SET THEORY AND FORCING LECTURE NOTES

JEAN-LOUIS KRIVINE

TRANSLATED BY: CHRISTIAN ROSENDAL

TYPESET BY: JESSICA SCHIRLE

Contents

Page

0. Typesetter's Introduction

These notes provide a great introduction to axiomatic set theory and topics therein appropriate for a first class for a graduate or upper level undergraduate student. I was taught set theory by Professor Anush Tserunyan at the university of Illinois at Urbana-Champaign in the Spring of 2018 following mostly these notes. My goal in TeXing these notes is to stick to Rosendal's original handwritten translation of Krivine's notes as closely as possible so as to represent them well and avoid making mistakes myself. However, I have taken the liberty in a couple sections to change minor details. These include things like grammar or notation (for example, when discussing forcing, Professor Tserunyan took the liberty of underlining variable names as opposed to barring the top, so as to be able to write vectors as well when dealing with multiple variables, and I also employ this). The overall message of these notes is unchanged. Where I include notes of my own, I precede them with *TN:*, for Typesetter's Note. I am also greatly indebted to both Professor Tserunyan and Ms. Jenna Zomback. Without Professor Tserunyan, I would not have any good grasp of the topics herein, and Ms. Zomback provided the last set of notes that were mysteriously missing from both Rosendal's handwritten pdf and my own notebook. If you happen to notice any typos, issues, compliments, or complaints with these notes, please feel free to email me at schirlem (at) uci (dot) edu.

1. Axioms of Set Theory

Definition 1.1. A universe of set theory is a non-empty collection of objects $\mathcal{U}$ equipped with a binary relation $\in$ called the membership relation. For x, y belonging to $\mathcal{U}$ (which we will not denote by $x, y \in \mathcal{U}$ to avoid confusion with the above membership relation), when

$$x \in y$$

we say that x is a member of y or x belongs to y.

The objects in $\mathcal{U}$ are called sets. Moreover, $\mathcal{U}$ and $\in$ are supposed to satisfy the following list of axioms that we detail individually:[a]

(1) Axiom of Extensionality

$$\forall x \forall y \left(\forall z (z \in x \longleftrightarrow z \in y) \longrightarrow x = y\right)$$

That is, two sets with the same members are identical.

(2) Pairing Axiom

$$\forall x \forall y \exists z \forall u \big(u \in z \longleftrightarrow (u = x \vee u = y)\big)$$

That is, for any two sets x and y, there is a set z whose members are exactly x and y. Since, by the axiom of extensionality, this set z is unique, we will denote it by the notation

$$\{x, y\}.$$

Note also that this applies to the case when $x = y$, and we get a set $\{x, x\}$, whose unique member is x. We will simplify this notation to $\{x\}$ and call it the singleton of x.
When $x \neq y$, we say that $\{x, y\}$ is a pair or a doubleton.

Definition 1.2. Given sets x and y, we can define the ordered pair or couple $\{\{x\}, \{x, y\}\}$ by repeated application of the pairing axiom. We denote this by (x, y).

Lemma 1.3. If x, y, a, b are sets and $(x, y) = (a, b)$, then $x = a$ and $y = b$.

Proof.

If $x = y$, then $(x, y) = \{\{x\}, \{x, y\}\} = \{\{x\}, \{x\}\} = \{\{x\}\}$ has only a single element, implying that also $(a, b) = \{\{a\}, \{a, b\}\}$ has a single element which must be $\{a\} = \{x\}$. Thus, $a = b = x = y$.

If $x \neq y$, then (x, y) has two elements, namely, a singleton $\{x\}$ and a doubleton $\{x, y\}$. It follows that also (a, b) must contain a unique singleton, namely $\{a\}$, and a unique doubleton, namely $\{a, b\}$. From the uniqueness and the axiom of extensionality, $a = x$ and then also $b = y$. □

[a] *TN:* For those more familiar with model theory, $(\mathcal{U}, \in)$ will be a model of ZFC, and objects within this model will be sets. It will take us a while to fully describe the theory ZFC, as we want the reader to have an appreciation for each of the axioms therein.

Definition 1.4. For any sets $x_1, x_2, ..., x_n$, we can inductively define n-tuples $(x_1, x_2, ..., x_n)$ by

$$(x_1, x_2, x_3) := (x_1, (x_2, x_3))$$

$$(x_1, x_2, x_3, x_4) := (x_1, (x_2, x_3, x_4))$$

etc.

Theorem 1.5. For any $n > 1$, we have

$$(x_1, x_2, ..., x_n) = (y_1, y_2, ..., y_n)$$

$$\implies (x_1 = y_1) \wedge (x_2 = y_2) \wedge ... \wedge (x_n = y_n)$$

Proof.

Omitted in original notes, but very straightforward.

□

(3) Union Axiom

$$\forall x \exists y \forall z \Big(z \in y \longleftrightarrow \exists u (z \in u \wedge u \in x) \Big)$$

That is, for every set x, there is a set y whose members are exactly the z that belong to a member of x.[a] Again, by extensionality, this set y is unique and is denoted

$$\bigcup x \qquad \text{or} \qquad \bigcup_{u \in x} x$$

We call it the union over x. For example, $\bigcup (x, y) = \bigcup \{\{x\}, \{x, y\}\} = \{x, y\}$.
So, by induction on n, we can for any sets $x_1, x_2, ..., x_n$ construct a set $\{x_1, x_2, ..., x_n\}$ whose members are exactly $\{x_1, x_2, ..., x_n\}$. For this,

$$\{x_1, x_2, ..., x_n\} := \bigcup \big\{ \{x_1\}, \{x_2, ..., x_n\} \big\}.$$[b]

Also, if x and y are sets, we denote by

$$x \cup y := \bigcup \{x, y\}$$

and call it the union of x and y.

Fact:
The operation $\cup$ is associative, i.e.,

$$x \cup (y \cup z) = (x \cup y) \cup z$$

so we denote $(x_1 \cup (x_2 \cup ...(x_{n-1} \cup x_n)...))$ unambiguously by $x_1 \cup x_2 \cup ... \cup x_n$.

[a] *TN:* Prof. Tserunyan explains this in a very intuitive way. Think of x as your car, and its members z are the grocery bags, and the sets u are the actual groceries. The set y is the fridge where you empty all of your groceries into. So, for any car full of groceries, there is a fridge where you've emptied all the grocery bags into.

[b] *TN:* Note that up to this point, there is no way to construct such a set. Repeated application of the pairing axiom will not allow for pairing a doubleton and singleton to achieve a set of three elements. One needs to do this and then apply the union axiom.

(4) Power Set Axiom[a]
First, to simplify notation, we write $x \subseteq y$ whenever the following holds:

$$\forall z(z \in x \longrightarrow z \in y).$$

When $x \subseteq y$, we say that x is a subset of y or is contained/included in y.

The Power Set Axiom is then the following statement:

$$\forall x \exists y \forall z(z \subseteq x \longleftrightarrow z \in y).$$

That is, for every x there is a (unique) set y whose elements are exactly the subsets of x. We call y the power set of x and denote it by $\mathcal{P}(x)$.

CAUTION: One must be careful when understanding the power set axiom. For the variable z only refers to objects in $\mathcal{U}$ and *not* subsets of x that happen not to be in $\mathcal{U}$. In fact, it is a basic idea in the construction of universes to make judicious choices of which subsets of a set to include in $\mathcal{U}$ and which to leave out. So, in such a $\mathcal{U}$, $\mathcal{P}(x)$ will only consist of the subsets of x that are actually in $\mathcal{U}$.

1.1. Class Relations:

TN: We take a short divergence to introduce class relations, which will become very important in stating the axiom schemata of replacement and comprehension.

The language of set theory is the usual first-order language including the logical symbol $=$ and the extra-logical symbol $\in$. Now, if $\phi(x, y, z)$ is a formula with at most three free variables x, y, z, and possibly having parameters $a_1, a_2, ..., a_n$, we have a corresponding relation R_ϕ on $\mathcal{U}$ defined by

$$R_\phi(a, b, c) \Longleftrightarrow \phi \text{ holds of } a, b, c \text{ in } \mathcal{U}.$$

For example, if ϕ is $\forall z(z \in x \longrightarrow z \in y)$, then R_ϕ is the inclusion relation $\subseteq$.

The relations so obtained are called class relations, and unary class relations are also just called classes. Later on, we shall misuse notation and write $(b_1, ..., b_n) \in R_\phi$ to denote that $\phi(b_1, ..., b_n)$ holds even though R_ϕ is in general not a set.

Example:

$$\phi(x) := \forall u\,(u \in x \longrightarrow \exists v(v \in x \wedge \forall t\,(t \in v \longleftrightarrow t = u \vee t \in u)))$$

defines the class of all sets x such that if $u \in x$, then also $u \cup \{u\} \in x$.

