INNER MODELS OF SET THEORY

1. Introduction

Set theory is supposed to provide a framework for encoding all of mathematics. Hence, it is desirable to know whether the accepted axioms of set theory are consistent, meaning that they do not contradict each other. In this regard, it is important to keep in mind that the currently accepted formulation of set theory is incredibly subtle. Statements such as Russell’s paradox serve as examples of contradictions that crept up in earlier versions of set theory; the modern approach professes to exorcise them, but how can we be sure that there are no other contradictions to worry about?

The short answer is, we can’t. According to the celebrated incompleteness theorem of Gödel, if, say, ZFC is consistent, then no proof of ZFC’s consistency can be carried out within ZFC itself. Hence, if we agree that all mathematics should, at least in principle, be doable in ZFC, then there is no “mathematical” justification for thinking that ZFC is consistent. As far as we are concerned, the consistency of ZFC remains an empirical fact: no one has found a contradiction in ZFC yet, hence it appears likely to be consistent.

Nevertheless, there are some nontrivial results concerning consistency of set theory that we can try to prove. For instance, we can isolate a particularly “suspicious” axiom and ask whether this axiom is consistent with the rest of set theory assuming that the rest of set theory is consistent on its own. For example, assuming ZF is consistent, can we argue that ZFC is consistent as well? It turns out that one can establish surprisingly strong results along these lines. For instance, we will prove in these notes that if ZF (i.e., “basic” set theory without Foundation or Choice) is consistent, then so is ZFC.

How can one prove such “relative consistency” results? Suppose, for example, that we want to show that the consistency of ZF implies the consistency of ZF. Let U be a universe of set theory satisfying ZF. We can then consider the class V, as defined in U. As such, V is a collection of sets equipped with the binary relation ∈, so it makes sense to ask which statements in the language of set theory V satisfies. Well, it turns out that V satisfies all of the axioms of ZF! Hence, ZF must be consistent. In these notes, we explore the power of arguments of this type.

2. Relativizing Formulas

Fix a universe U of set theory and let C be a class. Given a formula φ in the language of set theory, we write φC for the formula obtained from φ by restricting every quantifier to C, i.e., by replacing each quantifier of the form “∀x” (resp. “∃x”) by “∀x ∈ C” (resp. “∃x ∈ C”). If φ is a formula with no free variables and with parameters from C, we say that φ holds in C, or C satisfies φ, in symbols C ⊨ φ, if φC is true.

Example 2.1. Let A := {Ø, {Ø}, {Ø, {{Ø}}}}. Then A does not satisfy the Axiom of Extensionality Ext. Indeed, Ext is given by the following formula:

\[
\text{Ext} : \forall x \forall y (\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y).
\]

Thus, the relativized formula ExtA is

\[
\text{Ext}^A : \forall x \in A \forall y \in A (\forall z \in A (z \in x \leftrightarrow z \in y) \rightarrow x = y).
\]

1More precisely, his second incompleteness theorem.
In other words, $\text{Ext}^A$ says that if $x, y \in A$ have the same elements in $A$, then $x = y$. But this is false, since the distinct sets $x := \{\emptyset\}$ and $y := \{\emptyset, \{\emptyset\}\}$ both have a unique element in $A$, namely $\emptyset$. Another way in which $A$ “disagrees” with the ambient universe $\mathcal{U}$ is that, letting $x := \{\emptyset\}$ and $y := \{\emptyset, \{\emptyset\}\}$, we have $A \models y \subseteq x$. Indeed, “$y \subseteq x$” is really a shortcut for the formula
\[
y \subseteq x : \forall z (z \in y \rightarrow z \in x).
\]
Relativizing to $A$, we get
\[
(y \subseteq x)^A : \forall z \in A (z \in y \rightarrow z \in x).
\]
The only member of $A$ that is an element of $y$ is $\emptyset$, which also belongs to $x$, so $(y \subseteq x)^A$ is true.

To avoid the nastiness of Example 2.1, we usually restrict our attention to transitive classes. A class $C$ is transitive if for all $x \in C$, we have $x \subseteq C$; in other words, if $y \in x \in C$, then $y \in C$. Note that if $C$ is a set, then this agrees with our usual definition of a transitive set.

**Lemma 2.2.** If $\mathcal{U} \models \text{Ext}$ and $C$ is a transitive class, then $C \models \text{Ext}$ as well.

**Proof.** The formula $\text{Ext}^C$ looks like this:
\[
\text{Ext}^C : \forall x \in C \forall y \in C (\forall z \in C (z \in x \leftrightarrow z \in y) \rightarrow x = y).
\]
Take $x, y \in C$. Since $C$ is transitive, $x, y \subseteq C$. But this means that if for all $z$ in $C$, $z \in x \leftrightarrow z \in y$, then indeed for all $z$ in $\mathcal{U}$, $z \in x \leftrightarrow z \in y$, and hence $x = y$ by $\text{Ext}$ in $\mathcal{U}$. 

Many statements have the same “meaning” in $\mathcal{U}$ and in any transitive class $C$. Specifically, we say that $\varphi$ is a $\Delta_0$-formula if every quantifier in $\varphi$ is bounded, i.e., of the form “$\forall x \in y$” or “$\exists x \in y$”, where $x$ is a variable and $y$ is either a variable or a parameter. We often say (somewhat loosely) that a property or a relation is $\Delta_0$ if it can be expressed by a $\Delta_0$-formula. For example, the statement “$x \subseteq y$” can be expressed by the following $\Delta_0$-formula:
\[
x \subseteq y : \forall z \in x (z \in y).
\]
The utility of $\Delta_0$-formulas is captured by the following observation:

**Lemma 2.3.** If $\mathcal{C}$ is a transitive class and $\varphi$ is a $\Delta_0$-formula with no free variables and with parameters from $\mathcal{C}$, then $\mathcal{C} \models \varphi$ if and only if $\mathcal{U} \models \varphi$.

**Proof.** When a bounded quantifier such as “$\forall x \in y$” or “$\exists x \in y$” is restricted to $\mathcal{C}$, the result is “$\forall x \in \mathcal{C} \cap y$” or “$\exists x \in \mathcal{C} \cap y$.” But $\mathcal{C}$ is transitive, so for all $y \in \mathcal{C}$, $\mathcal{C} \cap y = y$. Hence, $\varphi^C$ has exactly the same meaning as $\varphi$.

As the following exercise shows, lots of useful properties can be expressed by $\Delta_0$-formulas.

**Exercise 2.4 (important!).** Assuming that $\mathcal{U} \models \text{ZF}^-$, show that the following properties and relations can be expressed by $\Delta_0$-formulas without parameters:

(a) $x = \emptyset$, $x \subseteq y$, $x = \{y, z\}$, $x = (y, z)$, $x = y \cap z$, $x = y \setminus z$, $x = y \times z$, $x = \bigcup y$;
(b) $f$ is a function, $y = f(x)$, $x = \text{dom}(f)$, $x = \text{ran}(f)$, $f$ is an injection, $f : x \rightarrow y$ is bijective;
(c) $<$ is a linear order on $x$, $x$ is a transitive set.

Additionally, assuming that $\mathcal{U} \models \text{ZF}$, show that the following properties and relations also can be expressed by $\Delta_0$-formulas without parameters:

(d) $\alpha$ is an ordinal, $\alpha$ is a limit ordinal, $\alpha$ is a successor ordinal;
(e) $n$ is a natural number, $x = \omega$.

Here $x, y, z, f, \alpha, n$ are to be treated as free variables.

The next exercise is somewhat long and tedious, but it will be quite useful, as it shows that for most axioms of $\text{ZF}^-$, it is rather clear how to check whether they hold in a given transitive class $\mathcal{C}$. 
**Exercise 2.5** (important!). Suppose $\mathcal{U} \models ZF^-$. Let $\mathcal{C}$ be a transitive class.

(a) Show that $\mathcal{C}$ satisfies the Empty Set Axiom if and only if $\emptyset \in \mathcal{C}$.
(b) Show that $\mathcal{C}$ satisfies the Pairing Axiom if and only if for all $x, y \in \mathcal{C}$, we have $\{x, y\} \in \mathcal{C}$.
(c) Show that $\mathcal{C}$ satisfies the Union Axiom if and only if for every $x \in \mathcal{C}$, we have $\bigcup x \in \mathcal{C}$.
(d) Show that $\mathcal{C}$ satisfies the Powerset Axiom if and only if for all $x \in \mathcal{C}$, we have $\{y \in \mathcal{C} : y \subseteq x\} \in \mathcal{C}$.

(e) Show that if $\omega \in \mathcal{C}$, then $\mathcal{C}$ satisfies the Infinity Axiom.
(f) Show that $\mathcal{C}$ satisfies the Comprehension Schema if and only if for all $x \in \mathcal{C}$ and for every formula $\varphi(z, \bar{a})$ with a single free variable $z$ and parameters $\bar{a} = (a_1, \ldots, a_k)$ from $\mathcal{C}$, we have $\{z \in x : \mathcal{C} \models \varphi(z, \bar{a})\} \in \mathcal{C}$.

Here you should notice that $\{z \in x : \mathcal{C} \models \varphi(z, \bar{a})\}$ is indeed a well-defined set in $\mathcal{U}$, because we can apply Comprehension in $\mathcal{U}$ to the relativized formula $\varphi^\mathcal{U}(z, \bar{a})$.

(g) Show that $\mathcal{C}$ satisfies the Replacement Schema if and only if the following statement holds. Let $\varphi(x, y, \bar{a})$ be a formula with two free variables $x, y$ and parameters $\bar{a} = (a_1, \ldots, a_k)$ from $\mathcal{C}$ such that $\varphi(x, y, \bar{a})$ defines a class function in $\mathcal{C}$; in other words,

$$\forall x \in \mathcal{C} \forall y \in \mathcal{C} \forall z \in \mathcal{C} (\varphi^\mathcal{C}(x, y, \bar{a}) \land \varphi^\mathcal{C}(x, z, \bar{a}) \rightarrow y = z).$$

Then for every set $X \in \mathcal{C}$, we have

$$\{y \in \mathcal{C} : \exists x \in X (\mathcal{C} \models \varphi(x, y, \bar{a}))\} \in \mathcal{C}.$$

Again, observe that $\{y \in \mathcal{C} : \exists x \in X (\mathcal{C} \models \varphi(x, y, \bar{a}))\}$ is a well-defined set in $\mathcal{U}$, because we can apply Replacement in $\mathcal{U}$ to the formula

$$x \in \mathcal{C} \land y \in \mathcal{C} \land \varphi^\mathcal{C}(x, y, \bar{a}),$$

which defines a class function in $\mathcal{U}$.

**Exercise 2.6** (important!). Recall that $AF$ is our notation for the Axiom of Foundation; that is,

$$AF : \forall x (x \neq \emptyset \rightarrow \exists y \in x (y \cap x = \emptyset)).$$

Suppose $\mathcal{U} \models AF$. Show that if $\mathcal{C}$ is any class (not necessarily transitive), then $\mathcal{C} \models AF$ as well.

A particularly useful type of potential universes of set theory are the so-called inner models, i.e., transitive classes $\mathcal{C}$ such that $Ord \subseteq \mathcal{C}$. The simplest interesting example of an inner model is the von Neumann universe $V$, which will be discussed in the next section. Another example that we’ll be working with is the so-called constructible universe, usually denoted by $L$.

**Example 2.7.** Consider the class $Ord$. We claim that $Ord$ satisfies the Powerset Axiom. Indeed, since $Ord$ is a transitive class, we can apply the result of Exercise 2.5(d) and only check that for each $\alpha \in Ord$, the following set is an ordinal:

$$\alpha' := \{\beta \in Ord : \beta \subseteq \alpha\}.$$  

But $\beta \subseteq \alpha$ if and only if $\beta \leq \alpha$, so $\alpha' = \{\beta \in Ord : \beta \leq \alpha\} = \alpha + 1$. In other words, from the point of view of $Ord$, the powerset of $\alpha$ is $\alpha + 1!$

**Exercise 2.8.** For a class $\mathcal{C}$, we let $\text{Ord}^\mathcal{C}$ and $\text{Card}^\mathcal{C}$ denote the ordinals and the cardinals “from the point of view of $\mathcal{C}$.” More precisely, define

$$\text{Ord}^\mathcal{C} := \{x \in \mathcal{C} : \mathcal{C} \models \text{“} x \text{ is an ordinal”}\} \quad \text{and} \quad \text{Card}^\mathcal{C} := \{x \in \mathcal{C} : \mathcal{C} \models \text{“} x \text{ is a cardinal”}\}.$$  

What are $\text{Ord}^{Ord}$ and $\text{Card}^{Ord}$?
3. The relative consistency of $\text{AF}$

**Figure 1.** The Axiom of Foundation holds in $V$.

**Theorem 3.1.** If $\mathcal{U} \models \text{ZF}^-$, then $V \models \text{ZF}$. Therefore, if $\text{ZF}^-$ is consistent, then so is $\text{ZF}$.

**Proof.** Let us first check that $V \models \text{ZF}^-$. Since $V$ is transitive, Exercise 2.5 makes this verification rather easy; the key observation is that for every set $x$, we have $x \in V$ if and only if $x \subseteq V$.

- $V \models \text{Ext}$ by Lemma 2.2.
- $V \models $ Empty: $\emptyset \in V$.
- $V \models $ Pair: If $x, y \in V$, then $\{x, y\} \subseteq V$, hence $\{x, y\} \in V$.
- $V \models $ Union: exercise!
- $V \models $ Pow: If $x \in V$, then $\mathcal{P}(x) \subseteq V$, so $\mathcal{P}(x) \in V$.
- $V \models $ Inf: $\omega \in V$.
- $V \models $ Comp: exercise!
- $V \models $ Rep: Let $\varphi(x, y, \vec{a})$ be a formula with free variables $x, y$ and parameters $\vec{a} = (a_1, \ldots, a_k)$ from $V$ such that $\varphi(x, y, \vec{a})$ defines a class function in $V$. We need to argue that for every $X \in V$, the following set belongs to $V$:

$$Y := \{ y \in V : \exists x \in X \ (V \models \varphi(x, y, \vec{a})) \}.$$

But $Y$ is, by definition, a subset of $V$, and hence $Y \in V$.

