a group isomorphism from $(E \backslash \{0\}, \cdot, 1)$ to $(F \backslash \{0\}, \cdot, 1)$.

2. E and F are **isomorphic** if there is an isomorphism from E to F.

Definition B.34. Given a field F and variables $x_1, \ldots, x_n$, we let $F[x_1, \ldots, x_n]$ denote the set of polynomials in the variables $x_1, \ldots, x_n$ with coefficients in F.

Fact B.35. *If F is a field then $F[x_1, \ldots, x_n]$ is a commutative ring under usual addition and multiplication of polynomials.*

Given $p(x) \in F[x]$, we let $\deg(p)$ denote the degree of $p(x)$. The constant 0 polynomial has degree -∞ by convention. Any nonzero constant polynomial has degree 0.

Fact B.36. *Let F be a field.*

(*a*) *For any polynomials $p(x), q(x) \in F(x)$, with $q(x)$ nonzero, there are polynomials $r(x), s(x) \in F[x]$ such that $p(x) = q(x)s(x) + r(x)$ and $\deg(r) < \deg(q)$.*

(*b*) *If $p(x) \in F[x]$ and $a \in F$ is such that $p(a) = 0$, then $p(x) = (x - a)q(x)$ for some $q(x) \in F[x]$.*

Definition B.37. A field F is **algebraically closed** if, for any polynomial $p(x) \in F[x]$, there is some $a \in F$ such that $p(a) = 0$.

Fact B.38. *Every algebraically closed field is infinite.*

Example B.39. $(\mathbb{Q}^{alg}, +, \cdot, 0, 1)$ and $(\mathbb{C}, +, \cdot, 0, 1)$ are algebraically closed fields (where $\mathbb{Q}^{alg}$ is the set of algebraic numbers).

Fact B.40. *For any field F, there is a field F^{alg}, called the* ***algebraic closure of*** *F, which is the smallest (up to isomorphism) algebraically closed field containing F as a subfield.*

Example B.41. $(\mathbb{Q}^{alg}, +, \cdot, 0, 1)$ is the algebraic closure of $(\mathbb{Q}, +, 0, 1)$.

Fact B.42. *If F is a field and X is a subset of F then the intersection of all subfields of F containing X is a field, called the* ***subfield of*** *F* ***generated by*** *X.*

Definition B.43. Given a field F, the **prime subfield of** F is the subfield generated by $\emptyset$.

Definition B.44. Let F be a field. Define $\mathrm{ch}(F) \subseteq \mathbb{Z}^+$ to be the set of $n > 0$ such that

$$\underbrace{1 + 1 + \ldots + 1}_{n \text{ times}} = 0.$$

1. If $\mathrm{ch}(F) = \emptyset$ then F has **characteristic** 0.

2. If $\mathrm{ch}(F) \neq \emptyset$ then F has **characteristic** p, where p is the minimal element of $\mathrm{ch}(F)$.

Example B.45.

1. $(\mathbb{Q}, +, \cdot, 0, 1)$, $(\mathbb{R}, +, \cdot, 0, 1)$, and $(\mathbb{C}, +, \cdot, 0, 1)$ have characteristic 0.

2. $\mathbb{F}_p$ has characteristic p.

Fact B.46. *Let F be a field.*

(*a*) *If F has characteristic 0 then F is infinite and the prime subfield of F is isomorphic to $(\mathbb{Q}, +, \cdot, 0, 1)$.*

(*b*) *If F has characteristic $p > 0$ then p is prime and the prime subfield of F is isomorphic to $\mathbb{F}_p$.*

Fact B.47 (Finite fields)**.** *Fix a prime $p > 0$. There is a family $(\mathbb{F}_{p^n})_{n>0}$ of fields satisfying the following properties.*

(*i*) *For all $n > 0$, $\mathbb{F}_{p^n}$ has characteristic p and cardinality p^n.*

(*ii*) *For $m, n > 0$, $\mathbb{F}_{p^m}$ is a subfield of $\mathbb{F}_{p^n}$ if and only if m divides n.*

(*iii*) *If F is a finite field of characteristic p then F is isomorphic to $\mathbb{F}_{p^n}$ for some $n > 0$.*

(*iv*) $\mathbb{F}_p^{alg} := \bigcup_{n>0} \mathbb{F}_{p^n}$ *is the algebraic closure of* $\mathbb{F}_p$.

Definition B.48. Let F be a field.

1. Given a subfield E of F and a subset $X \subseteq E$, X is **algebraically independent over** E if, for any $a_1, \ldots, a_n \in X$ and any nonzero $p(x_1, \ldots, x_n) \in E[x_1, \ldots, x_n]$,
$$p(a_1, \ldots, a_n) \neq 0.$$

2. A subset $X \subseteq F$ is **algebraically independent** if it is algebraically independent over the prime subfield of F.