Class Functions:
Suppose $\phi(x_1, x_2, ..., x_n, y)$ is a formula. We say that ϕ defines a class function R_ϕ if the following holds in $\mathcal{U}$:

$$\forall x_1 \forall x_2 ... \forall x_n \forall y \forall y'(\phi(x_1, ..., x_n, y) \wedge \phi(x_1, ..., x_n, y') \longrightarrow y = y').$$

In this case, the formula $\exists y \phi(x_1, ..., x_n, y)$ in variables $x_1, ...x_n$ defines the domain of R_ϕ while $\exists x_1 \exists x_2 ... \exists x_n \phi(x_1, ..., x_n, y)$ defines the image of R_ϕ. To simplify notation, we shall often write

[a] *TN:* It will be noted that both Prof. Tserunyan and I believe that, should ZFC be inconsistent, the inconsistency will be at the hands of this particular axiom (well, or maybe the set existence axiom, to come later). As we'll see later, the Axiom of Choice won't introduce any inconsistencies when added to ZF, so there's no need to fear choice.

$R_\phi(x_1, ..., x_n) = y$ to denote $\phi(x_1, ..., x_n, y)$.
TN: Further, when the length is obvious, we shall also often denote $\phi(x_1, ..., x_n)$, $\exists x_1 ... \exists x_n$, and $\forall x_1 ... \forall x_n$ as $\phi(\vec{x})$, $\exists \vec{x}$ and $\forall \vec{x}$, respectively.

(5) Axiom Schema of Replacement or Substitution
Suppose $\phi(x, y, \vec{a})$ is a formula with at most free variables x and y and having parameters $\vec{a}$ from $\mathcal{U}$. Suppose ϕ defines a class function on $\mathcal{U}$, i.e.
$$\forall x \forall y \forall y' (\phi(x, y, \vec{a}) \wedge \phi(x, y', \vec{a} \longrightarrow y = y')$$
then the following is an axiom:
$$\forall z \exists u \forall y \left(y \in u \longleftrightarrow \exists x \in z,\ \phi(x, y, \vec{a})\right).$$
That is, for every set z, we can take the image of z under the class function R_ϕ. Note that this does not say the image under R_ϕ, i.e. that the image of all of $\mathcal{U}$ under R_ϕ is a set, only that the image of a set z is a set u. As we shall see, the unrestricted axiom would lead to contradictions.

Russell's Paradox:
Note that $\phi(x, y)$ given by "$(x = y) \wedge (y \notin y)$" is a functional relation. So if we allowed the unrestricted replacement axiom, then there would be a set u such that $y \in u \longleftrightarrow y \notin y$. But then $u \in u \longrightarrow u \notin u$ and $u \notin u \longrightarrow u \in u$, i.e. $u \in u \longleftrightarrow u \notin u$, which is a contradiction. So, we must restrict the replacement axiom to a specific set z.[a]

(6) The Comprehension Scheme
Suppose now $\psi(x, \vec{a})$ is a formula with a single free variable x and with parameters $\vec{a}$. From ψ we construct a new formula $\phi(x, y, \vec{a})$ by
$$\phi(x, y, \vec{a}) := \psi(x, \vec{a}) \wedge (x = y).$$
Again, $\phi(x, y, \vec{a})$ defines a class function, so we can apply the replacement axiom to obtain
$$\forall z \exists u \forall v \big(v \in u \longleftrightarrow (v \in z \wedge \psi(v, \vec{a}))\big).$$
For every z, this u is unique and will be denoted $\{v \in z \mid \psi(v, \vec{a})\}$ (*TN:* Or, sometimes $\{v \in z : \psi(v, \vec{a})\}$).
That is, for any set z and class R_ψ (defined possibly with parameters), there is a set $\{v \in z \mid R_\psi(v)\}$ consisting of the elements of z belonging to the class R_ψ.

(7) Set Existence
Since we demand $\mathcal{U}$ to be a non-empty collection, we also include the axiom
$$\exists x (x = x).$$
From this follows that there is a unique set, denoted $\varnothing$, having no members. For let $\psi(x)$ be the formula $x \neq x$ and let z be any set. Then
$$\varnothing = \{x \in z \mid x \neq x\}$$
has no members and is the unique such set.

[a] *TN:* This particular paradox was a big deal in the 1900s, as up to that point, set theory basically only existed as naive set theory. This paradox led to the formalization and axiomatization of ZFC that we are currently exploring.

TN: At this point, we have covered a lot of the axioms of ZFC. There still remain the axiom of infinity, the axiom of foundation, and the axiom of choice. These will be introduced later as we build up more machinery to be able to discuss them in proper context.

Pairing from Replacement:
We shall now see how one can get the pairing axiom from replacement, extensionality, the power set axiom, and set existence. First, note that $\mathcal{P}(\varnothing) = \{\varnothing\}$ and $\mathcal{P}(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}$. Now, given any two sets a and b, define

$$\psi(x, y, a, b) := (x = \varnothing \wedge y = a) \vee (x = \varnothing \wedge y = b).$$

Then R_ψ is a class function and so taking $z = \{\varnothing, \{\varnothing\}\}$, the image becomes $\{a, b\}$.

Exercise:
Given two sets a and b, show how to define (and prove the existence of) the sets

$$a \setminus b \qquad a \cap b \qquad a \times b = \{(x, y) | x \in a \wedge y \in b\}.$$

1.2. **Functions.**
Suppose $\psi(x, y, \vec{a})$ is a formula defining a class function R_ψ whose domain is not just a class, but a set. I.e., there is a set z such that

$$\forall x \left(\exists y\, \psi(x, y, \vec{a}) \longleftrightarrow x \in z\right).$$

Then R_ψ is itself a set, i.e., for some set u,

$$\forall x \forall y \left(\psi(x, y, \vec{a}) \longleftrightarrow (x, y) \in u\right)$$

To see this, just let v be the image of z by R_ψ. Then v is the whole image of R_ψ and

$$u = \{(x, y) \in z \times v \mid R_\psi(x, y)\}$$

is a set.
We say then that u is a function from z to v and denote this by $u : z \to v$. So functions are always a set of pairs and hence are identified with their graphs.

Definition 1.6. A function $f : a \to b$ from a set a to a set b is a subset $f \subseteq a \times b$ satisfying

$$\forall x \,(x \in a \longrightarrow \exists y \in b\ (x, y) \in f)$$

$$\forall x \forall y \forall y' \left(\left((x, y) \in f \wedge (x, y') \in f\right) \longrightarrow (y = y')\right)$$

1.3. **Families of Sets and Cartesian Products.**
Suppose $a : I \to X$ is a function. For simplicity, we write a_i for the unique set $x \in X$ such that $(i, x) \in a$. We then define

$$\bigcup_{i \in I} a_i = \left\{z \in \bigcup X \,|\, \exists i \in I,\ z \in a_i\right\}$$

$$\bigcap_{i \in I} a_i = \left\{z \in \bigcup X \,|\, \forall i \in I,\ z \in a_i\right\}$$

$$\prod_{i \in I} a_i = \text{set of all functions } f : I \to \bigcup X \text{ such that } \forall i \in I, f(i) \in a_i.$$

Exercise: Show that $\prod_{i \in I} a_i$ exists.

CAUTION:
When $I = \varnothing$, then $\bigcap_{i \in I} a_i = \bigcup X$ and hence depends on the choice of X.

2. Ordinals and Cardinals

Definition 2.1. Suppose R is a binary class relation and C is a class. We say that R defines a strict ordering of C if for all sets x, y, z we have:

$$R(x,y) \longrightarrow \big(C(x) \wedge C(y)\big)$$
$$\neg\big(R(x,y) \wedge R(y,x)\big)$$
$$\big(R(x,y) \wedge R(y,z)\big) \longrightarrow R(x,z).$$

The ordering R is total or linear if, moreover,

$$\forall x \forall y \, (C(x) \wedge C(y) \longrightarrow \big(R(x,y) \vee R(y,x) \vee x = y\big)).$$

Suppose now R is a strict linear ordering on the class C and X is a set all of whose elements belong to C. We say that X is well-ordered by R if any non-empty subset $Y \subseteq X$ has a smallest element, i.e.

$$\forall Y \Big(\big((Y \subseteq X) \wedge (\varnothing \neq Y)\big) \longrightarrow \exists y \in Y \; \forall x \in Y \, \big(x = y \vee R(y,x)\big)\Big).$$

(Note that if R is a strict ordering on a class C and X is a subset of C, i.e. a set all of whose elements belong to C, then we can identify R's restriction to X with the set $\{(x,y) \in X \times X \mid R(x,y)\}$.)

Now, suppose X is a set well-ordered by R. A subset $Y \subseteq X$ is said to be an initial segment if

$$\forall x, y \in X \Big(\big(y \in Y \wedge R(x,y)\big) \longrightarrow x \in Y\Big).$$

Also, for every $x \in X$, let $\delta_x(X)$ denote the initial segment $\delta_x(X) = \{y \in X \mid R(y,x)\}$.
Note that since R is strict, $x \notin \delta_x(X)$.

Observation:
$Y \subseteq X$ is an initial segment if and only if $Y = X$ or $Y = \delta_x(X)$ for some $x \in X$.

Proof.
If $Y \neq X$, just let x be the minimum element in $X \setminus Y$.

□

2.1. Classes and Sets.

Recall that a class C is simply the collection of all x satisfying some formula $\phi(x, \vec{a})$ with parameters. Note that we do not give classes any formal existence, in the sense that they do not belong to $\mathcal{U}$, and any statement about the class C is just a shorthand for a more complex statement involving the formula $\phi(x, \vec{a})$.