It remains to verify that $V \models \text{AF}$. Consider any $\emptyset \neq x \in V$. Let $\alpha := \min \{ \text{rank}(z) : z \in x \}$ and let $y \in x$ be an arbitrary element with $\text{rank}(y) = \alpha$. We claim that $x \cap y = \emptyset$, as desired. Indeed, if $z \in y$, then $\text{rank}(z) < \text{rank}(y) = \alpha$, while if $z \in x$, then $\text{rank}(z) \geq \alpha$ by the choice of $\alpha$. Hence, $y$ and $x$ have no elements in common. ■

**Exercise 3.2.** Show that if $\mathcal{U} \models \text{ZF}^- + \text{AC}$, then $V \models \text{ZFC}$.

**Assumption.** In the sequel, unless explicitly stated otherwise, we assume that $\mathcal{U} \models \text{ZF}$.

4. Inaccessible cardinals and models of set theory

**Assumption.** Throughout this section we assume that $\mathcal{U} \models \text{ZFC}$.

Recall that an infinite cardinal $\kappa$ is **regular** if $\text{cf}(\kappa) = \kappa$. A cardinal $\kappa$ is **(strongly) inaccessible** if $\kappa > \aleph_0$, $\kappa$ is regular, and $\kappa > 2^\lambda$ for every cardinal $\lambda < \kappa$. 
Exercise 4.1. Let $\kappa$ be an infinite cardinal. Show that $\kappa$ is regular if and only if a set of cardinality $\kappa$ cannot be written as a union of strictly fewer than $\kappa$-many sets of cardinality strictly less than $\kappa$.

Theorem 4.2. If $\kappa$ is an inaccessible cardinal, then $V_\kappa \models \text{ZFC}$. What makes Theorem 4.2 so remarkable is that $V_\kappa$ is merely a set, but it is so “large” that all of our usual set theory could be done inside of it!

Recall how in the proof of Theorem 3.1 we heavily relied on the fact that $x \in V \iff x \subseteq V$. To prove Theorem 4.2, we need a version of this property for $V_\kappa$:

Lemma 4.3. If $\kappa$ is an inaccessible cardinal, then $|V_\kappa| = \kappa$ and for every set $x$, we have

\[ x \in V_\kappa \iff x \subseteq V_\kappa \text{ and } |x| < \kappa. \quad (4.4) \]

Proof. Since $\kappa = \text{Ord} \cap V_\kappa \subseteq V_\kappa$, we have $|V_\kappa| \geq \kappa$. To prove the opposite inequality, recall that, by definition, $V_\kappa = \bigcup \{V_\alpha : \alpha < \kappa\}$. We will show that $|V_\alpha| < \kappa$ for all $\alpha < \kappa$. This implies that $V_\kappa$ is a union of $\kappa$-many sets of cardinality less than $\kappa$, and hence $|V_\kappa| \leq \kappa \otimes \kappa = \kappa$, as desired. The proof that $|V_\alpha| < \kappa$ for all $\alpha$ is by induction on $\alpha$.

Case 1: $\alpha = 0$. We have $|V_0| = |\varnothing| = 0 < \kappa$.

Case 2: $\alpha = \beta + 1$. We have

\[ |V_{\beta + 1}| = |\mathcal{P}(V_\beta)| = 2^{|V_\beta|} < \kappa, \]

since $\kappa$ is inaccessible and $|V_\beta| < \kappa$ by the inductive hypothesis.

Case 3: $\alpha$ is a limit. Then $V_\alpha = \bigcup \{V_\gamma : \gamma < \alpha\}$. In other words, $V_\alpha$ is a union of $|\alpha|$-many sets, each of cardinality less than $\kappa$. Since $\kappa$ is regular, this implies that $|V_\alpha| < \kappa$ (see Exercise 4.1).

Now we turn to proving (4.4). First, suppose that $x \in V_\kappa$. Since $V_\kappa$ is transitive, we conclude that $x \subseteq V_\kappa$. Furthermore, if we let $\alpha := \text{rank}(x)$, then $\alpha < \kappa$ and $x \subseteq V_\alpha$, so $|x| \leq |V_\alpha| < \kappa$.

Finally, suppose that $x \subseteq V_\kappa$ and $|x| < \kappa$. Since $\kappa$ is regular, the function $x \to \kappa : y \mapsto \text{rank}(y)$ cannot be cofinal. Hence, $\alpha := \sup\{\text{rank}(y) : y \in x\} < \kappa$. Then $x \subseteq V_\alpha$, so $x \in V_{\alpha + 1} \subseteq V_\kappa$.

Proof of Theorem 4.2. We will only verify that $V_\kappa$ satisfies Replacement and Choice, leaving the rest of the axioms as exercises.

To show that $V_\kappa \models \text{Rep}$, let $\varphi(x, y, \vec{a})$ be a formula with free variables $x$, $y$ and parameters $\vec{a} = (a_1, \ldots, a_k)$ from $V_\kappa$ such that $\varphi(x, y, \vec{a})$ defines a class function in $V_\kappa$. We need to argue that for every $X \in V_\kappa$, the following set belongs to $V_\kappa$:

\[ Y := \{y \in V_\kappa : \exists x \in X (V_\kappa \models \varphi(x, y, \vec{a}))\}. \]

By definition, $Y \subseteq V_\kappa$. Also, $Y$ is the image of $X$ under a function, so $|Y| \leq |X| < \kappa$. Therefore, $Y \in V_\kappa$ by Lemma 4.3.

To prove that $V_\kappa \models \text{AC}$, let $F \in V_\kappa$ be a set of nonempty sets. Since AC holds in $\mathcal{U}$, there is a choice function $c$ for $F$ in $\mathcal{U}$, and we only need to show that $c \in V_\kappa$. To that end, note that $c \subseteq V_\kappa$ (why?) and $|c| = |F| < \kappa$, hence $c \in V_\kappa$ by Lemma 4.3.

Taken together, Theorem 4.2 and Lemma 4.3 paint a very satisfying picture of the structure of the set $V_\kappa$ for inaccessible $\kappa$. First, $V_\kappa$ satisfies $\text{ZFC}$. Second, every class in $V_\kappa$ (i.e., a collection of elements of $V_\kappa$ defined by a formula) is a subset of $V_\kappa$ from the point of view of $\mathcal{U}$, but $V_\kappa$ “thinks” that it is a proper class whenever its cardinality (in $\mathcal{U}$) is $\kappa$.

Exercise 4.5. Let $\kappa$ be the least inaccessible cardinal. Show that

\[ V_\kappa \models \text{“there are no inaccessible cardinals.”} \]

Deduce that if $\text{ZFC}$ is consistent, then so is $\text{ZFC}$ together with the assertion that there are no inaccessible cardinals.

Exercise 4.6. Show that $V_\omega$ satisfies all of the axioms of $\text{ZFC}$ except one (which one?).
5. Talking about logic in \( U \)

5.A. What are we trying to do here?

Our long-term goal is to define another inner model, denoted by \( L \), which is “as thin as possible” and only includes those sets that are “necessary” for \( \text{ZF} \) to hold. What do we mean by that? Due to the Comprehension Schema, a universe of set theory must include, together with every set \( X \), each of its subsets of the form \( \{ x \in X : \varphi(x) \text{ true} \} \), where \( \varphi \) is some formula with one free variable \( x \). Calling such subsets of \( X \) “definable,” we would like to form the “definable powerset”

\[
\mathcal{D}(X) := \{ A \subseteq X : A \text{ is definable} \}.
\]

(We are using quotes because we haven’t defined what “definable” means yet.) Then the universe of “necessary” sets can be built recursively as follows:

\[
L_0 := \emptyset;
L_{\beta+1} := \mathcal{D}(L_\beta);
L_\alpha := \bigcup\{L_\gamma : \gamma < \alpha\} \text{ if } \alpha \text{ is a limit};
L := \bigcup\{L_\alpha : \alpha \in \text{Ord}\}.
\]

In other words, we mimic the construction of \( V \) but replace each powerset operation by its “definable” counterpart.

In order to carry out this construction, we have to be able to talk about formulas and definitions within \( U \). We develop the necessary machinery in this section, by showing how one can “encode” formulas by sets in \( U \). In hindsight, this shouldn’t be surprising: if we believe that all math can be done inside a universe of set theory, then this should include working with formulas and logic.

5.B. \( U \)-formulas

Fix a universe \( U \) of set theory that satisfies \( \text{ZF} \). A \( U \)-formula is a set in \( U \) that “represents” a formula in the language of set theory. The precise definition is recursive. First, we need to fix sets in \( U \) that are going to represent logical symbols. Say, let

\[
\vdash := 0, \quad \varepsilon := 1, \quad \land := 2, \quad \neg := 3, \quad \exists := 4.
\]

In other words, we shall use the numbers 0, 1, 2, 3, 4 to play the roles of the symbols “=,” “\( \in \),” “\( \land \),” “\( \neg \),” “\( \exists \),” and we put a dot above each symbol to indicate that we are talking about the corresponding member of \( U \).\(^2\) Similarly, let

\[
\text{Var} := \{ n \in \omega : n \geq 5 \}.
\]

We refer to the elements of \( \text{Var} \) as \( U \)-variables. They will be used as variables in our \( U \)-formulas. Note that the set of all \( U \)-variables is countably infinite. Define

\[
_0\mathcal{F} := \{ (\vdash, x, y), (\varepsilon, x, y) : x, y \in \text{Var} \}.
\]

The elements of the set \( _0\mathcal{F} \) are called atomic \( U \)-formulas. An atomic \( U \)-formula of the form \( (\vdash, x, y) \) represents the formula “\( x = y \),” while \( (\varepsilon, x, y) \) represents “\( x \in y \).” The rest of the definition is recursive. Namely, if the set \( _n\mathcal{F} \) for some \( n \in \omega \) is defined, then \( _{n+1}\mathcal{F} \) is given by

\[
_{n+1}\mathcal{F} := _n\mathcal{F} \cup \{ (\land, f, g), (\neg, f), (\exists, x, f) : f, g \in _n\mathcal{F}, x \in \text{Var} \}.
\]

Here \( (\land, f, g) \), \( (\neg, f) \), and \( (\exists, x, f) \) represent “\( f \land g \),” “\( \neg f \),” and “\( \exists x \, f \),” respectively. Finally, we set

\[
\mathcal{F} := \bigcup\{ _n\mathcal{F} : n \in \omega \},
\]

\(^2\) “Where are \( \lor, \rightarrow, \) and \( \forall \)?” I hear you ask in indignation. In the interest of saving space, we’ve got rid of them. For instance, instead of “\( \forall x \ldots \),” we can always write “\( \neg \exists x \ldots \).”
and refer to the elements of the set $F$ as $U$-formulas. The “$U$” in “$U$-formulas” is supposed to emphasize that $U$-formulas are sets in $U$. It should be clear from this definition that every “actual” formula $\phi$ can be naturally represented by a $U$-formula, which we denote by '$\phi$'. For instance, if $\phi$ is a formula asserting that $x = \emptyset$, i.e., $\phi = (\neg \exists y (y \in x))$, then we have

$$'\phi' = '\neg \exists y (y \in x)' = (\neg, (\exists, y, (\vdash, y, x)))).$$

The precise details of the above construction of $U$-formulas are not really important; what matters is that such a construction exists. In particular, to improve readability, we will usually write simply “$x = y$” instead of “(\vdash, x, y),” “$f \land g$” instead of “(\land, f, g),” &c. We will also use symbols such as “\vdash” and “\land” as obvious shortcuts.

Exercise 5.1. Show that every $U$-formula is a finite set. Moreover, show that $F \subseteq V_\omega$. Conclude that the set $F$ is countable. Hint: $V_\omega$ is countable.

Exercise 5.2 ($U$-formulas with parameters). For a set $W$, recursively define the set $F_W$ of all $U$-formulas with parameters from $W$. You should express $F_W$ as a union $F_W = \bigcup \{nF_W : n \in \omega\}$, where $0F_W$ is the set of all atomic $U$-formulas with parameters from $W$.

Exercise 5.3. Let $C$ be a class. Define the class $F_C$ of all $U$-formulas with parameters from $C$. Hint: Every individual $U$-formula involves only a (finite) set of parameters.

Exercise 5.4. Assuming AC, show that if $W$ is an infinite set, then $|F_W| = |W|$.

5.C. Induction on the complexity of $U$-formulas

Recall that the set $F$ of all $U$-formulas is defined as the union

$$F = \bigcup \{nF : n \in \omega\},$$

where $\{nF : n \in \omega\}$ is a recursively defined sequence starting with the set $0F$ of atomic $U$-formulas. For a $U$-formula $f \in F$, define its complexity $\text{comp}(f)$ as the least $n \in \omega$ such that $f \in nF$. One similarly defines the complexity of a $U$-formula with parameters.

A standard way of proving things related to $U$-formulas is by induction on their complexity. To illustrate this idea, let us formally define the set of all free variables in a $U$-formula $f$. We recursively define a sequence of functions $\text{FVar}_n : nF \to \mathcal{P}(\text{Var})$, for all $n \in \omega$, as follows. For $n = 0$, let

$$\text{FVar}_0(f) := \{x, y\} \text{ if } f = (x \vdash y) \text{ or } f = (x \in y).$$

If $\text{FVar}_n$ is already defined, then for each $f \in n+1F$, let

$$\text{FVar}_{n+1}(f) := \begin{cases} \text{FVar}_n(f) & \text{if } \text{comp}(f) \leq n; \\ \text{FVar}_n(g) \cup \text{FVar}_n(h) & \text{if } f = (g \land h); \\ \text{FVar}_n(g) & \text{if } f = (\neg g); \\ \text{FVar}_n(g) \setminus \{x\} & \text{if } f = (\exists x g). \end{cases}$$

(5.5)

Finally, let

$$\text{FVar}(f) := \text{FVar}_n(f) \text{ whenever } \text{comp}(f) = n,$$

and call $\text{FVar}(f)$ the set of all free variables in $f$. The above definition is usually phrased (somewhat loosely) in the following more concise form: We define the set $\text{FVar}(f) \subseteq \text{Var}$ as follows:

$$\text{FVar}(f) := \begin{cases} \{x, y\} & \text{if } f = (x \vdash y) \text{ or } f = (x \in y); \\ \text{FVar}(g) \cup \text{FVar}(h) & \text{if } f = (g \land h); \\ \text{FVar}(g) & \text{if } f = (\neg g); \\ \text{FVar}(g) \setminus \{x\} & \text{if } f = (\exists x g). \end{cases}$$

(5.6)
Exercise 5.7. Extend the above recursive definition of the set $FVar(f)$ to all $U$-formulas $f$ with parameters from a given set $W$.