3. Given a subfield E of F, F is an **algebraic extension of** E if, for every $a \in F$, there is a polynomial $p(x) \in E[x]$ such that $p(a) = 0$.

4. A subset $X \subseteq F$ is a **transcendence basis for** F if X is algebraically independent and F is an algebraic extension of the subfield of F generated by X.

5. Given a subfield E of F, a subset $X \subseteq F$ is a **transcendence basis for** F **over** E if X is algebraically independent over E and F is an algebraic extension of the subfield of F generated by $E \cup X$.

Fact B.49. *If F is a field and $X, Y \subseteq F$ are both transcendence bases for F then $|X| = |Y|$.*

Definition B.50. The **transcendence degree** of a field F is the cardinality of a transcendence basis for F. If E is a subfield of F then the **transcendence degree of** F **over** E is the cardinality of a transcendence basis for F over E.

Fact B.51. *Let E and F be algebraically closed fields. Suppose $E_0 \subseteq E$ and $F_0 \subseteq F$ are subfields such that the transcendence degree of E over E_0 equals the transcendence degree of F over F_0. If $\sigma : E_0 \longrightarrow F_0$ is a field isomorphism, then σ extends to a field isomorphism $\hat{\sigma} : E \longrightarrow F$.*

Fact B.52. *Two algebraically closed fields are isomorphic if and only if they have the same characteristic and transcendence degree.*

Fact B.53. *Any integral domain R is a subring of a field. The smallest such field is called the* ***field of fractions of*** *R.*

Fact B.54. *Let F be a field.*

(*a*) (Hilbert Basis Theorem) *Any ideal in $F[\bar{x}]$ is finitely generated.*

(*b*) (Primary Decomposition) *If $I \subseteq F[\bar{x}]$ is a radical ideal then there are prime ideals $P_1, \ldots, P_m$ such that $I = P_1 \cap \ldots \cap P_m$.*

B.5 Vector Spaces

Fix a field F.

Definition B.55. A **vector space over** F is an abelian group $(V, \oplus, \mathbf{0})$, together with a function $F \times V \longrightarrow V$, called *scalar multiplication*, such that:

(i) for all $a, b \in F$ and $\boldsymbol{v} \in V$, $a(b\boldsymbol{v}) = (ab)\boldsymbol{v}$;

(ii) for all $a \in F$ and $\boldsymbol{v}, \boldsymbol{w} \in V$, $a(\boldsymbol{v} \oplus \boldsymbol{w}) = a\boldsymbol{v} \oplus a\boldsymbol{w}$;

(iii) for all $a, b \in F$ and $\boldsymbol{v} \in V$, $(a + b)\boldsymbol{v} = a\boldsymbol{v} \oplus b\boldsymbol{v}$;

(iv) for all $\boldsymbol{v} \in V$, $1\boldsymbol{v} = \boldsymbol{v}$.

Example B.56. For any field F and $n > 0$, $(F^n, \oplus, \mathbf{0})$ is a vector space over F, where $\oplus$ is coordinate addition of vectos and $\mathbf{0}$ is the tuple with every coordinate 0.

Definition B.57. Let V and W be vector spaces over F.

1. A function $\sigma : V \longrightarrow W$ is a **linear map** if σ is a group homomorphism from $(V, \oplus, \mathbf{0})$ to $(W, \oplus, \mathbf{0})$ and, for all $a \in F$ and $\boldsymbol{v} \in V$, $\sigma(a\boldsymbol{v}) = a\sigma(\boldsymbol{v})$.

2. V and W are **isomorphic** if there is a bijective linear map from V to W.

Definition B.58. Let V be a vector space over F.

1. A subset $X \subseteq V$ is **linearly independent** if, for all $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n \in V$ and $a_1, \dots, a_n \in F$,

$$a_1\boldsymbol{v}_1 \oplus \dots \oplus a_n\boldsymbol{v}_n = \mathbf{0} \quad \Leftrightarrow \quad a_i = 0 \text{ for all } 1 \leq i \leq n.$$

2. A subset $X \subseteq V$ is a **basis for** V if it is linearly independent and the subspace of V generated by X is all of V.

Fact B.59. *Suppose V and W are vector spaces of the same dimension. If $X \subseteq V$ and $Y \subseteq W$ are linearly independent subsets of the same cardinality, then any bijection from X to Y extends to a unique isomorphism from the subspace of V generated by X to the subspace of W generated by Y.*

Fact B.60. *If V is a vector space over F and $X, Y \subseteq V$ are both bases for V then $|X| = |Y|$.*

Definition B.61. The **dimension** of a vector space V is the cardinality of a basis for V.