On the other hand, suppose there is a set z containing all members in C, i.e.

$$\forall x(\phi(x, \vec{a}) \longrightarrow x \in z)^{\mathrm{a}}$$

Then, by comprehension, we can identify C with the set

$$y = \{x \in z \mid \phi(x, \vec{a})\}.$$

[a] *TN:* Note that z needs only to contain C. It may have other elements.

Similarly, any set z can be identified with a class, namely the class given by the formula $\phi(x) := (x \in z)$.

Slogan:
Classes are collections that sometimes are too large to be sets, while on the other hand, all sets are classes.

Definition 2.2. A class C is a proper class when it is not a set, i.e., when there is no set whose elements are exactly those belonging to C.

Examples:

- $\mathcal{U}$ is a proper class given by the formula "$x = x$"
- Russell's class given by the formula "$x \notin x$" is also a proper class (*TN:* Assuming AF, to come later, this class *is* $\mathcal{U}$)

Notation:
If $\phi(x, \vec{a})$ is a formula with a single free variable x and parameters $\vec{a}$, we let $\{x \mid \phi(x, \vec{a})\}$ denote the, possibly proper, class defined by $\phi(x, \vec{a})$.
So, sets are the special classes given by expressions $\{x \in z \mid \phi(x, \vec{a})\}$ where z is another set.

2.2. Well-Orderings and Ordinals.

Definition 2.3. A class relation R defining a strict linear ordering of a class C is said to be a well-ordering if for any x in C, the class initial segment $\delta_x(C) = \{y \mid R(y, x)\}$ is a set that is well-ordered by R.

Definition 2.4. A set x is said to be transitive if $\forall y\, (y \in x \longrightarrow y \subseteq x)$. I.e., $z \in y \in x \implies z \in x$.

Definition 2.5. An ordinal is a transitive set α that is well-ordered by the class relation $\in$.

Lemma 2.6. Ordinals form a class relation called Ord[a], or, sometimes On.

Proof.

$$\begin{gathered}\mathbf{Ord} = \Bigg\{\alpha \mid \forall y (y \in \alpha \longrightarrow y \subseteq \alpha) \wedge \\ \forall x \forall y \bigg(\Big((x \in \alpha) \wedge (y \in \alpha) \wedge (x \neq y) \Big) \longrightarrow \Big((x \in y) \vee (y \in x) \Big) \bigg) \wedge \\ \forall x (x \in \alpha \longrightarrow x \notin x) \wedge \\ \forall x \bigg(\Big((x \subseteq \alpha) \wedge (x \neq \varnothing) \Big) \longrightarrow \exists y \in x\, \forall z \in x\, (y \in z \vee y = z) \bigg) \Bigg\}.\end{gathered}$$

□

Example:
$\varnothing, \{\varnothing\}, \big\{\varnothing, \{\varnothing\}\big\}$ are ordinals.

[a] *TN:* Not to be confused with ORD, a major Chicago airport.

Facts:
Let α be an ordinal. Then...

- ...the initial segments of α are α itself and the elements of α.
- ...so is any $\beta \in \alpha$.
- $\alpha \notin \alpha$. (For if $\gamma \in \alpha$, then $\gamma \notin \gamma$ as otherwise α would not be well-ordered by $\in$.)

Lemma 2.7. If α and β are ordinals, then either

$$\alpha \in \beta \qquad \alpha = \beta \qquad \text{or} \qquad \alpha \ni \beta.$$

Proof.

Let $\gamma = \alpha \cap \beta$. Then γ is an initial segment of α, β. For if $x \in \gamma$ and $y \in x$, then as α, β are transitive, also $y \in \alpha \cap \beta = \gamma$.
So, as any initial segment of an ordinal is either the ordinal itself or an element of it, there are four possible cases:

(1) $\gamma = \alpha = \beta$
(2) $\gamma = \alpha$ and $\gamma \in \beta$, hence $\alpha \in \beta$
(3) $\gamma = \beta$ and $\gamma \in \alpha$, hence $\beta \in \alpha$
(4) $\gamma \in \alpha$ and $\gamma \in \beta$. But γ is an ordinal, and $\gamma \in \alpha \cap \beta = \gamma$, which is impossible.

□

Proposition 2.8. The class Ord is well-ordered by the class relation $\in$.

Proof.

We know that $\in$ linearly orders Ord, for if $\alpha \in \beta \in \gamma$, then by transitivity of γ, $\alpha \in \gamma$ and irreflexivity has been checked. Moreover, for any ordinal α,

$$\delta_\alpha(\mathrm{Ord}) = \{\beta \mid \beta \text{ is an ordinal and } \beta \in \alpha\} = \alpha$$

is a set well-ordered by $\in$.

□

Lemma 2.9. Ord is a proper class.

Proof.

For if Ord was a set a, then a would itself be an ordinal and hence $a \in a$, which is impossible. (To see that a is transitive, note that for $\alpha \in \beta \in a$, since any element of an ordinal is an ordinal, also $\alpha \in a$).

□

Note:
For ordinals α, β, $\alpha \subseteq \beta \iff \alpha \in \beta$ or $\alpha = \beta$.

Lemma 2.10. If α is an ordinal, then so is $\alpha + 1 := \alpha \cup \{\alpha\}$, and, moreover, $\alpha + 1$ is the successor of α in the ordering $\in$.

Proof.

That $\alpha \cup \{\alpha\}$ is an ordinal is trivial and $\alpha \in \alpha \cup \{\alpha\}$. Also, if $\alpha \in \beta$ for β an ordinal, then $\alpha \subseteq \beta$ and so $\alpha \cup \{\alpha\} \subseteq \beta$, hence $\alpha \cup \{\alpha\} \in \beta$ or $\alpha \cup \{\alpha\} = \beta$.

□

Terminology: In the following, we will always consider the class Ord with the well-ordering $\in$, and we sometimes also write $<$ instead. That is, for ordinals α, β:

$$\alpha < \beta :\Longleftrightarrow \alpha \in \beta \Longleftrightarrow \alpha \subsetneq \beta$$

So for any ordinal α, $\alpha = \{\gamma \mid \gamma \text{ is an ordinal and } \gamma < \alpha\}$.

Lemma 2.11. If X is a set of ordinals, then $\sup X = \bigcup X$. I.e., $\bigcup X$ is an ordinal greater than or equal to all elements of X.

Proof.

$\bigcup X = \{x \mid \exists \alpha \in X, x \in \alpha\}$[a] is a set of ordinals and hence is well-ordered by $\in$. To see that $\bigcup X$ is transitive, note that if $y \in x \in \bigcup X$, then there is $\alpha \in X$ such that $x \in \alpha$, hence, by transitivity of α, also $y \in \alpha$, i.e., $y \in \bigcup X$.

Notice, if $\alpha \in X$, then $\alpha \subseteq \bigcup X$, so $\alpha \leq \bigcup X$. And if $\beta < \bigcup X$, i.e. $\beta \in \bigcup X$, then $\exists \alpha \in X, \beta \in \alpha$, i.e. β is not an upper bound for X. So $\bigcup X = \sup X$.

□

Lemma 2.12. Suppose α, β are ordinals and $f : \alpha \to \beta$ is a strictly increasing function, i.e. for $\gamma, \xi < \alpha$, $\gamma < \xi \implies f(\gamma) < f(\xi)$. Then $\alpha \leq \beta$ and $\gamma \leq f(\gamma)$ for all $\gamma < \alpha$.

Proof.

Suppose towards a contradiction that there is some $\gamma < \alpha$ such that $f(\gamma) < \gamma$. Take the minimal such γ.[b] Then, since $f(\gamma) < \gamma$ and γ is minimal, $f\big(f(\gamma)\big) \geq f(\gamma)$, but, on the other hand, since f is strictly increasing

$$f(\gamma) < \gamma \implies f\big(f(\gamma)\big) < f(\gamma),$$

which is a contradiction.

But then if $\beta < \alpha$, $\beta \leq f(\beta) \in \beta$, i.e. $\beta \leq f(\beta) < \beta$, which is a contradiction again.

□

Theorem 2.13. Suppose $f : \alpha \to \beta$ is a function which is an isomorphism of the ordered sets $(\alpha, <)$ and $(\beta, <)$. Then $\alpha = \beta$ and the isomorphism is unique, i.e. $f = \mathrm{id}_\alpha$. In other words, the structure $(\alpha, <)$ is rigid.

Proof.

$\alpha = \beta$ follows from the lemma applied to f and f^{-1}. Also, this gives us for any $\gamma < \alpha$, $\gamma \leq f(\gamma)$ and $\gamma \leq f^{-1}(\gamma)$, hence also $f(\gamma) \leq f(f^{-1}(\gamma)) = \gamma$, i.e. $\gamma = f(\gamma)$.

□

[a] *TN:* One may reasonably be concerned that this notation denotes unrestricted comprehension, as we do not stipulate that the element x be in another set. However, recall that this notation is valid for classes, and it's been established that taking the union of a set is a valid set operation.