Exercise 5.8. Give a recursive definition of the set $Sub(f)$ of all subformulas of a $U$-formula $f$. Here, a subformula of $f$ is one of the $U$-formulas $g$ that appear in the construction of $f$. For instance, 

$$Sub(\neg \exists y (y \in x)) = \{ \forall x, \exists y (y \in x), \neg \exists y (y \in x) \}.$$  

Exercise 5.9. Show that if $f$ is a $U$-formula, then the sets $FVar(f)$ and $Sub(f)$ are finite.

A $U$-formula $f$ such that $FVar(f) = \emptyset$ is called a $U$-sentence. The set of all $U$-sentences with parameters from a set $W$ is denoted by $F^0_W$. More generally, $F^k_W$ is the set of all $U$-formulas with parameters from $W$ and with exactly $k$ free variables, and similarly, $F^k$ is the set of all $U$-formulas with $k$ variables and no parameters (in other words, $F^k := F^k_\emptyset$). This is why we are using the somewhat clumsy notation “$nF$” instead of the more agreeable options “$F_n$” and “$F^n$”: both these expressions are reserved for other purposes, as $F_n$ is the set of all $U$-formulas with parameters from $n$, while $F^n$ is the set of all $U$-formulas with no parameters and with $n$ free variables.

Exercise 5.10. Let $W$ be a set. For a $U$-formula $f \in F^1_W$ and an element $a \in W$, give a recursive definition of the $U$-sentence $f(a) \in F^0_W$, obtained by plugging in $a$ instead of the free variable in $f$.

Note that if $C$ is a proper class, then we can also define the classes

$$nF_C := \bigcup\{nF_W : W \text{ is a subset of } C\}.$$  

However, one has to be extremely careful with recursive definitions involving such classes. For instance, look at our definition of the function $FVar$. Formally, instead of defining $FVar$ “in one go,” as in (5.6), we actually had to first construct a sequence of functions $FVar_n: nF \to P(Var)$ using (5.5). This sequence is itself a function defined on $\omega$ (sending each $n$ to $FVar_n$), so it is important that each $FVar_n$ is a set (which makes it an allowable value for a function). Imagine now that we were trying to do something similar on a proper class $F_C$. Then we would need each $FVar_n$ to be defined on the proper class $nF_C$, and so $FVar_n$ could not be a set and it would be meaningless to try to define the “function” $n \to FVar_n$!

This is a general feature of recursive definitions. It is important that the value of a recursively defined function on any given input is determined by the values it takes on some set of previously considered inputs. For example, when we recursively define a class function $\Phi: \text{Ord} \to U$, each $\Phi(\alpha)$ has to be expressed in terms of the values $\Phi(\beta)$ for $\beta < \alpha$, and all such $\beta$ form a set (namely $\alpha$ itself). Similarly, (5.6) reduces computing $FVar(f)$ for a $U$-formula $f$ with $\text{comp}(f) = n$ to knowing the values $FVar(g)$ for all $g$ of complexity less than $n$, and such $g$ form a set.

Of course, we actually can define the set of all free variables in a $U$-formula $f \in F_C$, because in fact $f \in F_W$ for some subset $W \subseteq C$. This might make the above discussion seem like unnecessary pedantry; however, in the next subsection we will discover an example when extending recursive definitions to formulas with arbitrary parameters becomes not just technically problematic, but in fact impossible.

Exercise 5.11. Extend the result of Exercise 5.10 by defining a class function

$$F^1_U \times U \to F^0_U: (f,a) \mapsto f(a).$$

5.D. Truth of $U$-formulas

Formulas are useful because they assert something. Similarly, we want to view $U$-formulas as statements that have meaning and hence can be true or false. Unfortunately, an attempt to define what it means for a $U$-formula to be true leads to unexpected difficulties.\(^3\)

\(^3\) Or maybe they should have been expected? If we could define truth, would philosophy departments still exist?
Lemma 5.12 (Truth is not definable). Let $\mathcal{F}_U^0$ be the class of all $U$-sentences with arbitrary parameters. There is no class function $\text{Truth}: \mathcal{F}_U^0 \to \{0, 1\}$ satisfying the following equation:

$$\text{Truth}(f) = \begin{cases} 
1 & \text{if } f = (a \equiv b) \text{ and } a = b; \\
1 & \text{if } f = (a \in b) \text{ and } a \in b; \\
1 & \text{if } f = (g \land h) \text{ and } \text{Truth}(g) = \text{Truth}(h) = 1; \\
1 & \text{if } f = (\lnot g) \text{ and } \text{Truth}(g) = 0; \\
1 & \text{if } f = (\exists x \, g) \text{ and } \exists a \, (\text{Truth}(g(a)) = 1); \\
0 & \text{otherwise}. 
\end{cases} \quad (5.13)$$

Equation (5.13) looks an awful lot like a definition by recursion on the complexity of $f$ (similar to (5.6)), so it might seem strange that, according to Lemma 5.12, it doesn’t actually define anything. But remember the discussion from the end of the last subsection! To compute the value $\text{Truth}(\exists x \, g)$ using (5.13), we have to first determine the values $\text{Truth}(g(a))$ for every $a \in U$, i.e., we need to already know a proper class of values of $\text{Truth}$. That is why (5.13) cannot be converted into a correct recursive definition (such as (5.5)). This observation, of course, doesn’t by itself prove that there is no other way to define a class function satisfying (5.13); but this we shall do now:

Proof. Suppose that $\text{Truth}$ is such a class function. Consider an arbitrary $U$-formula $f \in \mathcal{F}_U^0$. Since $f$ has one free variable, we can plug in any set $a$ into $f$ and obtain a $U$-sentence $f(a)$ (see Exercise 5.11), which should be either true or false, depending on whether $\text{Truth}(f(a))$ is 1 or 0. Since $f$ itself is a set, we can plug $f$ into itself and get a $U$-sentence $f(f)$. For some $f$, $\text{Truth}(f(f)) = 1$, while for others, $\text{Truth}(f(f)) = 0$. For example, if $f = 'x$ is a $U$-formula’, then $\text{Truth}(f(f)) = 1$, because $f$ is a $U$-formula. On the other hand, if $f = 'x$ is infinite’, then $\text{Truth}(f(f)) = 0$, because, like any $U$-formula, $f$ is finite (see Exercise 5.1). Now we define the following $U$-formula with one free variable, denoted below by $f$:

$$g := 'f \text{ is a } U\text{-formula with one free variable and } \text{Truth}(f(f)) = 0'. \quad (5.14)$$

The statement within the ‘…’ in (5.14) can be written out as a formula in the language of set theory—because, by our assumption, $\text{Truth}$ is a class function; therefore, the above definition of $g$ makes sense. And now the punchline: What is $\text{Truth}(g(g))$? If $\text{Truth}(g(g)) = 1$, then $g$ is a $U$-formula satisfying $\text{Truth}(g(g)) = 0$; a contradiction. But, conversely, if $\text{Truth}(g(g)) = 0$, then $g$ is a $U$-formula such that $\text{Truth}(g(g)) \neq 0$; a contradiction again.

In view of Lemma 5.12, you may wonder if there is any point in talking about $U$-formulas at all, given that there is no coherent way to define what it means for a $U$-formula to be true. Thankfully, there is still something we can do. Recall that the troubles in Lemma 5.12 stem from our desire to treat $U$-formulas with parameters ranging over the entire universe. We could be more modest and restrict our parameters and quantifiers to some set $W$; indeed, this restriction turns (5.13) into an appropriate recursive definition:

Theorem 5.15 (Truth in a model). There exists a unique class function $\text{Truth}$ with domain \{(W, f) : f \in \mathcal{F}_W^0 \} and range \{0, 1\}, such that for every set $W$ and for all $f \in \mathcal{F}_W^0$,

$$\text{Truth}(W, f) = \begin{cases} 
1 & \text{if } f = (a \equiv b) \text{ and } a = b; \\
1 & \text{if } f = (a \in b) \text{ and } a \in b; \\
1 & \text{if } f = (g \land h) \text{ and } \text{Truth}(W, g) = \text{Truth}(W, h) = 1; \\
1 & \text{if } f = (\lnot g) \text{ and } \text{Truth}(W, g) = 0; \\
1 & \text{if } f = (\exists x \, g) \text{ and } \exists a \in W \, (\text{Truth}(W, g(a)) = 1); \\
0 & \text{otherwise}. 
\end{cases} \quad (5.16)$$

\footnote{You can probably see where this is going.}
Proof. The restriction to the set $\mathcal{F}_W^0$ turns (5.16) into a valid definition by recursion on the complexity of $f$ that can be converted into a standard recursive definition analogous to (5.5). To really drive the point home, we do this explicitly here.

Given a set $W$, we recursively define functions $T_n: \mathcal{F}_W^n \to \{0, 1\}$, for all $n \in \omega$. For $n = 0$, let

$$T_0(f) := \begin{cases} 1 & \text{if } f = (a \vdash b) \text{ and } a = b; \\ 1 & \text{if } f = (a \in b) \text{ and } a \in b; \\ 0 & \text{otherwise}. \end{cases}$$

If $T_n$ is already defined, then for each $f \in \mathcal{F}_W^{n+1}$, define

$$T_{n+1}(f) := \begin{cases} T_n(f) & \text{if } \text{comp}(f) \leq n; \\ 1 & \text{if } f = (g \land h) \text{ and } T_n(g) = T_n(h) = 1; \\ 1 & \text{if } f = (\neg g) \text{ and } T_n(g) = 0; \\ 1 & \text{if } f = (\exists x g) \text{ and } \exists a \in W (T_n(g(a)) = 1); \\ 0 & \text{otherwise}. \end{cases}$$

Finally, we let $\text{Truth}(W, f) := T_n(f)$ whenever $\text{comp}(f) = n$. Checking that this definition satisfies the requirements of the theorem and that such a class function $\text{Truth}$ is unique are left as exercises. ■

Let $\text{Truth}$ be the class function from Theorem 5.15. When $\text{Truth}(W, f) = 1$, we say that $f$ holds in $W$, or $W$ satisfies $f$, and write $W \models f$. Note that if $\varphi$ is a formula, then the expressions $W \models \varphi$ and $W \models \neg \varphi$ are equivalent. The moral of this story can be summed up as follows:

There is no meaningful sense in which a $\mathcal{U}$-formula is true or false “in the universe”; but we can nonetheless talk about $\mathcal{U}$-formulas that hold or don’t hold in a given set.

6. Gödel’s constructible universe $L$

Figure 2. $L$ is an inner model with a lot of nice properties.

6.A. Definable powersets

Let $W$ be a set. A subset $A \subseteq W$ is definable in $W$ if there is a $\mathcal{U}$-formula $f \in \mathcal{F}_W^1$ such that

$$A = \{a \in W : W \models f(a)\}.$$ 

In this case, we say that $f$ defines $A$ in $W$. The definable powerset of $W$ is the set

$$\mathcal{D}(W) := \{A \subseteq W : A \text{ is definable in } W\}.$$
Prototypical examples of definable subsets of $W$ are of the form \{ $a \in W : W \models \varphi(a)$ \}, where $\varphi$ is a formula with one free variable and with parameters from $W$. Indeed, such a set is defined by the corresponding $\mathcal{U}$-formula $'\varphi'$. Some concrete examples of definable subsets of $W$ are:

- $W$ itself: $W = \{ a \in W : W \models a = a \}$;
- the empty set: $\varnothing = \{ a \in W : W \models \neg(a = a) \}$;
- every one-element subset of $W$: if $b \in W$, then
  \[ \{ b \} = \{ a \in W : W \models a = b \} \]
  (here $b$ is used as a parameter).

The last example can be generalized to show that every finite subset of $W$ is definable in $W$; indeed, if $k \in \omega$ and $a_1, \ldots, a_k \in W$, then

\[ \{ a_1, \ldots, a_k \} = \{ a \in W : W \models (a = a_1) \lor \ldots \lor (a = a_k) \}. \]

In particular, if $W$ is finite, then $\mathcal{D}(W) = \mathcal{P}(W)$. Not so if $W$ is infinite:

**Lemma 6.1.** Assume AC. If $W$ is an infinite set, then $|\mathcal{D}(W)| = |W|$. In particular, $\mathcal{D}(W) \neq \mathcal{P}(W)$.

**Proof.** Since every one-element subset of $W$ is in $\mathcal{D}(W)$, we have $|W| \leq |\mathcal{D}(W)|$. For the opposite inequality, notice that $|\mathcal{D}(W)| \leq |\mathcal{F}_W| \leq |\mathcal{F}_W|$, since the function

\[ \mathcal{F}_W \to \mathcal{D}(W) : f \mapsto \{ a \in W : W \models f(a) \} \]

is surjective. By Exercise 5.4, $|\mathcal{F}_W| = |W|$, and we are done. \hfill $\blacksquare$

**Exercise 6.2.** Show, without using AC, that if $W$ is countable, then $\mathcal{D}(W)$ is also countable.

**Exercise 6.3.** Show that if $A, B \in \mathcal{D}(W)$, then $A \cup B, A \cap B, A \setminus B \in \mathcal{D}(W)$.

Note that if $X \subseteq Y$, then $\mathcal{P}(X) \subseteq \mathcal{P}(Y)$ as well. For definable powersets, this property may fail. Indeed, if $X \notin \mathcal{D}(Y)$, then $\mathcal{D}(X) \nsubseteq \mathcal{D}(Y)$ as $X \in \mathcal{D}(X)$. Nevertheless, we have the following:

**Lemma 6.4.** Let $X$ and $Y$ be sets such that $X \subseteq Y$ and $X \in Y$. Then $\mathcal{D}(X) \subseteq \mathcal{D}(Y)$.