Fact B.62. *Two vector spaces over F are isomorphic if and only if the have the same dimension.*

B.6 Graphs

Definition B.63. A **graph** is a set V together with a subset $E \subseteq V \times V$ such that:

(i) for all $v \in V$, $(v, v) \notin E$,

(ii) for all $v, w \in V$, if $(v, w) \in E$ then $(w, v) \in E$.

V is the set of **vertices** of the graph and E is the set of **edges** of the graph.

Visually, a graph (V, E) can be imagined as a collection of points V with a line drawn from v to w if (v, w) is an edge in E.

Definition B.64. Two graphs $\Gamma_1 = (V_1, E_1)$ and $\Gamma_2 = (V_2, E_2)$ are **isomorphic** if there is a bijection $f : V_1 \longrightarrow V_2$ such that, for all $v, w \in V_1$, $(v, w) \in E_1$ if and only if $(f(v), f(w)) \in E_2$.

Example B.65. Fix $n > 0$.

1. The **complete graph on n vertices** is the graph, denoted K_n, whose vertex set $V(K_n)$ has size n and whose edge set $E(K_n)$ is $(V \times V) \backslash \{(v,v) : v \in V\}$ (i.e. all possible edges).

2. The **empty graph on n vertices** is the graph, denoted $\overline{K}_n$, whose vertex set $V(\overline{K}_n)$ has size n and whose edge set is $\emptyset$.

Fact B.66 (Ramsey's Theorem)**.** *Given integers $m_1, m_2 > 0$ there is an integer $R(m_1, m_2)$ such that any graph of cardinality at least $R(m_1, m_2)$ either contains a complete graph of size m_1 or an empty graph of size m_2. Formally: for any graph $\Gamma = (V, E)$ with $|V| \geq R(m_1, m_2)$, there is a subset $W \subseteq V$ such that, if $\Gamma_0 = (W, E \cap (W \times W))$, then Γ_0 is either isomorphic to K_{m_1} or $\overline{K}_{m_2}$.*

Definition B.67. Given a graph $\Gamma = (V, E)$, an **edge-coloring** of Γ is a function $f : E \longrightarrow X$, where X is some set, such that, for any $(v, w) \in E$, $f((v, w)) = f((w, v))$.

Fact B.68 (Ramsey's Theorem (general form))**.** *Given integers $k > 0$ and $m_1, \ldots, m_k > 0$, there is an integer $R(m_1, \ldots, m_k)$ such that, for any $n \geq R(m_1, \ldots, m_k)$, if the edges of K_n are colored with k colors $\{1, \ldots, k\}$ then, for some $1 \leq t \leq k$, there there is copy of K_{m_t} all of whose edges are colored t. Formally: for any $n \geq R(m_1, \ldots, m_k)$ and any edge-coloring $f : E(K_n) \longrightarrow \{1, \ldots, k\}$, there is some $1 \leq t \leq k$ and a set $W \subseteq V(K_n)$ of cardinality m_t such that, for any distinct $v, w \in W$, $f((v, w)) = t$.*

Definition B.69. Let $\Gamma = (V, E)$ be a graph.

1. Given $k > 0$, a **k-coloring** of Γ is a function $c : V \longrightarrow \{1, \ldots, k\}$ such that if $(v, w) \in E$ then $f(v) \neq f(w)$.

2. Γ is **triangle-free** if there do not exists three distinct vertices $u, v, w \in V$ such that

$$(u, v), (v, w), (w, u) \in E.$$

3. Γ is **bipartite** if there is a partition $V = V_1 \cup V_2$ such that, for any $(v, w) \in E$, either $v \in V_1$ and $w \in V_2$ or $w \in V_1$ and $v \in V_2$.

4. Γ is **planar** if it can be drawn in the Euclidean plane in such a way that no two edges cross. Formally: there is an injection $f : V \longrightarrow \mathbb{R}^2$ and a family $(e_{(v,w)})_{(v,w) \in E}$ of functions such that:

   (i) for all $(v, w) \in E$, $e_{(v,w)} : [0, 1] \longrightarrow \mathbb{R}^2$ is a homeomorphism with $e_{(v,w)}(0) = f(v)$ and $e_{(v,w)}(1) = f(w)$;

   (ii) for all $(v_1, w_1), (v_2, w_2) \in E$, with $\{v_1, w_1\} \neq \{v_2, w_2\}$,

$$e_{(v_1,w_1)}((0, 1)) \cap e_{(v_2,w_2)}((0, 1)) = \emptyset.$$

Fact B.70.

(a) *A graph is bipartite if and only if it has a 2-coloring.*

(b) *Any bipartite graph is triangle-free.*

Fact B.71 (Four-Color Theorem)**.** *Any finite planar graph has a 4-coloring.*