[b] *TN:* As a quick comment, the fact that ordinals are well-ordered becomes a phenomenally useful fact it proofs, as then one can always take the minimal ordinal that satisfies some property. Keep this trick in your back pocket, or front pocket, or, really, just always on hand.

Theorem 2.14. For any well-ordered set $(X, <_X)$, there is a unique isomorphism onto an ordinal $(\alpha, <)$.

Proof.

Uniqueness: Follows from the theorem above. ■

Existence: Let $Y = \{x \in X \mid (\delta_x, <_X) \text{ is isomorphic to an ordinal}\}$[a] where $\delta_x := \delta_x(X)$.
By the uniqueness part, for any $x \in Y$, there is a unique ordinal $\beta(x)$ isomorphic to $(\delta_x, <_X)$. Note that for $x <_X y$ and $y \in Y$, by the isomorphism of δ_y with $\beta(y)$, $\delta_x \subseteq \delta_y$ is set to an initial segment of $\beta(y)$, which is itself an ordinal $\beta(x) < \beta(y)$. So $x \in Y$ and Y is an initial segment of X.

Now $x \longmapsto \beta(x)$ is a function defined on Y, so by the axiom of replacement, $Z = \{\beta(x) \mid x \in Y\}$ is a set. Moreover, Z is an initial segment of the ordinal numbers. For if $\gamma < \beta(x)$ for some $x \in Y$, then the isomorphism from δ_x to $\beta(x)$ takes some $y \in \delta_x$ to γ, hence δ_y is mapped onto the set of ordinals below γ, i.e. γ itself. So $\gamma = \beta(y)$.

So Z itself, being a set and an initial segment, is an ordinal $Z = \alpha$ and $x \in Y \longmapsto \beta(x) < \alpha$ is an isomorphism from Y to α. Now, if $Y \subsetneq X$, let x_0 be minimal in $X \setminus Y$[b], so $\delta_{x_0} = Y \cong \alpha$, contradicting that $x_0 \notin Y$.

□

There is also a class version of this theorem. Namely, if $R(x, y)$ is a class relation that well-orders a proper class C, then there is a class function F from C to Ord which is an isomorphism of the orderings R and $\in$. Just define F by $F(x) = \alpha \iff$ there is an order-isomorphism between $(\delta_x(C), R)$ and $(\alpha, \in)$.

2.3. Inductive Definitions.

Suppose $\phi(x)$ is a formula (possibly with parameters). Then

$$\phi(x) \text{ holds for every ordinal } \alpha$$

$$\text{iff}$$

$$(*) \qquad \forall\alpha \Big(\forall\beta \left(\beta < \alpha \longrightarrow \phi(\beta)\right) \longrightarrow \phi(\alpha)\Big)$$

For the nontrivial implication, if $(*)$ holds, but $\phi(\alpha)$ is not true for all ordinals, then one gets a contradiction by looking at the least α such that $\neg\phi(\alpha)$.
Whereas proving $\forall\alpha\, \phi(\alpha)$ by proving $(*)$ constitutes a proof by induction on ordinals, we can also give definitions by induction.

Suppose F is a class function of one variable and a is a subset of the domain of F, i.e., $\forall x \in \alpha\, \exists y,\ F(x) = y$. Then we let $F\restriction_a$ denote the function obtained by restricting F to a, i.e. for $b = \{F(x) \mid x \in a\}$, which is a set by replacement, $F\restriction_a = \{(x, y) \in a \times b \mid F(x) = y\}$.

Now, suppose H is any class function of one variable. We say that a function f is H-inductive if:

- $\alpha = \mathrm{dom}(f)$ is an ordinal, and
- $\forall\beta < \alpha$, $f\restriction_\beta$ is in the domain of H and $f(\beta) = H(f\restriction_\beta)$.

[a] *TN:* Perhaps it ought to be mentioned that these sets where English is used in the set notation are also well-defined. One needs only to find first order formulas which describe the English appropriately, and this is a fairly simple exercise for the reader should they not already be convinced of this truth.

[b] *TN:* I told you this would come up a lot.

So we can think of H as associating to each function $f : \beta \to X$ defined on an ordinal β an extension $\widetilde{f} : \beta + 1 \to \widetilde{X}$, by $\tilde{f}(\beta) = H(f)$. An H-inductive function just satisfies this equation whenever defined.

Lemma 2.15. For any class function H and ordinal α, there is at most one H-inductive function f with $\mathrm{dom}(f) = \alpha$.

Proof.

Suppose $f : \alpha \to X$ and $g : \alpha \to Y$ are distinct H-inductive functions and let $\beta < \alpha$ be minimal such that $f(\beta) \neq g(\beta)$. Then $f\restriction_\beta = g\restriction_\beta$ and so $f(\beta) = H(f\restriction_\beta) = H(g\restriction_\beta) = g(\beta)$, a contradiction.

□

Lemma 2.16. Suppose H is a class function and α is an ordinal such that any function $f : \beta \to X$ with domain β will belong to the domain of H. Then there is an H-inductive function $g : \alpha \to Y$.

Proof.

Let $\tau = \{\beta < \alpha \mid \text{there is an } H\text{-inductive function } f_\beta : \beta \to X\}$. Then τ is a set and is seen to be an initial segment of α, hence τ is an ordinal $\leq \alpha$. Moreover, by uniqueness, the assignment $\beta < \tau \longmapsto f_\beta$ is well-defined, and since for any $\gamma < \beta$, $f_\beta\restriction_\gamma$ is also H-inductive, by uniqueness, we have

$$\gamma < \beta < \tau \implies f_\beta\restriction_\gamma = f_\gamma.$$

Thus, $f = \bigcup_{\beta<\tau} f_\beta$ is an H-inductive function with domain $\sigma = \sup_{\beta<\tau} \beta = \bigcup_{\beta<\tau} \beta$.

Assume towards a contradiction that $\sigma < \alpha$. Then $\tilde{f}$ defined by $\tilde{f}\restriction_\sigma = f$ and $\tilde{f}(\sigma) = H(f)$ is an H-inductive function with domain $\sigma + 1 = \sigma \cup \{\sigma\} \notin \tau$, which is impossible.

□

With only slight adjustments, it suffices that for some class A, H takes values in A and any function $f : \beta \to X$, where X is a sub<u>set</u> of A, belongs to the domain of H.

Theorem 2.17. Suppose A is a class and M is the class of all functions $f : \alpha \to X$ with domain an ordinal and range X a subset of A. Suppose H is a class function of one variable defined on all of M and with values in A. Then there is a unique class function F, i.e., given by a formula of set theory, such that

- F is defined on Ord
- $\forall\alpha\, F(\alpha) = H(F\restriction_\alpha)$

Proof.

The class function F is defined by

$$y = F(\alpha) \iff \text{there is an } H\text{-inductive function } f : \alpha \to X, X \subseteq A, \text{ and } y = H(f).$$

□

2.4. **Stratified or Ranked Classes.**

A class W is said to be <u>stratified</u> or <u>ranked</u> if there is a class function ρ with domain W and taking values in Ord such that for any ordinal α, the following class is a set.

$$W_\alpha = \{x \mid W(x) \wedge \rho(x) < \alpha\}$$

Theorem 2.18. Suppose W is a stratified class with corresponding stratification ρ. Also let M be the class of all functions with domain W_α for some α and $H(\alpha, f) = y$ a class function with domain $W \times M$. Then there is a unique class function F with domain W such that for any a in W: $F(a) = H(a, F\restriction_{W_{\rho(a)}})$.

Note:
For a in W, $a \in W_{\rho(a)+1} \setminus W_{\rho(a)}$.

Proof.
Set $W'_\alpha = \{x \mid W(x) \wedge \rho(x) = \alpha\}$ and note that W'_α is a set and $W_\alpha = \bigcup_{\beta<\alpha} W'_\beta$. By induction on Ord, i.e., by applying the preceding theorem, we find a class function G defined on ordinals such that for every α, $G(\alpha) =$ "the function ϕ with domain W'_α such that $\phi(x) := H\Big[x, \bigcup_{\beta<\alpha} G(\beta)\Big]$ for all $x \in W'_\alpha$".
(Note that $G(\alpha)$ is written as a function of $G\restriction_\alpha$, so the theorem applies.)

Also, for any a in W,

$$G\big(\rho(a)\big) = H[-, \bigcup_{\beta<\rho(a)} G(\beta)] \text{ with domain } W'_{\rho(a)}$$

We can now write

$$F(a) = b \iff W(a) \wedge b = G\big(\rho(a)\big)(a).$$

To see that this works, note that $F\restriction_{W_\alpha} = \bigcup_{\beta<\alpha} G(\beta)$; for if $a \in W_\alpha$, say $a \in W'_\beta$ for some $\beta = \rho(a) < \alpha$ and so $F(a) = G(\beta)(a)$.
It thus follows that for a in W, $F(a) = G\big(\rho(a)\big)(a) = H[a, F\restriction_{W_{\rho(a)}}]$.
Uniqueness is an exercise.

□

(8) Axiom of Choice or (AC)

For any set X and $A \subseteq \mathcal{P}(X)$ set of pairwise disjoint non-empty subsets of X, there is a set $T \subseteq X$ such that $\forall a \in A$, $T \cap a$ contains exactly one element.