**Proof.** Take any $A \in \mathcal{D}(X)$. We want to show that $A \in \mathcal{D}(Y)$. Write

\[ A = \{ a \in X : X \models f(a) \} \]

for some $f \in \mathcal{F}_X$. We can (exercise!) recursively build a $\mathcal{U}$-formula $f^X$ by restricting the domain of every quantifier in $f$ to $X$. Then all the parameters in $f^X$ come from $Y$. (Note that $f^X$ uses $X$ as a parameter, but, thankfully, $X \in Y$.) This allows us to write

\[ A = \{ a \in Y : Y \models (a \in X) \land f^X(a) \} \in \mathcal{D}(Y). \]

\hfill $\blacksquare$

### 6.B. The definition and basic properties of $L$

Finally, we can define $L$. For each ordinal $\alpha$, we recursively build a set $L_\alpha$ as follows:

\[ \begin{align*}
L_0 & := \varnothing; \\
L_{\beta + 1} & := \mathcal{D}(L_\beta); \\
L_\alpha & := \bigcup \{ L_\gamma : \gamma < \alpha \} \text{ if } \alpha \text{ is a limit.}
\end{align*} \]

**Gödel’s constructible universe** is the class $L$ given by

\[ L := \bigcup \{ L_\alpha : \alpha \in \text{Ord} \}. \]

A set $x$ is called **constructible** if $x \in L$. The **order** of a constructible set $x$ is

\[ \text{order}(x) := \min \{ \alpha \in \text{Ord} : x \in L_\alpha \}. \]

(This is analogous to the definition of rank for sets in $V$.)
Lemma 6.5. Let \( \alpha \) be an ordinal. Then:

- \( L_\gamma \subseteq L_\alpha \) for all \( \gamma \leq \alpha \); and
- the set \( L_\alpha \) is transitive.

Therefore, \( L \) is a transitive class.

Proof. There is a reason why we combined these two statements in one lemma: we will prove them simultaneously by induction on \( \alpha \). The cases when \( \alpha = 0 \) and \( \alpha \) is a limit ordinal are clear (exercise!), so consider the case when \( \alpha = \beta + 1 \). By the inductive hypothesis, \( L_\gamma \subseteq L_\beta \) for all \( \gamma < \beta \), so to prove the first part of the lemma, we just need to show that \( L_\beta \subseteq L_{\beta + 1} \). Let \( A \in L_\beta \). By the inductive assumption, \( L_\beta \) is transitive, so \( A \subseteq L_{\beta + 1} \). Then we may use \( A \) as a parameter and obtain

\[
A = \{ a \in L_\beta : L_\beta \models a \in A \} \in \mathcal{D}(L_\beta) = L_{\beta + 1},
\]

as desired. To show that \( L_{\beta + 1} \) is transitive, consider any element \( A \in L_{\beta + 1} \). Then \( A \) is a definable subset of \( L_\beta \). But we have already shown that \( L_\beta \subseteq L_{\beta + 1} \), so \( A \subseteq L_\beta \subseteq L_{\beta + 1} \), and we are done. \( \blacksquare \)

Exercise 6.6. Show that for each \( \alpha \in \text{Ord} \), \( L_\alpha = \bigcup \{ \mathcal{D}(L_\gamma) : \gamma < \alpha \} \).

Exercise 6.7. Show that for every constructible set \( x \), order(\( x \)) is a successor ordinal, and if \( y \in x \), then \( y \) is also constructible with order(\( y \)) < order(\( x \)).

Exercise 6.8. Show that \( L_\alpha \subseteq V_\alpha \) for all \( \alpha \in \text{Ord} \), and hence \( L \subseteq V \).

Since for a finite set \( W \), \( \mathcal{D}(W) = \mathcal{P}(W) \), we conclude that

\[
L_n = V_n \text{ for all } n \in \omega,
\]

and hence also \( L_\omega = V_\omega \). But on the next level the two hierarchies separate, as \( L_{\omega + 1} \neq V_{\omega + 1} \), because \( L_{\omega + 1} = \mathcal{D}(L_\omega) \) is countable (see Exercise 6.2), while \( V_{\omega + 1} = \mathcal{P}(V_\omega) \) is not.

Lemma 6.9. Every ordinal \( \alpha \) is constructible with order(\( \alpha \)) = \( \alpha + 1 \). In other words, for all \( \alpha \in \text{Ord} \), we have \( \text{Ord} \cap L_\alpha = \alpha \).

Proof. The proof is by induction on \( \alpha \), so suppose that for all \( \gamma < \alpha \), we have \( \text{Ord} \cap L_\gamma = \gamma \). Since \( \text{Ord} \cap L_\alpha \subseteq \text{Ord} \cap V_\alpha = \alpha \), we only need to show that \( \alpha \subseteq L_\alpha \). In other words, given any \( \gamma < \alpha \), we have to argue that \( \gamma \in L_\alpha \). But

\[
\gamma = \text{Ord} \cap L_\gamma = \{ x \in L_\gamma : x \text{ is an ordinal} \}
\]

\[
= \{ x \in L_\gamma : L_\gamma \models \text{‘}x \text{ is an ordinal’} \} \in \mathcal{D}(L_\gamma) = L_{\gamma + 1} \subseteq L_\alpha,
\]

where we are using that being an ordinal can be defined by a \( \Delta_0 \)-formula (see Exercise 2.4(d)). \( \blacksquare \)

The proof of Lemma 6.9 can be summarized as follows: to construct the next ordinal, simply take the set of all ordinals that you have already constructed.

In view of Lemmas 6.5 and 6.9, we are justified in calling \( L \) an inner model: \( L \) is a transitive class and \( \text{Ord} \subseteq L \).

6.C. A preview of further results about \( L \)

Here we summarize the main properties of \( L \) that we will eventually demonstrate:

Theorem 6.10. If \( \mathcal{U} \models \text{ZF} \), then \( L \models \text{ZFC + GCH} \). Hence, if \( \text{ZF} \) is consistent, then so is \( \text{ZFC + GCH} \).

To advertise \( L \) even more, let us state a simple application of the above result. We say that a formula \( \varphi \) without parameters is \textit{arithmetical} if all of its quantifiers are bounded by \( V_\omega \), i.e., are of the form “\( \exists x \in V_\omega \)” or “\( \forall x \in V_\omega \)." The word “arithmetical" is used because \( V_\omega \) is a countable
set all of whose elements are finite, so arithmetical formulas are really statements about finite sets, and as such can be “encoded” as statements about natural numbers. Here are some examples of statements that can be formulated as arithmetical formulas:

**Example 6.11.** *Fermat’s Last Theorem* is the assertion that

\[ \forall x, y, z, n \in \omega \ (x, y, z, \geq 1 \land n \geq 3 \rightarrow x^n + y^n \neq z^n). \]

This statement was a notorious open problem for over 350 years (even though Pierre de Fermat claimed in 1637 that he had found “a truly marvelous proof of this, which this margin is too narrow to contain”), until it was finally proved by Andrew Wiles in 1994.

**Example 6.12.** Prime numbers \( p \) and \( q \) are called *twin primes* if \( |p - q| = 2 \). The *Twin Primes Conjecture* is the statement that there are infinitely many pairs of twin primes. It may seem “there are proved in...”

We begin our analysis by showing that stronger set-theoretic assumptions can simplify their solution. Recall, for example, this statement was a notorious open problem for over 350 years (even though Pierre de Fermat claimed in 1637 that he had found “a truly marvelous proof of this, which this margin is too narrow to contain”), until it was finally proved by Andrew Wiles in 1994.

**Example 6.13.** Another famous open problem that can be formulated as an arithmetical statement is the \( \mathsf{P} \) vs. \( \mathsf{NP} \) *Problem*. Roughly speaking, \( \mathsf{P} \neq \mathsf{NP} \) is the statement that a certain computational problem called \( \mathsf{SAT} \) (for “satisfiability”) cannot be solved by a computer program whose running time is polynomial in the size of the input. In contrast to the previous examples, it is less clear why this is an arithmetical statement; the idea is that a computer program is a finite piece of text and so can be represented by a finite set in \( V_\omega \). The problem of whether \( \mathsf{P} \neq \mathsf{NP} \) is one of the most celebrated open questions in mathematics. In particular, it is included on the list of the Millennium Prize Problems, so if you manage to solve it, the Clay Mathematics Institute will pay you $10^6.

Since many important open problems can be formulated as arithmetical statements, it is natural to wonder whether stronger set-theoretic assumptions can simplify their solution. Recall, for example, Goodstein’s theorem: it is an arithmetical statement about natural numbers, and yet it’s proof involves properties of infinite ordinals! However, Theorem 6.10 shows that the truth of arithmetical statements cannot be contingent on assumptions such as \( \mathsf{AC} \) or \( \mathsf{CH} \):

**Corollary 6.14.** Let \( \varphi \) be an arithmetical sentence. If \( \varphi \) can be proved in \( \mathsf{ZFC} + \mathsf{GCH} \), then \( \varphi \) can be proved in \( \mathsf{ZF} \) alone.

**Proof.** Suppose that \( \varphi \) is provable in \( \mathsf{ZFC} + \mathsf{GCH} \) and let \( \mathcal{U} \models \mathsf{ZF} \). By Theorem 6.10, \( L \models \mathsf{ZFC} + \mathsf{GCH} \), and so \( L \models \varphi \). But all quantifiers in \( \varphi \) are restricted to \( V_\omega \), and “\( V_\omega \)” has the same meaning in \( \mathcal{U} \) and in \( L \) (why?), so \( \varphi^L \) is equivalent to \( \varphi \). Hence, \( \mathcal{U} \models \varphi \), as desired.

\[ \square \]

6.D. \( L \) satisfies \( \mathsf{ZF} \); the Reflection Principle

We begin our analysis by showing that \( L \) satisfies all of the axioms of \( \mathsf{ZF} \) except for the Axiom Schemas of Comprehension and Replacement.

**Lemma 6.15.** If \( \mathcal{U} \models \mathsf{ZF} \), then \( L \models \mathsf{ZF} - \mathsf{Comp} - \mathsf{Rep} \).

**Proof.** The axioms \( \mathsf{Ext} \), \( \mathsf{Empty} \), \( \mathsf{Inf} \), and \( \mathsf{AF} \) are immediate.

- \( L \models \mathsf{Pair} \): Take \( x, y \in L \) and let \( \alpha := \max\{\text{order}(x), \text{order}(y)\} \). Then \( \{x, y\} \) is a finite (hence definable) subset of \( L_\alpha \), so \( \{x, y\} \in L_{\alpha+1} \).
- \( L \models \mathsf{Union} \): If \( x \in L_\alpha \), then, since \( L_\alpha \) is transitive, we have \( \bigcup x \subseteq L_\alpha \), and thus \( \bigcup x = \{z \in L_\alpha : L_\alpha \models \exists y \in x(z \in y)\} \in \mathcal{D}(L_\alpha) = L_{\alpha+1} \).
• \( L \models \text{Pow} \): Let \( x \in L \). We must argue that the following set is in \( L \):
\[
P := \{ y \in L : y \subseteq x \}.
\]
To that end, let \( \alpha := \max\{\text{order}(y) : y \in P\} \). Note that \( \alpha \) is a well-defined ordinal, because \( P \) is a set and each \( y \in P \) is constructible (so \( \text{order}(y) \) makes sense). Then \( P \subseteq L_\alpha \), so
\[
P = \{ y \in L_\alpha : y \subseteq x \} \in \mathcal{D}(L_\alpha) = L_{\alpha+1}.
\]

Next we want to show that \( L \) satisfies the Axiom Schema of Comprehension. Let \( X, a_1, \ldots, a_k \in L \) and let \( \varphi(x, a_1, \ldots, a_k) \) be a formula with a free variable \( x \) and parameters \( a_1, \ldots, a_k \). We want to prove that the following set is in \( L \):
\[
Y := \{ a \in X : L \models \varphi(a, a_1, \ldots, a_k) \}.
\]
Take \( \alpha \in \text{Ord} \) such that \( X, a_1, \ldots, a_k \in L_\alpha \) and define
\[
Y' := \{ a \in X : L_\alpha \models \varphi(a, a_1, \ldots, a_k) \} = \{ a \in L_\alpha : L_\alpha \models \varphi(a, a_1, \ldots, a_k) \land a \in X \} \in L_{\alpha+1}.
\]

It would be great if we could show that \( Y = Y' \), i.e., that \( L_\alpha \) and \( L \) “agree” in their opinions on the formula \( \varphi \). But how can we ensure that \( L_\alpha \) “knows” what \( L \) will “think” is true? The solution is provided by a general result known as the Reflection Principle.

First, we require some terminology. We say that \( \mathcal{C} = \bigcup\{ C_\alpha : \alpha \in \text{Ord} \} \) is a stratified class if \( \text{Ord} \to \mathcal{U} : \alpha \to C_\alpha \) is a class function such that:

- if \( \alpha \leq \beta \), then \( C_\alpha \subseteq C_\beta \);
- if \( \alpha \) is a limit ordinal, then \( C_\alpha = \bigcup\{ C_\gamma : \gamma < \alpha \} \).

Prototypical examples of stratified classes are
\[
V = \bigcup \{ V_\alpha : \alpha \in \text{Ord} \} \quad \text{and} \quad L = \bigcup \{ L_\alpha : \alpha \in \text{Ord} \}.
\]

We wish to prove that if \( \mathcal{C} = \bigcup\{ C_\alpha : \alpha \in \text{Ord} \} \) is a stratified class, then we can find an ordinal \( \beta \) such that the set \( C_\beta \) in some sense “reflects” the properties of the entire class \( \mathcal{C} \).

Let \( \mathcal{D} \subseteq \mathcal{C} \) be classes. Let \( \varphi(x_1, \ldots, x_k) \) be a formula with free variables \( x_1, \ldots, x_k \) and without parameters. We say that \( \varphi \) is absolute between \( \mathcal{D} \) and \( \mathcal{C} \) if for all \( a_1, \ldots, a_k \in \mathcal{D} \),
\[
\mathcal{D} \models \varphi(a_1, \ldots, a_k) \iff \mathcal{C} \models \varphi(a_1, \ldots, a_k).
\]

In other words, \( \varphi \) is absolute between \( \mathcal{D} \) and \( \mathcal{C} \) if \( \mathcal{D} \) and \( \mathcal{C} \) agree on the truth value of \( \varphi \) whenever it is applied to elements of \( \mathcal{D} \). Now we can state the aforementioned Reflection Principle:

**Theorem 6.16 (Reflection Principle).** Let \( \mathcal{C} = \bigcup\{ C_\alpha : \alpha \in \text{Ord} \} \) be a stratified class and let \( \varphi_1, \ldots, \varphi_n \) be a finite list of formulas without parameters. Then there is an ordinal \( \beta \) such that the formulas \( \varphi_1, \ldots, \varphi_n \) are absolute between \( C_\beta \) and \( \mathcal{C} \).