Define also the statements

(AC′) For any set X, there is a function $\pi : \mathcal{P}(X) \setminus \{\varnothing\} \to X$ such that $\pi(a) \in a$ for every $\varnothing \neq a \subseteq X$.

(AC″) If $(X_i)_{i\in I}$ is an indexed family of non-empty sets, then $\prod_{i\in I} X_i \neq \varnothing$.

Proposition 2.19. AC$\iff$AC′ $\iff$AC″ (given the background theory of axioms (1)–(7))[a]

[a] *TN:* The proof of this is omitted from the original notes, and I do not provide it here, as this would make for a good exercise. As a further note, there are many other statements to come which are also equivalent to the axiom of choice. The reader is encouraged to attempt equivalence proofs of these later statements.

Theorem 2.20 (Zermelo)**.** Every set can be well-ordered.

Proof.

Suppose X is a set and let $\pi : \mathcal{P}(X) \setminus \{\varnothing\} \to X$ be a choice function, i.e., $\pi(a) \in a$ for all non-empty $a \subseteq X$.
Assume for contradiction that X cannot be well-ordered. Let H be the class function defined by

$$H(f) = y \iff f \text{ is a function with } \mathrm{dom}(f)=\alpha \text{ an ordinal, } \mathrm{Im}(f) \subsetneq X \text{ and } y = \pi\big(X \setminus \mathrm{Im}(f)\big).$$

($\star$) H is defined on the class of H-inductive functions. Also, H-inductive functions are injective. For if $f : \alpha \to X$ is H-inductive, then for any $\beta < \alpha$, $f(\beta) = H(f\restriction_\beta) \in X \setminus \mathrm{Im}(f\restriction_\beta)$, hence for any $\gamma < \beta$, $f(\gamma) \neq f(\beta)$. It thus follows that f is injective. If also f were surjective, then this would induce a well-ordering of X, contradicting our assumption.

By ($\star$), we know that there is an H-inductive class function $F : \mathrm{Ord} \to X$, which is injective by ($\star$). But then since $\mathrm{Im}(f)$ is a set (*TN:* Why?), $F^{-1} : \mathrm{Im}(F) \to \mathrm{Ord}$ is a function from a set onto Ord, which is impossible.

□

Note:
Any well-ordered set admits a choice function.

Theorem 2.21 (Zorn's Lemma)**.** Suppose $(X, \leqslant)$ is a partially ordered set (or poset) all of whose linearly ordered subsets admits an upper bound. Then $(X, \leqslant)$ has a maximal element, i.e. there is $y \in X$ such that $\forall x \in X$, $y \not< x$.

Proof.

Let $A = \{Y \subseteq X \mid \exists x \in X, \forall y \in Y, y < x\}$ and let $\pi : \mathcal{P}(X) \setminus \{\varnothing\} \to X$ be a choice function.
Define $p : A \to X$ by $p(Y) = \pi\big(\{x \in X \mid \forall y \in Y, y < x\}\big)$.
Define a class function H by

$$H(f) = y \iff f \text{ is a function with } \mathrm{dom}(f) \text{ an ordinal and } \mathrm{Im}(f) \in A \text{ and } y = p\big(\mathrm{Im}(f)\big).$$

(†) Any H-inductive function $f : \alpha \to X$ is strictly increasing, i.e., $\beta < \gamma < \alpha \implies f(\beta) < f(\gamma)$.

It follows from (†) that the image of any H-inductive function $f : \alpha \to X$ is linearly ordered, and hence has an upper bound $x_f \in X$, i.e. $\forall \beta < \alpha$, $f(\beta) < x_f$. Now, if $f : \alpha \to X$ is H-inductive, but H is not defined on f, then $\mathrm{Im}(f)$ has no strict majorant, hence $f(\beta) = x_f$ for some $\beta < \alpha$ and x_f thus is the maximum in $\mathrm{Im}(f)$ and a maximal element of X.

If, on the other hand, H is defined on the class of H-inductive functions, then we can find an H-inductive class function $F : \mathrm{Ord} \to X$, which by (†) is strictly increasing. But then F would define an injection of a proper class into the set X, which is impossible.

□

Proposition 2.22. Zorn's Lemma $\Longrightarrow$ (AC):

Proof.

Suppose A is a collection of pairwise disjoint subsets of a set X and let

$$B = \{T \subseteq X \mid T \text{ is a partial transversal for } A^{\text{a}}\}$$

We order B by inclusion and note that any linearly ordered subset of B has an upper bound. So, by Zorn's Lemma, B has a maximal element, which is seen to be a transversal for A, i.e. intersects every element of A in a singleton.

$\square$

TN: Before moving onto ordinal arithmetic, it should be noted that Zorn's Lemma, Zermelo's Theorem, and the three statements for choice are all equivalent. The reader should verify this if they have not already done so. Further questions and observations include wondering what is yellow and equivalent to the axiom of choice,[b] and noting that Jesus was in fact pro-choice.[c]

2.5. **Ordinal Arithmetic.** Recall that if α is an ordinal, the successor of α, i.e. the smallest ordinal strictly larger than α, is $\alpha + 1 = \alpha \cup \{\alpha\}$.

Definition 2.23. An ordinal β is a successor ordinal if $\beta = \alpha + 1$ for some ordinal α.
Let also 0 denote the smallest ordinal, i.e. $0 = \varnothing$.
β is a limit ordinal if $\beta \neq 0$ and β is not a successor.
β is a natural number or a finite ordinal if $\forall \alpha \leq \beta(\alpha = 0 \vee \alpha \text{ is a successor})$.
Let also $n = (\cdots(0 + \underbrace{1) + 1) + \cdots + 1)}_{n \text{ times}}$[d]

Note that if $(X, <_X)$ and $(Y, <_Y)$ are well-ordered sets, then we can define well-orderings on the sets

$$(X \times \{0\}) \cup (Y \times \{1\}) \qquad \text{and} \qquad X \times Y$$

by

$$(a, i) < (b, j) \Longleftrightarrow \begin{cases} i = j = 0 \text{ and } a <_X b, \text{ or} \\ i = j = 1 \text{ and } a <_Y b, \text{ or} \\ i < j \end{cases}$$

and

$$(x_0, y_0) < (x_1, y_1) \Longleftrightarrow \begin{cases} y_0 < y_1, \text{ or} \\ (y_0 = y_1) \wedge (x_0 < x_1) \end{cases}$$

The first ordering, i.e. on $(X \times \{0\}) \cup (Y \times \{1\})$, is said to be the sum of $(X, <_X)$ and $(Y, <_Y)$, while the latter is the product of the two.

[a] I.e. $\forall a \in A, T \cap a$ has at most a single element

[b] Zorn's lemon!

[c] Because how else would you explain the loaves and fishes without Banach-Tarski!

[d] *TN:* If this seems circular, that's because it is. We get away with this by saying that it's an intuitive definition and we don't care if it's *technically* circular.

Definition 2.24. For ordinals α, β

- $\alpha + \beta$ denotes the unique ordinal γ that is order-isomorphic to the sum of α and β.
- $\alpha\beta$ or $\alpha \cdot \beta$ denotes the unique ordinal γ that is order-isomorphic to the product of α and β.

Lemma 2.25. $+$ is associative and 0 is a two-sided additive identity. $\cdot$ is associative, $\alpha \cdot 0 = 0$, $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$, $\alpha \cdot (\beta + \gamma) = \alpha\beta + \alpha\gamma$, and if λ is a limit, then $\alpha\lambda = \sup_{\beta<\lambda} \alpha\beta$.[a]

For all we know till now, multiplication could be commutative, but this fails after the next axiom.

(9) Axiom of Infinity[b]

There is an ordinal which is not a natural number, i.e. an infinite ordinal

$$\exists \alpha \Big(\alpha \text{ is an ordinal} \wedge \exists \beta \leq \alpha \, (\beta \neq 0 \wedge \beta \text{ is not a successor}) \Big)$$

Note that the natural numbers form an initial segment of the ordinals. So, if we let ω denote the smallest infinite ordinal, we see that ω is a limit ordinal.

Remark:
$\omega \cdot 2 = \omega + \omega \neq 2 \cdot \omega = \omega$
$\omega + 2 \neq \omega = 2 + \omega$

Exponentiation: We define α^β by recursion on β uniformly on α by

- $\alpha^0 = 1$[c]
- $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$
- $\alpha^\lambda = \sup_{\beta<\lambda} \alpha^\beta$ whenever λ is a limit ordinal.

Note:
Formally, we can see ordinal exponentiation as being given by a class function
Exp : Ord $\times$ Ord $\to$ Ord defined by the first order formula

$$\operatorname{Exp}(\alpha,\beta) = \gamma \iff \begin{cases} \text{there is a function } f : \beta+1 \to \mathrm{Ord} \text{ such that } \forall \zeta \leq \beta \\ \quad \bullet\, f(\beta) = \gamma \\ \quad \bullet\, \zeta = 0 \implies f(\zeta) = 1 \\ \quad \bullet\, \zeta = \delta + 1 \implies f(\zeta) = f(\delta) \cdot \alpha \\ \quad \bullet\, \zeta \text{ a limit} \implies f(\zeta) = \sup_{\delta<\zeta} f(\delta). \end{cases}$$

The proof of uniqueness and existence of f follows along the same lines as the general proof for inductive definitions (or can be deduced from it directly).

[a] *TN:* The proof of this is omitted, but it is simple to check.
[b] *TN:* Finitists, beware!
[c] *TN:* Don't tell the analysts, but this does actually mean that $0^0 = 1$. Ok...Fine. *Technically* it's just shorthand for something that looks an awful lot like exponentiation, and we just stipulate that $\varnothing^\varnothing = \{\varnothing\}$, but that's way less fun than pissing off an analyst.

2.6. **Cardinals and Their Arithmetic.**

Definition 2.26. Given a set X, we define the cardinality of X, denoted $|X|$ or $\text{card}(X)$, to be the smallest ordinal κ such that there is a bijection from X to κ.[a]

Definition 2.27. Two sets are said to be equinumerous if there is a bijection between them. So X and Y are equinumerous if and only if $|X| = |Y|$.

Theorem 2.28. Suppose X and Y are non-empty sets. The following are equivalent:

(i) There is an injection from X to Y
(ii) There is a surjection from Y to X
(iii) $|X| \leq |Y|$.

Note:
$|X|$ and $|Y|$ are defined using (AC).
(i) $\implies$ (ii) does not require (AC)
(ii) $\implies$ (i) does.

TN: Even further, the statement "every set X has a cardinality" is equivalent to (AC). To see this, we actually prove the equivalence to Zermelo's theorem, which is equivalent to AC. Note that if every set can be well-ordered, then there is a bijection from a set to a well-ordered set, and as seen previously, an isomorphism from a well-ordered set to an ordinal, and then of course, from an ordinal to its cardinality. On the other hand, if an arbitrary set X has a cardinality, then there is a bijection from it to an ordinal, and this induces a well-ordering on X.

A common joke is that Zermelo's theorem is clearly false, the axiom of choice is clearly true, and Zorn's lemma is too difficult to understand to say one way or another. I find this equivalence of Zermelo's theorem to the fact that every set has a cardinality very compelling in convincing one of the truth of Zermelo's theorem.

Theorem 2.29 (Cantor-Schröder-Bernstein)**.** X and Y are equinumerous if and only if X injects into Y and vice versa.

Proof.

This is trivial given (AC). But it is possible to give a proof without. (*TN:* The reader is encouraged to try this for themselves. A back-and-forth technique should prove useful.)

□

Theorem 2.30 (Cantor)**.** For any set X, $|X| < |\mathcal{P}(X)|$.

Proof.

Suppose for contradiction that there is a surjection $\pi : X \twoheadrightarrow \mathcal{P}(X)$ and define $Y := \{x \in X \mid x \notin \pi(x)\}$. Then if $Y = \pi(y)$, we have

$$y \in \pi(y) \iff y \in Y \iff y \notin \pi(y)$$

which is impossible.

□

Definition 2.31. An ordinal number κ is said to be a cardinal if $|\kappa| = \kappa$.

[a] *TN:* It's not immediately obvious why the cardinality of a set should always exist. That is, why must there exist *any* bijection to an ordinal for every set? This follows from Zermelo's Theorem, which is equivalent to AC, which we've included as an axiom.

Corollary 2.32. The class of cardinal numbers is a proper class.

Proof.

Suppose towards a contradiction that A were the set of all cardinal numbers. Then $\gamma = \sup A$ is an ordinal and $\kappa = |\mathcal{P}(\gamma)| > \gamma \geq \lambda$ for any $\lambda \in A$. Since κ is a cardinal, this is impossible. □

Definition 2.33. A set X is finite if $|X|$ is a finite ordinal and is infinite otherwise. So, X is infinite if and only if ω injects into X.

Proposition 2.34 (Galileo)**.** A set X is infinite if and only if it properly injects into itself.

Proof.

For one direction, it suffices to note that $\alpha \longmapsto \alpha + 1$ is a proper injection of ω into itself.
For the converse, one shows by induction on finite ordinals that they are cardinal numbers. □

The $\aleph$-Function:
The class of infinite cardinals, being cofinal in the proper class of ordinals,[a] is itself a proper class well-ordered by the usual ordering of ordinals. It thus follows that there is a uniquely defined class function $\aleph : \mathrm{Ord} \to$ (Class of infinite cardinals) preserving the ordering.

For example, $\aleph_0 = \omega$ and $\aleph_{\alpha+1}$ is the smallest cardinal number larger than $\aleph_\alpha$.

Since the cardinal numbers are well-ordered and unbounded, for any cardinal number κ there is a smallest cardinal greater than κ, which we denote κ^+. So, for example, $n^+ = n+1$ for n finite, and $\aleph_\alpha^+ = \aleph_{\alpha+1}$.

Note:
For any ordinal γ, we have $|\gamma| \leq \gamma < |\gamma|^+$. For if $|\gamma|^+ \leq \gamma$, then $|\gamma|^+ \subseteq \gamma$ and thus $|\gamma|^+$ would inject into the smaller cardinal $|\gamma|$, which is impossible.

Definition 2.35. Cardinals of the form κ^+ are called successors, while non-zero non-successors are called limit cardinals.

Proposition 2.36. The function $\aleph$ is continuous with respect to the order-topology, that is, for any limit ordinal λ,

$$\aleph_\lambda = \sup_{\xi<\lambda} \aleph_\xi$$

Proof.

Set $\gamma = \sup_{\xi<\lambda} \aleph_\xi$. Then $|\gamma| \leq \gamma < |\gamma|^+$. Now, if $\gamma < \aleph_\lambda$, then there is some $\xi_0 < \lambda$ such that $\aleph_{\xi_0} = |\gamma|$, hence

$$\sup_{\xi<\lambda} \aleph_\xi = \gamma < |\gamma|^+ = \aleph_{\xi_0+1} \leq \sup_{\xi<\lambda} \aleph_\xi$$

which is impossible. So

$$\sup_{\xi<\lambda} \aleph_\xi \leq \aleph_\lambda \leq \sup_{\xi<\lambda} \aleph_\xi.$$

□

[a] *TN:* Cofinality is defined a bit later, but all you need to know is essentially that if you pick some ordinal, there's a cardinal above it.

Definition 2.37. A set X is countable if $|X| \leq \aleph_0$ and uncountable otherwise.

The Continuum Hypothesis (CH) is the statement $|\mathcal{P}(\omega)| = \aleph_1$.[a]

Cardinal Arithmetic.

Definition 2.38. For cardinal numbers κ and λ, we set

$$\kappa \otimes \lambda := |\kappa \times \lambda|$$
$$\kappa \oplus \lambda := \Big|\big(\kappa \times \{0\}\big) \cup \big(\lambda \times \{1\}\big)\Big|$$

Note that both cardinal multiplication and addition are commutative and associative.

Theorem 2.39. If κ is an infinite cardinal number, then $\kappa \otimes \kappa = \kappa$.[b]

Proof.

This is by induction on κ (i.e., by induction on α in $\kappa = \aleph_\alpha$). So suppose that for all $\beta < \kappa$,

$$|\beta \times \beta| = |\beta| \otimes |\beta| = |\beta|$$

and define a well-ordering $\prec$ on $\kappa \times \kappa$ by

$$(\alpha, \beta) \prec (\alpha', \beta') \iff \begin{cases} \max(\alpha, \beta) < \max(\alpha', \beta') \\ \max(\alpha, \beta) = \max(\alpha', \beta') \wedge (\alpha < \alpha') \\ \max(\alpha, \beta) = \max(\alpha', \beta') \wedge (\alpha = \alpha') \wedge (\beta < \beta') \end{cases}$$

Since $\kappa \times \kappa$ is the increasing union of $\xi \times \xi$ for $\xi < \kappa$, each $\xi \times \xi$ is an initial segment of $(\kappa \times \kappa, \prec)$ and $(\xi \times \xi, \prec)$ is isomorphic to some ordinal γ with $|\gamma| = |\xi \times \xi| = |\xi| < \kappa$, and thus $\gamma < \kappa$, we see that $(\kappa \times \kappa, \prec)$ is order-isomorphic to κ with the usual order. So $|\kappa \times \kappa| = \kappa$.

$\square$

Corollary 2.40. $\kappa \oplus \kappa = \kappa$ for all infinite κ.[c]

Proof.

$\kappa \oplus \kappa = |\kappa \times 2| \leq |\kappa \times \kappa| = \kappa \otimes \kappa = \kappa$.

$\square$

Definition 2.41. For cardinal numbers κ and λ, let $\kappa^\lambda = \Big|\{f \mid f \text{ is a function from } \lambda \text{ to } \kappa\}\Big|$.

Thus, $2^\kappa = |\mathcal{P}(\kappa)|$ by identifying a set with its characteristic function.[d]

---

[a] *TN:* This statement is well-known to be independent of the axioms of ZFC (meaning there are models in which it is true, and models in which it is false). We shall prove this by the end of these notes.

[b] *TN:* This is not quite equivalent to choice, but the statement that for every infinite *set* A, that $A \times A$ and A are equinumerous *is* equivalent to choice.