Note that Theorem 6.16 is really a “theorem schema,” in the sense that it cannot be phrased as a single formula in the language of set theory, since it involves quantification over finite lists of formulas. On a related note, it is impossible to replace “formulas” by “\( \mathcal{U} \)-formulas” in Theorem 6.16, because there is no sense in which \( \mathcal{C} \models f \) if \( f \) is a \( \mathcal{U} \)-formula and \( \mathcal{C} \) is a proper class (see §5.D).

**Example 6.17.** Suppose that \( \varphi_1, \ldots, \varphi_n \) are sentences satisfied by \( V \) (for instance, \( \varphi_1, \ldots, \varphi_n \) could be some of the axioms of ZF). By the Reflection Principle, there is an ordinal \( \beta \) such that \( \varphi_1, \ldots, \varphi_n \) are absolute between \( V_\beta \) and \( V \); in other words, \( V_\beta \models \varphi_1 \land \ldots \land \varphi_n \). Thus, we can always find \( V_\beta \) in which any given finite part of ZF holds! (Compare this to Theorem 4.2.)

The Reflection Principle is often applied in the form of the following corollary:

**Corollary 6.18.** Let \( \mathcal{C} = \bigcup\{ C_\alpha : \alpha \in \text{Ord} \} \) be a stratified class and let \( \varphi_1, \ldots, \varphi_n \) be a finite list of formulas without parameters. Then for every \( \gamma \in \text{Ord} \), there is \( \beta \geq \gamma \) such that \( \varphi_1, \ldots, \varphi_n \) are absolute between \( C_\beta \) and \( \mathcal{C} \).
Theorem 6.20. Without parameters. For simplicity, we will assume that the formulas are absolute between $C'_{\delta} = C_{\gamma+\delta}$ and $C$. Setting $\beta := \gamma + \delta$ finishes the proof. \hfill \blacksquare

Exercise 6.19. Deduce from Theorem 6.16 the following strengthening: Let $C = \bigcup \{ C_\alpha : \alpha \in \text{Ord} \}$ be a stratified class and let $\varphi_1, \ldots, \varphi_n$ be a finite list of formulas without parameters. Then there is an infinite cardinal $\kappa$ such that $\varphi_1, \ldots, \varphi_n$ are absolute between $C_\kappa$ and $C$. Hint: Define $C'_\alpha := C_\kappa$. With the help of the Reflection Principle, we can now finish the proof that $L \models \text{ZF}$:

Theorem 6.20. If $U \models \text{ZF}$, then $L \models \text{ZF}$.

Proof. Thanks to Lemma 6.15, it only remains to show that $L$ satisfies Comprehension and Replacement. We shall prove $L \models \text{Comp}$, leaving $L \models \text{Rep}$ as an exercise.

Let $X, a_1, \ldots, a_k \in L$ and let $\varphi(x, a_1, \ldots, a_k)$ be a formula with a free variable $x$ and parameters $a_1, \ldots, a_k$. As discussed earlier, we want to prove that the following set is in $L$:

$$Y := \{ a \in X : L \models \varphi(a, a_1, \ldots, a_k) \}.$$

Let $\gamma$ be an ordinal such that $X, a_1, \ldots, a_k \in L_\gamma$. By the Reflection Principle, there is an ordinal $\beta \geq \gamma$ such that $\varphi$ is absolute between $L_\beta$ and $L$. Then

$$Y = \{ a \in L_\beta : L \models \varphi(a, a_1, \ldots, a_k) \land a \in X \}$$

[by absoluteness]$$= \{ a \in L_\beta : L_\beta \models \varphi(a, a_1, \ldots, a_k) \land a \in X \} \in L_{\beta+1}.$$ \hfill \blacksquare

In the next subsection, we shall prove Theorem 6.16.

6.E. Proof of the Reflection Principle

Suppose that $C = \bigcup \{ C_\alpha : \alpha \in \text{Ord} \}$ is a stratified class and let $\varphi_1, \ldots, \varphi_n$ be a finite list of formulas without parameters. For simplicity, we will assume that the formulas $\varphi_i$ are constructed using only $\in, =, \land, \neg$, and $\exists$ as logical symbols (but not $\lor$ and $\forall$). (See footnote 2.) More importantly, we shall assume that the list $\varphi_1, \ldots, \varphi_n$ together with each formula $\varphi$ contains all of $\varphi$’s subformulas. For instance, if the formula

$$\exists y \ (y \in x \land \neg \exists z (z \in y))$$

is on the list, then so should be

$$y \in x \land \neg \exists z (z \in y), \quad y \in x, \quad \neg \exists z (z \in y), \quad \exists z (z \in y), \quad \text{and} \quad z \in y.$$

Finally, we reorder the list $\varphi_1, \ldots, \varphi_n$ in such a way that for each formula $\varphi_i$, all of its subformulas appear among $\varphi_1, \ldots, \varphi_i$. For instance, we could have $n = 6$ and

$$\varphi_1 = z \in y;$$
$$\varphi_2 = \exists z (z \in y);$$
$$\varphi_3 = \neg \exists z (z \in y);$$
$$\varphi_4 = y \in x;$$
$$\varphi_5 = y \in x \land \neg \exists z (z \in y);$$
$$\varphi_6 = \exists y \ (y \in x \land \neg \exists z (z \in y)).$$ (6.21)

Consider any $\alpha \in \text{Ord}$ and suppose that not all of $\varphi_1, \ldots, \varphi_n$ are absolute between $C_\alpha$ and $C$. Let $i$ be the smallest index such that $\varphi_i$ is not absolute between $C_\alpha$ and $C$ and set $\varphi := \varphi_i$. Suppose that $\varphi$ has $k$ free variables $x_1, \ldots, x_k$. By the choice of $\varphi$, all of $\varphi$’s subformulas, except for $\varphi$ itself, are absolute between $C_\alpha$ and $C$. What can $\varphi$ look like? It certainly cannot be a basic formula of the form $x = y$ or $x \in y$, because such formulas are always absolute. Neither can it be of the form
\( \varphi = \psi_1 \land \psi_2 \): if it were, then the subformulas \( \psi_1 \) and \( \psi_2 \) would be absolute between \( C_\alpha \) and \( C \), and thus, for all \( a_1, \ldots, a_k \in C_\alpha \), we would have
\[
C_\alpha \models \varphi(a_1, \ldots, a_k) \iff C_\alpha \models \psi_1(a_1, \ldots, a_k) \text{ and } C_\alpha \models \psi_2(a_1, \ldots, a_k) \iff C \models \psi_1(a_1, \ldots, a_k) \text{ and } C \models \psi_2(a_1, \ldots, a_k) \iff C \models \varphi(a_1, \ldots, a_k).
\]
Similarly, \( \varphi \) cannot be of the form \( \varphi = \neg \psi \), for then \( \psi \) would be absolute between \( C_\alpha \) and \( C \), and so \( \varphi \) would be absolute as well. Therefore, \( \varphi \) must be of the form
\[
\varphi(x_1, \ldots, x_k) = \exists x \psi(x, x_1, \ldots, x_k).
\]
(6.22)
Again, the subformula \( \psi \) is absolute between \( C_\alpha \) and \( C \), which yields that for all \( a_1, \ldots, a_k \in C_\alpha \),
\[
C_\alpha \models \varphi(a_1, \ldots, a_k) \iff \exists a \in C_\alpha \text{ such that } C_\alpha \models \psi(a, a_1, \ldots, a_k) \iff \exists a \in C_\alpha \text{ such that } C \models \psi(a, a_1, \ldots, a_k) \iff \exists a \in C \text{ such that } C \models \psi(a, a_1, \ldots, a_k) \iff C \models \varphi(a_1, \ldots, a_k).
\]
Thus, the only reason why \( \varphi \) is not absolute is that for some \( a_1, \ldots, a_k \in C_\alpha \), we have
\[
\exists a \in C \text{ such that } C \models \psi(a, a_1, \ldots, a_k), \quad \text{but} \quad \neg \exists a \in C_\alpha \text{ such that } C \models \psi(a, a_1, \ldots, a_k).
\]
This discussion motivates the following definition: Given a formula \( \varphi \) of the form (6.22) and elements \( a_1, \ldots, a_k \in C \), let \( W_{\varphi}(a_1, \ldots, a_k) \) denote the least ordinal \( \alpha \) such that
\[
\exists a \in C_\alpha \text{ with } C \models \psi(a, a_1, \ldots, a_k),
\]
if such \( \alpha \) exists (i.e., if \( C \models \varphi(a_1, \ldots, a_k) \)), and \( 0 \) otherwise. The letter \( \text{“} W \text{”} \) here stands for \( \text{“} \text{witness} \text{”} \): the truth of the statement \( \text{“} C \models \varphi(a_1, \ldots, a_k) \text{”} \) is witnessed by some \( a \in C_{W_{\varphi}(a_1, \ldots, a_k)} \). Given an ordinal \( \alpha \), define
\[
W_{\varphi}(\alpha) := \sup\{W_{\varphi}(a_1, \ldots, a_k) : a_1, \ldots, a_k \in C_\alpha\},
\]
and let
\[
W(\alpha) := \max \{W_{\varphi}(\alpha) : \varphi \text{ is on the list } \varphi_1, \ldots, \varphi_n\}.
\]
(6.23)
At this point, we should emphasize that the expression on the right-hand side of (6.23) is defined by listing explicitly the relevant formulas (rather than by quantifying over them, which is not allowed).
For instance, if our list of formulas were given by (6.21), then we would write
\[
W(\alpha) := \max\{W_{\exists z (z \in y)}(\alpha), W_{\exists y (y \in x \land \neg \exists z (z \in y))}(\alpha)\}.
\]
Now we recursively define ordinals \( \beta_m, m \in \omega \), by setting
\[
\beta_0 := 0, \quad \text{and} \quad \beta_{m+1} := W(\beta_m) \text{ for all } m \in \omega.
\]
Finally, let
\[
\beta := \sup\{\beta_m : m \in \omega\}.
\]
We claim that the formulas \( \varphi_1, \ldots, \varphi_n \) are absolute between \( C_\beta \) and \( C \), as desired. As observed above, we only have to argue that if \( \varphi \) is one of the formulas among \( \varphi_1, \ldots, \varphi_n \) of the form (6.22), then for all \( a_1, \ldots, a_k \in C_\beta \),
\[
C \models \varphi(a_1, \ldots, a_k) \quad \Longrightarrow \quad \exists a \in C_\beta \text{ such that } C \models \psi(a, a_1, \ldots, a_k).
\]
(6.24)
To that end, notice that \( 0 = \beta_0 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta \) (exercise!), so
\[
C_\beta = \bigcup\{C_{\beta_m} : m \in \omega\}.
\]
This means that for any \( a_1, \ldots, a_k \in C_\beta \), there is some \( m \in \omega \) with \( a_1, \ldots, a_k \in C_{\beta_m} \). Assuming that \( C \models \varphi(a_1, \ldots, a_k) \), this implies that there is some \( a \in C_{\beta_{m+1}} \subseteq C_\beta \) satisfying \( C \models \psi(a, a_1, \ldots, a_k) \), which yields (6.24) and finishes the proof.
Exercise 6.25. Pinpoint exactly every place in this argument where we used the assumption that $\mathcal{C} = \bigcup\{C_\alpha : \alpha \in \text{Ord}\}$ is a stratified class.

6.F. $\Sigma_1$-formulas and the Axiom of Constructibility

The Axiom of Constructibility is the assertion that every set is constructible, i.e., that $\mathcal{U} = L$. The goal of this section is to prove that the Axiom of Constructibility holds in $L$:

Theorem 6.26. $L \models \text{"every set is constructible."}$

Theorem 6.26 can be understood as follows. Consider a universe $\mathcal{U}$ satisfying ZF. Then, by Theorem 6.20, the class $L$, as defined in $\mathcal{U}$, also satisfies ZF. This means that the definition of $L$ can be interpreted inside $L$, producing some subclass $L' \subseteq L$. In other words, $L'$ is $L$’s version of $L$: it is the class of all sets $x \in L$ such that $L \models \text{"x is constructible."}$ Theorem 6.26 then asserts that, in fact, $L' = L$; or, using the terminology developed in §6.D, the property of being constructible is absolute between $L$ and $\mathcal{U}$.

How can we prove Theorem 6.26? It would be ideal if the statement “$x$ is constructible” were equivalent to a $\Delta_0$-formula, because then it would automatically be absolute between $L$ and $\mathcal{U}$. Unfortunately, it is not necessarily equivalent to a $\Delta_0$-formula; however, it is equivalent to a formula in a somewhat wider class—namely to a $\Sigma_1$-formula.

Let $\varphi$ be a formula in the language of set theory. We say that $\varphi$ is a $\Sigma_1$-formula if it is obtained from a $\Delta_0$-formula by adding a sequence of existential and bounded universal quantifiers. For instance, if $\psi$ is a $\Delta_0$-formula, then

$$\exists x \exists y \forall z \in y \exists u \forall v \in u \exists w \psi$$

is a $\Sigma_1$-formula. Also, a $\Pi_1$-formula is a formula that is obtained from $\Delta_0$-formula by adding a sequence of universal and bounded existential quantifiers. For instance, if $\psi$ is $\Delta_0$, then

$$\forall x \forall y \exists z \in y \forall u \exists v \in u \forall w \psi$$

is $\Pi_1$. Observe that the negation of any $\Sigma_1$-formula is equivalent to a $\Pi_1$-formula; conversely, every $\Pi_1$-formula is equivalent to the negation of some $\Sigma_1$-formula.

Example 6.27. The statement “$x$ is finite” can be expressed by a $\Sigma_1$-formula:

$$x \text{ is finite} \iff \exists n \exists f (n \in \omega \text{ and } f \text{ is a bijection from } n \text{ to } x).$$

Assuming AC, it can also be expressed by a negation of a $\Sigma_1$-formula (and hence by a $\Pi_1$-formula):

$$x \text{ is finite} \iff \neg \exists f (f \text{ is an injection from } \omega \text{ to } x).$$

Example 6.28. The statement “$x$ is countable” can be expressed by a $\Sigma_1$-formula:

$$x \text{ is countable} \iff \exists f (f \text{ is an injection from } x \text{ to } \omega).$$

Exercise 6.29. Let $\varphi$ and $\psi$ be $\Sigma_1$-formulas (resp. $\Pi_1$-formulas). Show that $\varphi \land \psi$ and $\varphi \lor \psi$ are also equivalent to $\Sigma_1$-formulas (resp. $\Pi_1$-formulas).