[c] *TN:* This is *not* equivalent to AC. Try a Cantor-Schröder-Bernstein argument.

[d] *TN:* Yes, this is ambiguous with ordinal exponentiation. In general, this tends to be the more common usage of this notation, and if you mean ordinal exponentiation, you should specify. Well, in general, you should always specify, but c'est la vie.

Lemma 2.42. If $\lambda \geq \omega$ and $2 \leq \kappa \leq \lambda$, then $\kappa^\lambda = 2^\lambda$.

Proof.

Note that $2^\lambda = 2^{\lambda \otimes \lambda} = (2^\lambda)^\lambda \geq \lambda^\lambda \geq \kappa^\lambda$.

□

Definition 2.43. A function $f : \alpha \to \beta$ between ordinals α and β is said to be cofinal if $\mathrm{Im}(f)$ is unbounded in β, i.e., $\forall \gamma < \beta, \exists \xi < \alpha, f(\xi) \geq \gamma$.

The cofinality of an ordinal β, denoted $\mathrm{cof}(\beta)$, is the smallest ordinal α such that there is a cofinal map from α to β. Thus, $\mathrm{cof}(\beta) \leq \beta$.

Note:
$\mathrm{cof}(\beta)$ is always a cardinal.

Observation:
There is a strictly increasing cofinal map $h : \mathrm{cof}(\beta) \to \beta$.

Proof.

If $f : \mathrm{cof}(\beta) \to \beta$ is cofinal, define $h(\xi) := \max\Big(f(\xi), \sup_{\gamma<\xi}\big(h(\gamma)+1\big)\Big)$.

□

Proposition 2.44. If α and β are limit ordinals and $f : \alpha \to \beta$ is a strictly increasing cofinal map, then $\mathrm{cof}(\alpha) = \mathrm{cof}(\beta)$.

Proof.

The fact that $\mathrm{cof}(\beta) \leq \mathrm{cof}(\alpha)$ is clear. Conversely, if $g : \mathrm{cof}(\beta) \to \beta$ is a cofinal map, then we can define $h : \mathrm{cof}(\beta) \to \alpha$ by

$$h(\xi) = \min\big(\gamma < \alpha \mid f(\gamma) > g(\xi)\big).$$

Then h is cofinal in α.

□

FIGURE 1. Rosendal's Visualization of the Proof

Corollary 2.45. For α a limit, $\mathrm{cof}(\aleph_\alpha) = \mathrm{cof}(\alpha)$.

Corollary 2.46. For any β, $\mathrm{cof}(\beta) = \mathrm{cof}(\mathrm{cof}(\beta))$.

So, the cofinality operation is idempotent, so has plenty of fixed points.

Definition 2.47. An ordinal β is said to be regular if $\mathrm{cof}(\beta) = \beta$ and β is a limit.

Note:
Regular ordinals are cardinals and the first regular cardinal is ω.

TN: As the next lemma will show, "most" cardinals are regular (in a loose sense of the word "most"), but regular cardinals become very nice to work with when doing more advanced set theory.

Lemma 2.48. κ^+ is regular for $\kappa \geq \omega$.

Proof.

If $f : \alpha \to \kappa^+$ is a cofinal map, then $\kappa^+ = \sup_{\gamma<\alpha} f(\gamma) = \bigcup_{\gamma<\alpha} f(\gamma)$. Since each ordinal $f(\gamma)$ is a set of cardinality $< \kappa^+$, i.e. $\leq \kappa$, we see that

$$\kappa^+ = \left|\bigcup_{\gamma<\alpha} f(\gamma)\right| \leq |\alpha \times \kappa| = \max\big(|\alpha|, |\kappa|\big)$$

and so $|\alpha| = \kappa^+$, hence $\kappa^+ \leq \alpha$.

□

Note:
Since for a limit ordinal α, we have $\mathrm{cof}(\aleph_\alpha) = \mathrm{cof}(\alpha)$, if $\aleph_\alpha$ is regular, then

$$\aleph_\alpha = \mathrm{cof}(\aleph_\alpha) = \mathrm{cof}(\alpha) \leq \alpha \leq \aleph_\alpha,$$

so $\aleph_\alpha = \alpha$.

Definition 2.49. A cardinal κ is weakly inaccessible if κ is a regular limit cardinal $> \omega$.
Moreover, κ is (strongly) inaccessible if $\kappa > \omega$, κ is regular, and for any $\lambda < \kappa$, $2^\lambda < \kappa$.

One cannot prove with the axioms given (nor with the axiom of foundation) that weakly inaccessible cardinals exist.

Lemma 2.50 (König)**.** If κ is an infinite cardinal and $\lambda \geq \mathrm{cof}(\kappa)$, then $\kappa^\lambda > \kappa$.

TN: We first give a proof of this Lemma, and then state König's Theorem (excluded from the original notes) along with a proof sketch, from which this Lemma is fairly simple to conclude. The importance of König's Theorem is partially given anecdotally by Professor Tserunyan, who mentioned that someone's question at a conference was quickly solved with this theorem, and also motivated by the fact that it is also equivalent to choice!

Proof.

Fix a cofinal map $f : \lambda \to \kappa$ and consider any function $G : \kappa \to \kappa^\lambda$. We have to show that G is not surjective. We think of κ^λ as the set of functions from λ to κ. So define $h : \lambda \to \kappa$ by

$$h(\xi) = \min(\kappa \setminus \big\{G(\alpha)(\xi) \mid \alpha \leq f(\xi)\big\})$$

Note then that if $h = G(\alpha)$ for some $\alpha < \kappa$, pick $\xi < \lambda$ such that $f(\xi) \geq \alpha$. Then $h(\xi) \neq G(\alpha)(\xi)$, which is a contradiction.

□

Corollary 2.51. If $\lambda \geq \omega$, then $\mathrm{cof}(2^\lambda) > \lambda$.

Proof.

$(2^\lambda)^\lambda = 2^{\lambda \otimes \lambda} = 2^\lambda$, so if $\lambda \geq \mathrm{cof}(2^\lambda)$, then the lemma would give

$$2^\lambda = (2^\lambda)^\lambda > 2^\lambda,$$

a contradiction.

□

Notation:[a]
Let I be a set and $(A_i)_{i\in I}$ be a sequence of sets. We denote by $\bigsqcup_{i\in I} A_i$ the disjoint union of A_i, i.e.,

$$\bigsqcup_{i\in I} A_i := \big\{(i,a) : a \in A_i, i \in I\big\}.$$

The product, $\prod_{i\in I} A_i$ is defined in the usual way.

Theorem 2.52 (König)**.** Let I be a set and $(A_i)_{i\in I}, (B_i)_{i\in I}$ be sequences of sets. If, for all $i \in I$, we have that $|A_i| < |B_i|$, then

$$\Big|\bigsqcup_{i\in I} A_i\Big| < \Big|\prod_{i\in I} B_i\Big|.$$

Proof.

First, use (AC) more than once to define an injection $\bigsqcup_{i\in I} A_i \hookrightarrow \prod_{i\in I} B_i$.
Suppose towards a contradiction that there is a surjection $g : \bigsqcup_{i\in I} A_i \twoheadrightarrow \prod_{i\in I} B_i$, and then define $x \in \prod_{i\in I} B_i$ that is not in the image of g by choosing the value $x(i)$ such that it precludes x from being in the g-image of $\{i\} \times A_i$. The main point is that for each $i \in I$, the map $g_i : A_i \to B_i$ defined by $a \longmapsto g(i,a)(i)$ is not surjective, so one can choose a value from $B_i \setminus g_i[A_i]$ as $x(i)$ (where $g_i[A_i]$ denotes the image of A_i under the map g_i).

□

2.7. **Foundation.**

(10) Axiom of Foundation or (AF)
The axiom of foundation states that $\in$ admits no infinite descending chains, or, equivalently

$$\forall x \Big((x \neq \varnothing) \longrightarrow \exists y \in x\ \forall z \in y, z \notin x\Big)$$

Note that, if $(u_n)_{n\in\omega}$ were an infinite sequence, i.e., formally a function with domain ω, such that $u_{n+1} \in u_n$ for every $n \in \omega$, then the set $x = \{u_n\}_{n\in\omega}$ would contradict (AF).

Note:
(AF) $\Longrightarrow \forall x,\ x \notin x$.

Without assuming the axiom of foundation, we can define a class function $V : \mathrm{Ord} \to \mathcal{U}$ by transfinite induction:

$$V_\beta := \bigcup_{\alpha<\beta} \mathcal{P}(V_\alpha).$$

[a] *TN:* Here I am explicitly following the notation and proof sketch given during a homework exercise in the class I took with Professor Tserunyan.

Thus, $V_0 = \varnothing$ and $\alpha \leq \beta \implies V_\alpha \subseteq V_\beta$. From this, it follows that $V_{\beta+1} = \bigcup\limits_{\alpha \leq \beta} \mathcal{P}(V_\alpha) = \mathcal{P}(V_\beta)$.
Then, for λ a limit, we get $V_\lambda = \bigcup\limits_{\alpha<\lambda} V_\alpha$.
Let also V be the class defined by

$$V(x) \iff \exists \alpha \text{ an ordinal such that } x \in V_\alpha.$$

Definition 2.53. For every set x in V, we let $\mathrm{rank}(x) = \mathrm{rk}(x) = \min(\alpha \mid x \in V_\alpha)$. Note that $\mathrm{rk}(x)$ is always a successor ordinal.