Example 6.30. The statement “$\kappa$ is a cardinal” can be expressed by a $\Pi_1$-formula:

$$\kappa \text{ is a cardinal} \iff \kappa \in \text{Ord} \land \neg \exists \alpha \in \kappa \exists f (f \text{ is a bijection from } \alpha \text{ to } \kappa).$$

---

6Often it is stated as “$V = L$,” because $\mathcal{U} = V$ by AF.
Example 6.31. The statement “\( y = \mathcal{P}(x) \)” can be expressed by a \( \Pi_1 \)-formula:
\[
y = \mathcal{P}(x) \iff \forall z (z \in y \iff z \subseteq x).
\]

Exercise 6.32. Show that the statement “\( R \) is a well-ordering on a set \( S \)” can be expressed both by a \( \Sigma_1 \)-formula and by a \( \Pi_1 \)-formula.

Exercise 6.33. Assuming \( \text{AC} \), show that every \( \Sigma_1 \)-formula is equivalent to a formula of the form \( \exists x \psi \), where \( \psi \) is \( \Delta_0 \).

In contrast to \( \Delta_0 \)-formulas, \( \Sigma_1 \)-formulas may not have the same truth value in \( \mathcal{U} \) and in a transitive class \( \mathcal{C} \). However, if a transitive class \( \mathcal{C} \) satisfies a \( \Sigma_1 \)-formula \( \varphi \), then \( \varphi \) holds in \( \mathcal{U} \) as well:

Exercise 6.34 (important!). Show that if \( \mathcal{C} \) is a transitive class and \( \varphi \) is a \( \Sigma_1 \)-formula with no free variables and with parameters from \( \mathcal{C} \), then
\[
\mathcal{C} \models \varphi \implies \mathcal{U} \models \varphi.
\]

Example 6.35. Suppose that \( \mathcal{C} \) is a transitive class containing \( \text{Ord} \). Then \( \text{Card}^\mathcal{C} \supseteq \text{Card} \), because the statement “\( x \) is not a cardinal” is \( \Sigma_1 \).

We say that a class \( \mathcal{C} \) is \( \Sigma_1 \) (resp. \( \Pi_1 \)) if there exists a \( \Sigma_1 \)-formula (resp. a \( \Pi_1 \) formula) \( \varphi(x) \) without parameters such that \( x \in \mathcal{C} \iff \varphi(x) \). Similarly, we say that a class function \( \Phi \) is \( \Sigma_1 \) (resp. \( \Pi_1 \)) if there is a \( \Sigma_1 \)-formula (resp. a \( \Pi_1 \) formula) \( \varphi(x,y) \) without parameters such that \( \Phi(x) = y \iff \varphi(x,y) \). Unless explicitly stated otherwise, we always assume that the “background theory” is \( \text{ZF} \). For instance, the class of all finite sets is \( \Sigma_1 \), but it is not \( \Pi_1 \) unless we additionally assume the Axiom of Choice (see Example 6.27).

Lemma 6.36. Let \( \Phi \) be a \( \Sigma_1 \)-class function. Suppose that \( \text{dom}(\Phi) \) is a \( \Pi_1 \)-class. Then \( \Phi \) is also \( \Pi_1 \).

Proof. Since \( \Phi \) is a class function, for every \( x \in \text{dom}(\Phi) \), there is precisely one set \( y \) such that \( \Phi(x) = y \). Hence, we can write
\[
\Phi(x) = y \iff x \in \text{dom}(\Phi) \land \forall z (z = y \lor \exists \underbrace{\underbrace{x \in \text{dom}(\Phi)}_{\Pi_1} \land \forall z (z = y \lor \neg (\underbrace{\Phi(x) = z}_{\Sigma_1})))}.
\]

Before moving on, let us make a few observations that will help us show that certain classes and class functions are \( \Sigma_1 \). We say that a set \( A \) is \( \Sigma_1 \)-identifiable if the class \( \{A\} \) is \( \Sigma_1 \); in other words, if there is a \( \Sigma_1 \)-formula \( \varphi(x) \) without parameters such that
\[
x = A \iff \varphi(x).
\]

For instance, \( \omega \) is \( \Sigma_1 \)-identifiable, since the property “\( x = \omega \)” can be expressed by a \( \Delta_0 \)-formula (and every \( \Delta_0 \)-formula is \( \Sigma_1 \)). If \( A \) is a \( \Sigma_1 \)-identifiable set, then quantifiers ranging over \( A \) (i.e., of the form “\( \exists x \in A \)” and “\( \forall x \in A \)” can be used in \( \Sigma_1 \)-definitions:
\[
\exists x \in A (\ldots) \iff \exists z (z = A \land \exists x \in z (\ldots));
\]
\[
\forall x \in A (\ldots) \iff \exists z (z = A \land \forall x \in z (\ldots)).
\]

In particular, we can freely quantify over \( \omega \) when defining \( \Sigma_1 \)-classes and \( \Sigma_1 \)-class functions. Another observation is that if \( \Phi \) is a \( \Sigma_1 \) class function (i.e., the statement “\( \Phi(x) = y \)” is given by a \( \Sigma_1 \)-formula without parameters), then the statement “\( z \in \Phi(x) \)” can be expressed by a \( \Sigma_1 \)-formula as follows:
\[
z \in \Phi(x) \iff \exists y (\Phi(x) = y \land z \in y).
\]

Exercise 6.37. Let \( \Phi \) and \( \Psi \) be \( \Sigma_1 \)-functions. Show that the composition \( \Phi \circ \Psi \) is also \( \Sigma_1 \); i.e., the statement “\( z = \Phi(\Psi(x)) \)” is equivalent to a \( \Sigma_1 \)-formula without parameters.
Our next observation is that \textit{recursive definitions} are naturally $\Sigma_1$. To make this statement precise, we need to recall some terminology. Let $H$ be a class function. A function $f$ is called \textit{$H$-inductive} if the domain of $f$ is an ordinal $\alpha := \text{dom}(f)$ and, for all $\beta < \alpha$, we have $f(\beta) = H(f|\beta)$. We also say that a class function $F: \text{Ord} \to U$ is \textit{$H$-inductive} if $F(\beta) = H(F|\beta)$ for all $\beta \in \text{Ord}$; i.e., if for all $\alpha \in \text{Ord}$, the function $F|\alpha$ is $H$-inductive. The class version of the Transfinite Recursion Theorem then says that if every $H$-inductive function belongs to $\text{dom}(H)$, then there is a unique $H$-inductive class function $F: \text{Ord} \to U$. It turns out that if $H$ is $\Sigma_1$, then $F$ is $\Sigma_1$ as well:

\textbf{Theorem 6.38 (\(\Sigma_1\)-Recursion).} Let $H$ be a $\Sigma_1$-class function such that every $H$-inductive function belongs to $\text{dom}(H)$. Then the unique $H$-inductive class function $F: \text{Ord} \to U$ is also $\Sigma_1$.

\textbf{Proof.} We have

$$F(\alpha) = y \iff \exists f \left( f \text{ is a function } \land \text{dom}(f) = \alpha + 1 \land f \text{ is } H\text{-inductive } \land f(\alpha) = y \right).$$

It remains to check that \textquotedblleft $f$ is $H$-inductive\textquotedblright{} is a $\Sigma_1$-statement. By definition,

$$f \text{ is } H\text{-inductive } \iff \forall \beta \leq \alpha \left( f(\beta) = H(f|\beta) \right) \iff \forall \beta \in \alpha \left( f(\beta) = H(f|\beta) \land f(\alpha) = H(f|\alpha) \right).$$

At this point, we just have to show that the statement \textquotedblleft $f(\beta) = H(f|\beta)$\textquotedblright{} is $\Sigma_1$. This follows from Exercise 6.37; explicitly, we can write

$$f(\beta) = H(f|\beta) \iff \exists g \exists z \left( g = f|\beta \land z = H(g) \land f(\beta) = z \right).$$

\textbf{Exercise 6.39.} State and prove the version of Theorem 6.38 for recursion up to $\omega$.

\textbf{Example 6.40.} For a set $x$, let $\text{cl}(x)$ denote the transitive closure of $x$, i.e., the smallest transitive set $y$ such that $x \subseteq y$. The class function $x \mapsto \text{cl}(x)$ is clearly $\Pi_1$:

$$y = \text{cl}(x) \iff y \text{ is transitive } \land x \subseteq y \land \forall z \left( (z \text{ is transitive } \land x \subseteq z) \implies y \subseteq z \right).$$

On the other hand, the transitive closure of $x$ can be defined recursively, showing that the class function $x \mapsto \text{cl}(x)$ is $\Sigma_1$. Indeed, if we set

$$f(0) := x, \quad \text{and } f(n + 1) := \bigcup f(n) \text{ for all } n \in \omega,$$

then $\text{cl}(x) = \bigcup \{ f(n) : n \in \omega \}$, so we may invoke Theorem 6.38 (or rather Exercise 6.39) to conclude that $x \mapsto \text{cl}(x)$ is a $\Sigma_1$-class function. Explicitly, we can write

$$y = \text{cl}(x) \iff \exists f \left( f \text{ is a function } \land \text{dom}(f) = \omega \land f(0) = x \land \forall n \in \omega \left( f(n + 1) = \bigcup f(n) \right) \land y = \bigcup \{ f(n) : n \in \omega \} \right),$$

and it is not hard to verify that this definition is $\Sigma_1$ (exercise!).

Now we can apply these results to the study of the constructible universe $L$:

\textbf{Lemma 6.41.} The class function $\text{Ord} \to U: \alpha \mapsto L_\alpha$ is $\Sigma_1$.

\textbf{Proof.} We begin by noting that various constructions that have to do with $U$-formulas are defined recursively and hence are $\Sigma_1$. These include:

- the class $\{ F \}$;
• the class function $W \mapsto \mathcal{F}_W$;
• the class function $(f, a) \mapsto f(a)$;
• &c.

Proving that these classes and class functions are $\Sigma_1$ involves an analysis similar to that performed in Example 6.40, and we omit it (but the reader is encouraged to do at least some of it as an exercise). Of particular importance is the class function $\text{Truth}$ given by Theorem 5.15:

$$\text{Truth}(W, f) = \begin{cases} 1 & \text{if } W \models f; \\ 0 & \text{if } W \models \neg f. \end{cases}$$

To see that the class function $\text{Truth}$ is $\Sigma_1$, one can either invoke Theorem 6.38, or explicitly write

$$\text{Truth}(W, f) = i \iff f \in \mathcal{F}_W^0 \land \exists T (T \text{ is a function } \land \text{ dom}(T) = \mathcal{F}_W^0 \land \forall g \in \mathcal{F}_W^0 T(g) = \begin{cases} 1 & \text{if } g = (a \equiv b) \text{ and } a = b; \\ 1 & \text{if } g = (a \in b) \text{ and } a \in b; \\ 1 & \text{if } g = (h \land h') \text{ and } T(h) = T(h') = 1; \\ 1 & \text{if } g = (\exists x h) \text{ and } T(h) = 0; \\ 1 & \text{if } g = (\exists x h) \text{ and } \exists a \in W (T(h(a)) = 1); \\ 0 & \text{otherwise.} \end{cases} \land T(f) = i).$$

Next we observe that the statement “$A = \{ a \in W : W \models f(a) \}$,” where $f \in \mathcal{F}_W^1$, is $\Sigma_1$:

$$A = \{ a \in W : W \models f(a) \} \iff f \in \mathcal{F}_W^1 \land \forall a \in A (a \in W \land \text{Truth}(W, f(a)) = 1) \land \forall a \in W (\text{Truth}(W, f(a)) = 0 \lor a \in A).$$

Therefore, the statement “$A$ is a definable subset of $W$” is also $\Sigma_1$:

$$A \text{ is definable in } W \iff \exists f \in \mathcal{F}_W^1 (A = \{ a \in W : W \models f(a) \}).$$

Hence, the class function $W \mapsto \mathcal{D}(W)$ is $\Sigma_1$:

$$\mathcal{D}(W) = Y \iff \forall A \in Y (A \text{ is definable in } W) \land \forall f \in \mathcal{F}_W^1 \exists A \in Y (A = \{ a \in W : W \models f(a) \}).$$

Notice the slight subtlety in this definition. Its second line is intended to say that every definable subset of $W$ is in $Y$, but we are not allowed to quantify over all subsets of $W$, so instead we use a quantifier ranging over $\mathcal{F}_W^1$, which is acceptable because the class function $W \mapsto \mathcal{F}_W$ is $\Sigma_1$.

Finally, the class function $\alpha \mapsto L_\alpha$ is defined recursively by repeatedly applying the definable powerset operation, so, by Theorem 6.38, it is $\Sigma_1$. Again, the explicit $\Sigma_1$-definition of this class function is as follows:

$$W = L_\alpha \iff \alpha \in \text{Ord} \land \exists f \left( f \text{ is a function } \land \text{ dom}(f) = \alpha + 1 \land f(0) = \emptyset \land \forall \beta < \alpha (f(\beta + 1) = \mathcal{D}(f(\beta))) \land \forall \beta \leq \alpha \left( \beta \text{ is a limit } \rightarrow f(\beta) = \bigcup \{ f(\gamma) : \gamma < \beta \} \right) \land f(\alpha) = W \right).$$

Now it’s time to reap the harvest:

**Proof of Theorem 6.26.** By Lemma 6.41, there is a $\Sigma_1$-formula $\varphi(x, y)$ such that $\varphi(\alpha, W)$ holds if and only if $\alpha \in \text{Ord}$ and $W = L_\alpha$. Consider the class $L$ and let $\alpha$ be an arbitrary ordinal. By Theorem 6.20, $L \models \text{ZF}$. In particular,

$L \models \text{“there is a unique set } W \text{ such that } \varphi(\alpha, W).$"
since the statement that for every ordinal $\alpha$, there is a unique set $W$ with $\varphi(\alpha, W)$, is a \textit{theorem} of ZF. Let $W_\alpha \in L$ be the unique set such that $L \models \varphi(\alpha, W_\alpha)$; in other words, $W_\alpha$ is “$L$’s version” of $L_\alpha$. Since $\varphi$ is a $\Sigma_1$-formula, by Exercise 6.34, the truth of $\varphi$ “lifts” from $L$ to $U$, and thus

\[ L \models \varphi(\alpha, W_\alpha) \implies U \models \varphi(\alpha, W_\alpha). \]

But $U \models \varphi(\alpha, W_\alpha)$ means that $W_\alpha = L_\alpha$; i.e., “$L$’s version” of $L_\alpha$ must actually coincide with the “real” $L_\alpha$. And now we are done, since $L \models $ “every set is constructible” is equivalent to

\[ L = \bigcup \{ W_\alpha : \alpha \in \text{Ord} \}, \]

which is true since $W_\alpha = L_\alpha$ for all $\alpha \in \text{Ord}$ and $L = \bigcup \{ L_\alpha : \alpha \in \text{Ord} \}$ by definition. $\blacksquare$

**Exercise 6.42.** Let $C \supseteq \text{Ord}$ be a transitive class such that $C \models \text{ZF}$. Show that $L \subseteq C$.

**Exercise 6.43 (Condensation lemma).** Let $S$ be a transitive set such that $S \models \text{ZF}$, “all sets are constructible.” Show that $S = L_\alpha$ for some ordinal $\alpha$.

### 6.G. $L$ satisfies AC and the Principle of Global Choice

The goal of this section is to prove that $L \models \text{AC}$, thus showing that if ZF is consistent, then ZFC is consistent as well. In fact, we will show that $L$ satisfies a stronger version of AC, called the \textbf{Principle of Global Choice} (GC):

\[
\text{GC: There is a class function } gc : U \rightarrow U \text{ such that for every nonempty set } A, gc(A) \in A.
\]

The above statement of GC involves quantification over class functions; nevertheless, it turns out to be equivalent to a certain single axiom—see Problem 4 on Homework 6.

**Exercise 6.44.** Show that the following statements are equivalent:

(i) GC holds;

(ii) there is a bijective class function $\text{Ord} \rightarrow U$.

(iii) for every two proper classes $C, D$, there is a bijective class function $C \rightarrow D$.

(iv) for every proper class $C$, there is a bijective class function $\text{Ord} \rightarrow C$.

(v) AC holds and for every proper class $C$, there is an injective class function $\text{Ord} \rightarrow C$.

**Hints:** In (i) $\implies$ (ii), you will have to use $\text{AF}$ to ensure surjectivity. For (v) $\implies$ (ii), consider

\[ W := \{ < : \exists \alpha \in \text{Ord} (< \text{ is a well-ordering on } V_\alpha) \}. \]

By AC, $W$ is a proper class. Use an injection $\text{Ord} \rightarrow W$ to construct a well-ordering of $V = U$.

**Theorem 6.45.** \textit{Suppose that every set is constructible. Then GC holds.}

From Theorems 6.26 and 6.45, it follows that $L$ always satisfies GC (and hence AC), even if $U$ does not. This is why we proved Theorem 6.26 first: now instead of worrying about the differences between $L$ and $U$, we may just assume that every set is constructible, i.e., $U = L$.

The proof of Theorem 6.45 hinges on the following observation:

**Lemma 6.46.** \textit{There exists a class function that, given a set $W$ and a well-ordering $<$ on $W$, outputs a well-ordering $<^*$ on $D(W)$.}

**Proof.** Every definable subset of $W$ is given by a $U$-formula with one free variable and with parameters from $W$, so we just have to well-order the set $F_W^1$. To that end, we first fix an arbitrary well-ordering $\prec$ of $F$, which exists since $F$ is countable (see Exercise 5.1). Every $U$-formula in
\( \mathcal{F}_W \) can be written as \( f(x, \bar{a}) \), where \( x \) is the free variable and \( \bar{a} = (a_1, \ldots, a_k) \) is a finite tuple of parameters from \( W \). Hence, we may identify the elements of \( \mathcal{F}_W \) with pairs of the form \( (f, \bar{a}) \), where 
\[
  f = f(x, x_1, \ldots, x_k) \in \mathcal{F} \quad \text{and} \quad \bar{a} = (a_1, \ldots, a_k) \in \omega W.
\]
The set \( \omega W \) can be well-ordered lexicographically:
\[
  (a_1, \ldots, a_k) \prec_{\text{lex}} (b_1, \ldots, b_\ell) \iff (k < \ell) \text{ or } (k = \ell \text{ and } a_i < b_i, \text{ where } i \text{ is the least index such that } a_i \neq b_i).
\]
This allows us to put a well-ordering \( \prec_{(W, \prec)} \) on \( \mathcal{F}_W \) as follows:
\[
  (f, \bar{a}) \prec_{(W, \prec)} (g, \bar{b}) \iff (f \prec g) \text{ or } (f = g \text{ and } \bar{a} \prec_{\text{lex}} \bar{b}).
\]
Finally, for each \( A \in \mathcal{D}(W) \), let \( f_A \) be the \( \prec_{(W, \prec)} \)-least \( \mathcal{U} \)-formula in \( \mathcal{F}_W \) such that
\[
  A = \{ a \in W : W \models f_A(a) \}.
\]
Then we obtain a desired well-ordering on \( \mathcal{D}(W) \) by setting
\[
  A \prec^g B \iff f_A \prec_{(W, \prec)} f_B.
\]

**Proof of Theorem 6.45.** Suppose that every set is constructible, i.e., \( \mathcal{U} = L \). We recursively define well-orderings \( \prec_\alpha \) on \( L_\alpha, \alpha \in \text{Ord} \), as follows. To begin with, let \( \prec_0 \) be the empty ordering on \( \emptyset = L_0 \). Next, let \( \beta \in \text{Ord} \) and suppose that \( \prec_\beta \) is already defined. Let \( \prec_\beta^g \) be the well-ordering of \( \mathcal{D}(L_\beta) = L_{\beta+1} \) given by Lemma 6.46. Then for each \( x, y \in L_{\beta+1} \), we let
\[
  x \prec_{\beta+1} y \iff (x, y \in L_\beta \text{ and } x \prec_\beta y) \text{ or } (x \in L_\beta \text{ and } y \in L_{\beta+1} \setminus L_\beta) \text{ or } (x, y \in L_{\beta+1} \setminus L_\beta \text{ and } x \prec^g_\beta y).
\]
This definition ensures that the ordering \( \prec_{\beta+1} \) extends \( \prec_\beta \) and that \( L_\beta \) is \( \prec_{\beta+1} \)-downward closed. This allows us to define, for limit ordinals \( \alpha \),
\[
  \prec_\alpha := \bigcup \{ \prec_\gamma : \gamma < \alpha \}.
\]
Now let \( \prec := \bigcup \{ \prec_\alpha : \alpha \in \text{Ord} \} \). Then \( \prec \) is a well-ordering of \( L = \mathcal{U} \), and thus GC holds.

**6.H. L satisfies GCH**

Finally, we will prove in this subsection that \( L \) satisfies the Generalized Continuum Hypothesis:

**Theorem 6.47.** If every set is constructible (i.e., if \( \mathcal{U} = L \)), then GCH holds.

Again, by Theorem 6.26, \( L \models \) “every set is constructible,” so Theorem 6.47 implies that \( L \models \) GCH. This, in turn, means that if \( \text{ZF} \) is consistent, then so is \( \text{ZF} + \text{GCH} \).

Recall that GCH implies AC, and hence Theorem 6.47 implies that \( L \models \) AC; however, we will actually use that \( L \) satisfies AC in the proof of Theorem 6.47. The main ingredient in the proof of Theorem 6.47 is a remarkable result concerning the cardinality of sets that can be identified using a \( \Sigma_1 \)-formula. Let \( a \) be a set. We say that a set \( b \) is \( \Sigma_1 \)-identifiable over \( a \) if there is a \( \Sigma_1 \)-formula \( \varphi(x, y) \) with two free variables \( x, y \) and without parameters such that
\[
  y = b \iff \varphi(a, y),
\]
i.e., if \( \{ y : \varphi(a, y) \} = \{ b \} \). How large, in terms of its cardinality, can a set that is \( \Sigma_1 \)-identifiable over a given set \( a \) be? On the one hand, \( \omega \) is \( \Sigma_1 \)-identifiable (even without using \( a \)), and \( |\omega| = \aleph_0 \). On the other hand, \( a \) itself is \( \Sigma_1 \)-identifiable over \( a \) (via the formula \( y = a \)); furthermore, by Example 6.40, the set \( \text{cl}(a) \) (i.e., the transitive closure of \( a \)) is also \( \Sigma_1 \)-identifiable over \( a \). It turns out that no set of strictly greater cardinality could be \( \Sigma_1 \)-identifiable over \( a \):
Theorem 6.48 (Gödel’s Magic Theorem). Assume that $\mathcal{U} \models \text{ZFC}$. Let $a$ and $b$ be sets and suppose that $b$ is $\Sigma_1$-identifiable over $a$. Then

$$|b| \leq \max\{\aleph_0, |\text{cl}(a)|\}.$$  

We will prove Theorem 6.48 in §6.1. Theorem 6.48 is a perfect example of the power of mathematical logic. Using this theorem one can, simply by looking at a definition of a set $b$ and making sure that it involves no unbounded universal quantifiers, obtain a sharp upper bound on $|b|$, without trying to understand what the definition means at all!

Proof of Theorem 6.47 (from Theorem 6.48). Assume $\mathcal{U} = L$. By Theorem 6.45, this implies that AC holds, so we can talk about cardinalities and use Theorem 6.48.

We start by making two observations:

Observation 6.49. For every infinite ordinal $\alpha$, we have $|L_\alpha| = |\alpha|$.

Proof. Since $\alpha \subseteq L_\alpha$, $|\alpha| \leq |L_\alpha|$. On the other hand, by Lemma 6.41, the set $L_\alpha$ is $\Sigma_1$-identifiable over $\alpha$, and thus, by Theorem 6.48, we have

$$|L_\alpha| \leq \max\{\aleph_0, |\text{cl}(\alpha)|\} = |\alpha|,$$

where we are using that $\text{cl}(\alpha) = \alpha \geq \omega$.

Recall that for a constructible set $a$, order$(a)$ is the least ordinal $\alpha$ such that $a \in L_\alpha$.

Observation 6.50. For every $a \in L$, order$(a)$ is $\Sigma_1$-identifiable over $a$. And indeed, using Lemma 6.41, we obtain:

$$\alpha = \text{order}(a) \iff \begin{aligned} \alpha &\in \text{Ord} \land a \in L_\alpha \land \forall \beta < \alpha (a \notin L_\beta) \\
&\iff \alpha \in \text{Ord} \land \exists W \left(W = L_\alpha \land a \in W \right) \land \forall \beta < \alpha \exists U \left(U = L_\beta \land a \notin U \right). \end{aligned}$$

Let $\kappa$ be an infinite cardinal. We have to show that $2^\kappa = |\mathcal{P}(\kappa)| \leq \kappa^+$. To that end, consider an arbitrary subset $A \subseteq \kappa$. Since $\mathcal{U} = L$, $A$ is constructible, and, by Observation 6.50,

$$|\text{order}(A)| \leq \max\{\aleph_0, |\text{cl}(A)|\} \leq \kappa,$$

where we are using that $\kappa$ is a transitive set and hence $\text{cl}(A) \subseteq \kappa$. Therefore, order$(A) < \kappa^+$, so $A \in L_{\kappa^+}$.

Since this holds for every subset $A \subseteq \kappa$, we conclude that $|\mathcal{P}(\kappa)| \leq |L_{\kappa^+}| = \kappa^+$. \qed

Exercise 6.51. Suppose that $\mathcal{U} = L$. Show that for every ordinal $\alpha$, $V_{\omega+\alpha} \subseteq L_{\aleph_\alpha}$.

Exercise 6.52. Suppose that $\mathcal{U} = L$ and let $\kappa$ be a cardinal such that $\kappa = \aleph_\kappa$. Show that $V_\kappa = L_\kappa$.

6.1. Proof of Gödel’s Magic Theorem

The proof of Theorem 6.48 relies on a combination of two important tools, the first of which is the so-called Löwenheim–Skolem theorem:

Theorem 6.53 (Löwenheim–Skolem). Assume that $\mathcal{U} \models \text{ZFC}$. Let $A \subseteq B$ be sets. Then there exists a set $A^*$ such that $A \subseteq A^* \subseteq B$, $|A^*| \leq \max\{\aleph_0, |A|\}$, and for all $f \in \mathcal{F}_{A^*}$, we have

$$A^* \models f \iff B \models f.$$  

Theorem 6.53 might seem similar to the Reflection Principle, and for good reason: it’s proof relies on a similar process of adding witnesses to existential statements (called Skolemization).
Proof. Let $S \subseteq B$ and suppose that there is some $\mathcal{U}$-formula $f \in \mathcal{F}_S^0$ such that $S \models f \iff B \models f$. Consider such $f$ of the lowest possible complexity. The same analysis as in the proof of the Reflection Principle (see §6.E) shows that $f$ must be of the form $f = \exists x \ g(x)$ for some $g \in \mathcal{F}_S^1$, which yields the following result, known as the Tarski–Vaught test:

**Exercise 6.54 (Tarski–Vaught test).** Let $S \subseteq B$ and suppose that for all $g \in \mathcal{F}_S^1$,

$$B \models \exists x \ g(x) \implies \exists a \in S \text{ such that } B \models g(a).$$

Show that for all $f \in \mathcal{F}_S^0$, we have $S \models f \iff B \models f$.

Fix a choice function $\text{ch}: \mathcal{P}(B) \setminus \emptyset \to B$. For a $\mathcal{U}$-formula $g \in \mathcal{F}_B^1$ such that $B \models \exists x \ g(x)$, define

$$W(g) := \text{ch} \{ a \in B : B \models g(a) \},$$

and for a subset $S \subseteq B$, let

$$W(S) := \{ W(g) : g \in \mathcal{F}_S^1 \text{ such that } B \models \exists x \ g(x) \}.$$

Note that $S \subseteq W(S)$, since for each $a \in S$, we have $a = W(x = a)$.

Now we recursively define a sequence of sets $A_n, n \in \omega$, by

$$A_0 := A, \quad \text{and} \quad A_{n+1} := W(A_n) \text{ for all } n \in \omega.$$

Finally, set

$$A^* := \bigcup \{ A_n : n \in \omega \}.$$ 

We claim that the set $A^*$ has all the required properties. By definition, $A \subseteq A^* \subseteq B$. If $g \in \mathcal{F}_{A^*}^1$ is such that $B \models \exists x g(x)$, then, since $g$ involves only finitely many parameters from $A^*$, there is some $n \in \omega$ such that all the parameters in $g$ come from $A_n$, i.e., $g \in \mathcal{F}_{A_n}^0$. But then $W(g) \in A_{n+1} \subseteq A^*$, and, by definition, $B \models g(W(a))$, so $A^*$ passes the Tarski–Vaught test. Hence, for all $f \in \mathcal{F}_{A^*}^0$, we have $A^* \models f \iff B \models f$.

It remains to bound $|A^*|$. To that end, recall that for every set $S$, we have $|\mathcal{F}_S| = \max\{\aleph_0, |S|\}$ (see Exercise 5.4). Hence, we can inductively show that for all $n \in \omega$,

$$|A_{n+1}| \leq |\mathcal{F}_{A_n}^1| = \max\{\aleph_0, |A_n|\} \leq \max\{\aleph_0, |A|\}.$$ 

Thus, $A^*$ is a union of countably many sets, each of cardinality at most $\max\{\aleph_0, |A|\}$, and

$$|A^*| \leq \aleph_0 \times \max\{\aleph_0, |A|\} = \max\{\aleph_0, |A|\}. \quad \blacksquare$$

Here’s an interesting application of the Löwenheim–Skolem theorem. Suppose that $\mathcal{U} \models \text{ZFC}$ and let $\kappa$ be an inaccessible cardinal. By Theorem 4.2, we have $V_\kappa \models \text{ZFC}$. Applying the Löwenheim–Skolem theorem with $A = \emptyset$ and $B = V_\kappa$ gives a countable set $A^* \subset V_\kappa$ such that for all $f \in \mathcal{F}_{A^*}^0$, we have $A^* \models f \iff V_\kappa \models f$. In particular,

$$A^* \models \text{ZFC}.$$

In other words, under the assumption that there is an inaccessible cardinal, we can find a countable set that satisfies all the axioms of ZFC. This observation is known as Skolem’s paradox.\(^7\) Skolem himself found this paradox so counter-intuitive that he believed it fully discredits the logical foundations of set theory. For instance, the existence of an uncountable set is a theorem of ZFC, and thus

$$A^* \models \text{“there is an uncountable set,”}$$

even though $A^*$ itself is countable. However, we know that there is nothing paradoxical about this: the statement “there is an uncountable set” really means that

$$\exists S \neg \exists f \ (f \text{ is an injection } S \to \omega),$$

\(^7\)Named after the Norwegian mathematician Thoralf Skolem.
so \( A^* \models \) “there is an uncountable set” simply means that there is some \( S \in A^* \) such that there is no \( f \) in \( A^* \) with \( A^* \models \) “\( f \) is an injection \( S \to \omega \),” which has nothing to do with whether \( A^* \) itself is a countable set from the point of view of \( U \).

Let’s investigate the structure of this set \( A^* \) a bit more. Since \( A^* \models \text{ZFC} \), we have

\[ A^* \models \exists S (S = \omega), \]

where “\( S = \omega \)” is a shorthand for any formula that defines \( \omega \). Let \( S \in A^* \) be the set such that \( A^* \models S = \omega \). The statement “\( S = \omega \)” can be expressed by a \( U \)-formula using \( S \) as a parameter, and thus we must have \( V_\kappa \models S = \omega \). But this means that \( S \) actually is \( \omega \); in other words, \( \omega \in A^* \).

Similarly, \( A^* \) satisfies the Powerset Axiom, and hence

\[ A^* \models \exists P (P = \mathcal{P}(\omega)). \]

Again, letting \( P \in A^* \) be the set such that \( A^* \models P = \mathcal{P}(\omega) \), we conclude that \( V_\kappa \models P = \mathcal{P}(\omega) \), which implies that \( P \) is the powerset of \( \omega \) (exercise!). But this means that \( \mathcal{P}(\omega) \in A^* \). So, even though \( A^* \) itself is a countable set, it contains as an element the uncountable set \( \mathcal{P}(\omega) \). Continuing in like manner, we see that the following sets are elements of \( A^* \):

\[ V_\omega, \ V_{\omega+1}, \ V_{\omega+\omega}, \ V_{\omega\omega}, \ V_{\aleph_1}, \ V_{\aleph_\omega}, \ & \text{etc}. \]

One consequence of the above discussion is that \( A^* \) is definitely not a transitive set, as it has lots of uncountable elements. Nevertheless, it turns out that we can replace \( A^* \) by an isomorphic transitive set using a trick known as the Mostowski collapse, which is the second tool we need to establish the Magic Theorem:

**Theorem/Definition 6.55 (Mostowski collapse).** Let \( C \) be a class such that \( C \models \text{Ext} \). Then there is a unique class function \( j : C \to U \) such that:

(W1) \( j \) is injective;
(W2) \( \text{ran}(j) \) is a transitive class; and
(W3) for all \( x, y \in C \), we have \( y \in x \iff j(y) \in j(x) \).

The class function \( j : C \to \text{ran}(j) \) is called the **Mostowski collapse** of \( C \).

**Proof.** By recursion on the rank of \( x \in C \), define

\[ j(x) := \{j(y) : y \in x \cap C\}. \tag{6.56} \]

The fact that \( C \models \text{Ext} \) guarantees that \( j \) is injective. The details are left as an exercise. \( \square \)

**Example 6.57.** To see why it’s necessary to assume in Theorem 6.55 that \( C \models \text{Ext} \), consider the set

\[ A := \{\varnothing, \varnothing, \varnothing, \varnothing\} \]

from Example 2.1. Recall that \( \text{Ext} \) fails in \( A \), because the sets \( \{\varnothing\} \) and \( \{\varnothing, \{\varnothing\}\} \) have the same (unique) element in \( A \), namely \( \varnothing \). Thus, if we were to define a function \( j : A \to U \) using (6.56), it won’t be injective:

\[ j(\varnothing) = \varnothing; \quad j(\{\varnothing\}) = j(\varnothing, \{\varnothing\}) = \{\varnothing\}. \]

**Example 6.58.** Consider the set \( E := \{n \in \omega : n \text{ is even}\} \). Then \( E \models \text{Ext} \) (exercise!). We claim that the Mostowski collapse of \( E \) is the function \( j : E \to \omega : n \mapsto n/2 \). Indeed, \( j \) is certainly injective and its range, \( \omega \), is a transitive set. Furthermore, if \( n, m \in E \), then

\[ n \in m \iff n < m \iff n/2 < m/2 \iff n/2 \in m/2. \]

On the other hand, we could obtain the same result by induction using formula (6.56):

\[ j(0) = j(\varnothing) = \varnothing = 0; \]
\[ j(2) = j(\{0,1\}) = j(\{0\}) = \{0\} = 1; \]
\[ j(4) = j(\{0,1,2,3\}) = j(\{0,1\}) = \{0,1\} = 2; \]
\[ j(6) = j(\{0,1,2,3,4,5\}) = j(\{0,1,2,3\}) = \{0,1,2\} = 3; \quad \& \text{c.} \]
Exercise 6.59. Let $C \subseteq \text{Ord}$ be a proper class. Show that the unique order-isomorphism $C \to \text{Ord}$ coincides with the Mostowski collapse of $C$.

Exercise 6.60. Let $C$ be a transitive class. Show that the Mostowski collapse of $C$ is the identity class function $\text{id}_C : C \to C$. More generally, let $C$ be a class such that $C \models \text{Ext}$ and let $A \subseteq C$ be a transitive subclass. Let $j$ be the Mostowski collapse of $C$. Show that $j\restriction A = \text{id}_A$.

Let $C$ be a class such that $C \models \text{Ext}$ and let $j : C \to C'$ be the Mostowski collapse of $C$, where $\text{ran}(j) = C'$. By definition, $j$ is an isomorphism between the structures

$$(C, \in) \quad \text{and} \quad (C', \in).$$

Hence, any statement that holds in $C$ must also hold in $C'$, and vice versa. More precisely, let $\varphi(a_1, \ldots, a_k)$ be a formula without free variables and with parameters $a_1, \ldots, a_k \in C$. Then

$$C \models \varphi(a_1, \ldots, a_k) \iff C' \models \varphi(j(a_1), \ldots, j(a_k)),$$

where $j(a_1), \ldots, j(a_k)$ are treated as parameters from $C'$. Similarly, if $A$ is a set such that $A \models \text{Ext}$ and $j : A \to A'$ is the Mostowski collapse of $A$ with $\text{ran}(j) = A'$, then the above observation extends to $\mathcal{U}$-formulas; that is, for every $\mathcal{U}$-formula $f(a_1, \ldots, a_k) \in \mathcal{F}_A$ with parameters $a_1, \ldots, a_k \in A$,

$$A \models f(a_1, \ldots, a_k) \iff A' \models f(j(a_1), \ldots, j(a_k)),$$

where $j(a_1), \ldots, j(a_k)$ are treated as parameters from $A'$ (and so $f(j(a_1), \ldots, j(a_k)) \in \mathcal{F}_{A'}$).

Corollary 6.61 (Löwenheim and Skolem meet Mostowski). Assume that $\mathcal{U} \models \text{ZFC}$. Let $A \subseteq B$ be sets. Suppose that $A$ is transitive and $B \models \text{Ext}$. Then there exists a transitive set $A' \supseteq A$ such that $|A'| \leq \max\{|\mathbb{N}_0|, |A|\}$ and for all $f \in \mathcal{F}_A$, we have

$$A' \models f \iff B \models f.$$

Note that Corollary 6.61 does not claim that $A' \subseteq B$.

Proof. First we apply the Löwenheim–Skolem theorem to obtain a set $A^*$ such that $A \subseteq A^* \subseteq B$, $|A^*| \leq \max\{|\mathbb{N}_0|, |A|\}$, and for all $f \in \mathcal{F}_{A^*}$, we have $A^* \models f \iff B \models f$. Since $B \models \text{Ext}$, we conclude that $A^* \models \text{Ext}$ as well, and thus we can consider the Mostowski collapse $j : A^* \to A'$, where $A' := \text{ran}(j)$. By construction, $A'$ is a transitive set with $|A'| = |A^*| \leq \max\{|\mathbb{N}_0|, |A|\}$. By Exercise 6.60, since $A$ is a transitive subset of $A^*$, we have $j\restriction A = \text{id}_A$ and hence $A = j\[A] \subseteq A'$. Finally, if $f \in \mathcal{F}_{A^*}$, then $A' \models f \iff A^* \models f \iff B \models f$, as desired. (Here we are using again that $j$ acts as the identity function on the elements of $A$.)

Example 6.62. It follows that if $\mathcal{U} \models \text{ZFC}$ and there is an inaccessible cardinal $\kappa$, then there exists a transitive countable set that satisfies $\text{ZFC}$: simply apply Corollary 6.61 with $A = \emptyset$ and $B = V_\kappa$. This is a nice example of a combinatorial statement about the existence of a countable structure with certain properties that relies on the existence of a large uncountable cardinal $\kappa$. (Also, see Exercise 6.65.)

Exercise 6.63. Suppose that $\mathcal{U} \models \text{ZFC}$ and let $\varphi_1, \ldots, \varphi_n$ be a finite list of axioms of $\text{ZFC}$. Show that there is a transitive countable set $S$ such that $S \models \varphi_1 \land \ldots \land \varphi_n$.

We are now ready to prove Theorem 6.48.

Proof of Gödel’s Magic Theorem. Recall the set-up: we are assuming that $\mathcal{U} \models \text{ZFC}$, $a$ and $b$ are sets, and $b$ is $\Sigma_1$-identifiable over $a$. Fix a $\Sigma_1$-formula $\varphi(x, y)$ such that

$$y = b \iff \varphi(a, y). \quad (6.64)$$

Consider the formula $\psi(x) := \exists y \varphi(x, y)$ with one free variable $x$. By the Reflection Principle, there is an ordinal $\beta$ such that $a \in V_\beta$ and $\psi$ is absolute between $V_\beta$ and $V = \mathcal{U}$. In particular, since $\mathcal{U} \models \psi(a)$ (as witnessed by $b$), we also have $V_\beta \models \psi(a)$. Let $A := \text{cl}(\{a\}) = \text{cl}(a) \cup \{a\}$ (i.e., $A$ is
the smallest transitive set that has $a$ as an element). Since $V_\beta$ is transitive, we have $A \subseteq V_\beta$ and $V_\beta \models \text{Ext}$, so we may apply Corollary 6.61 to $A$ and $V_\beta$, obtaining a transitive set $A'$ such that:

- $A \subseteq A'$;
- $|A'| \leq \max\{|\mathfrak{N}_0, |A|\} = \max\{|\mathfrak{N}_0, |\text{cl}(a)|\}$; and
- for all $f \in F_A^0$, $A' \models f \iff V_\beta \models f$.

In particular, since $a \in A$ and $V_\beta \models \psi(a)$, we also have $A' \models \psi(a)$, i.e.,

$$A' \models \exists y \varphi(a, y).$$

Let $b' \in A'$ be a set such that $A' \models \varphi(a, b')$. Since $A'$ is transitive and $\varphi$ is a $\Sigma_1$-formula, the truth of $\varphi(a, b')$ “lifts” from $A'$ to $\mathcal{U}$ (see Exercise 6.34), and hence $\mathcal{U} \models \varphi(a, b')$. But, by (6.64), this means that $b' = b$. It remains to notice that, by the transitivity of $A'$, $b = b' \subseteq A'$, and hence

$$|b| \leq |A'| \leq \max\{|\mathfrak{N}_0, |A|\}.$$  

**Exercise 6.65.** Suppose that there is a transitive set $S$ such that $S \models \text{ZF}$. (This happens, e.g., if there exists an inaccessible cardinal.) Show that there is a countable ordinal $\alpha$ such that $L_\alpha \models \text{ZFC}$. Hint: Use Exercise 6.43.