Lemma 2.54. $V(x) \iff \forall y \in x, V(y)$.
Also, if $V(x)$, then $\mathrm{rk}(y) < \mathrm{rk}(x)$ for all $y \in x$.

Proof.

Suppose $V(x)$ and $\mathrm{rk}(x) = \beta + 1$. Then $x \in V_{\beta+1} = \mathcal{P}(V_\beta)$ and so $x \subseteq V_\beta$. Thus, $y \in x$ gives $y \in V_\beta$ and $\mathrm{rk}(y) \leq \beta < \mathrm{rk}(x)$.

Conversely, if $\forall y \in x, V(y)$, the class function rk: $V \to \mathrm{Ord}$ will be bounded on the set x, hence for some β, $x \subseteq V_\beta$. Thus, $x \in V_{\beta+1} \implies V(x)$.

□

One can check (as an exercise) that

Lemma 2.55. For every ordinal α, $V(\alpha)$ and $\mathrm{rk}(\alpha) = \alpha + 1$.

Theorem 2.56. AF $\iff \forall x$, $V(x)$.

Proof.

($\Leftarrow$) : Suppose $\forall x, V(x)$ and let a be any non-empty set. Let $b \in a$ be a set of minimal rank. Then for any $c \in b$, we see that $\mathrm{rk}(c) < \mathrm{rk}(b) \implies c \notin a$ (since b had minimal rank), thus showing AF. ■

For the converse, we need the following definition.

Definition 2.57. For any set X, define a function f with domain ω by $f(0) = X$, $f(n+1) = \bigcup_{x \in f(n)} x = \bigcup f(n)$. We let $\mathrm{cl}(X) = \bigcup_{n \in \omega} f(n)$. Then notice $X \subseteq \mathrm{cl}(X)$ and $\mathrm{cl}(X)$ is transitive. Moreover, if Z is any transitive set containing X, then $Z \supseteq \mathrm{cl}(X)$. So $\mathrm{cl}(X)$ is the unique <u>transitive closure</u> of X.

($\Rightarrow$) : Now suppose that X does not belong to V, but that AF holds. Then $X \subseteq \mathrm{cl}(X)$ and we claim that $Y = \{y \in \mathrm{cl}(X) \mid \neg V(y)\}$ is non-empty. For if not, then $\mathrm{cl}(X)$ and hence also X would be a subset of V and thus belong to V itself, which is not the case.
On the other hand, if $y \in Y$, then $\neg V(y)$ and hence y cannot be a subset of V. Thus, for some $z \in y$, $\neg V(z)$. Since $\mathrm{cl}(X)$ is transitive, also $z \in \mathrm{cl}(x)$ and thus $z \in Y$ too. It follows that Y is a counterexample to AF.

□

Proposition 2.58. $\mathrm{cl}(X) = X \cup \bigcup_{y\in X} \mathrm{cl}(y)$.

Proof.

$X \subseteq \mathrm{cl}(X)$ and the latter is transitive, so $y \subseteq \mathrm{cl}(x)$ for all $y \in X$. Since $\mathrm{cl}(y)$ is the minimal transitive set containing y, also $\mathrm{cl}(y) \subseteq \mathrm{cl}(x)$ for any $y \in X$. So $\mathrm{cl}(X) \supseteq X \cup \bigcup_{y\in X} \mathrm{cl}(y)$.

For the other direction, it suffices to note that the right hand side is transitive. □

Definition 2.59. The theory ZF^- consists of all axioms given previously except for AC and AF. Moreover,

$\mathrm{ZF} = \mathrm{ZF}^- + \mathrm{AF}$
$\mathrm{ZFC} = \mathrm{ZF} + \mathrm{AC} = \mathrm{ZF}^- + \mathrm{AF} + \mathrm{AC}$
$\mathrm{ZFC}^- = \mathrm{ZF}^- + \mathrm{AC}$.

Definition 2.60. A set X is extensional if

$$\forall x, y \in X \ \big((x \cap X = y \cap X) \longrightarrow x = y\big)$$

That is, $(X, \in) \models$ axiom of extensionality.

Note that every transitive set is extensional simply because the axiom of extensionality holds.

Theorem 2.61 (The Mostowski Collapse (ZF))**.** For any extensional set X, there is a unique isomorphism $\pi : (X, \in) \to (Y, \in)$ onto a transitive set Y.

Proof.

Uniqueness: Follows by induction on the rank of elements of X.
Existence: We define $\pi(x)$ by induction on the rank of $x \in X$.

$$\pi(x) = \big\{\pi(a) : a \in x \cap X\big\}$$

(Formally, this is induction on a stratified class).

That is, note that by extensionality of X, there is a unique element $a_0 \in X$ of minimal rank. Set $\pi(a_0) = \varnothing$. Now, if $x \in X$ and $\pi(a)$ has been defined for all $a \in X$ with $\mathrm{rk}(a) < \mathrm{rk}(x)$, we set $\pi(x) = \big\{\pi(a) : a \in x \cap X\big\}$.

By induction on the rank, we see that π is injective and also that $\pi[X] = Y$ is a transitive set. Finally, π is an isomorphism; for if $x, y \in X$ and $x \in y$, then $x \in y \cap X$ and so

$$\pi(x) \in \big\{\pi(a) : a \in y \cap X\big\} = \pi(y);$$

and, conversely, if $x, y \in X$ and $\pi(x) \in \pi(y)$, then $\pi(x) = \pi(a)$ for some $a \in y \cap X$, so by injectivity, $x = a \in y$.

□

3. Relativization

Suppose C is a class. We define for every formula $\phi(\vec{x},\vec{a})$ in parameters $\vec{a}=(a_1,...a_n)$, a_i belonging to C, the relativized formula $\phi^C(\vec{x},\vec{a})$ by induction on the construction of ϕ:

- if ϕ is quantifier-free, then $\phi^C=\phi$
- $(\neg\phi)^C=\neg\phi^C,\ (\phi\vee\psi)^C=\phi^C\vee\psi^C$
- $(\exists y\phi)^C=\exists y\big(C(y)\big)\wedge\phi^C$
- $(\forall y\phi)^C=\forall y\big(C(y)\longrightarrow\phi^C\big)$

So ϕ^C simply relativizes all quantifications to C. So for any $\phi(\vec{x})$ without parameters and any parameters $\vec{a}$ in C, we have

$$(\mathcal{U},\in)\models\phi^C(\vec{a})\Longleftrightarrow(C,\in)\models\phi(\vec{a})$$

3.1. **Consistency of the Axiom of Foundation.**

Something that should be on our mind is the consistency of the axioms we have stated so far. So proofs of relative consistency are highly important. It follows from Gödel's Second Incompleteness Theorem that we cannot prove the consistency of ZFC from ZFC itself, so the only results we can hope for are relative consistency results, such as

Theorem 3.1. If ZF^- is consistent, then so is ZF.

In fact, we shall prove the following more precise statement:

Theorem (3.1*). Suppose $\mathcal{U}$ is a universe of sets satisfying the theory ZF^-. Then the class V constructed in $\mathcal{U}$ will be a universe of sets in which ZF holds.

Proof.

Recall that ZF^- is axiomatized by[a]

(i) the axiom of extensionality
(ii) the union axiom
(iii) the powerset axiom
(iv) the axiom scheme of replacement
(v) set existence
(vi) the axiom of infinity

We thus have to prove that if $\mathcal{U}$ is a universe of sets and V is the class of sets defined by $V=\bigcup_{\alpha\in\mathrm{Ord}}V_\alpha$, then $(V,\in)\models$ (i), ..., (vi).[b] I.e., for any axiom ϕ among (i)-(vi), ϕ^V holds.

(i): Suppose x,y belong to V and that for any z in V, $z\in x\longleftrightarrow z\in y$. Then as x,y are subsets of V, we have $\forall z\,(z\in x\longleftrightarrow z\in y)$, so since extensionality holds in $\mathcal{U}$, we have $x=y$ and thus $(x=y)^V$ too.

(ii): Suppose x belongs to V. Then $\bigcup x=\big\{z:\exists y\in x, z\in x\big\}$ is a subset of V and hence itself belongs to V.

[a] *TN:* And recall that pairing and comprehension follow from these as well.

[b] *TN:* This notation is slightly improper, as replacement actually has infinitely many axioms, but the idea is clear. Note that it's obviously not enough just to show that AF holds. We still want V to be a model of ZF^-!

(iii): If X belongs to V, then so does $\mathcal{P}(x)$; for any subset of x will be a subset of V and thus belong to V, hence $\mathcal{P}(x)$ is a subset of V and thus an element of V.

(iv): Suppose $\phi(x, y)$ is a formula with parameters in V that defines a class function in V, i.e.

$$\left(\forall x \exists^{\leq 1} y\ \phi(x, y)\right)^V$$

or more explicitly

$$\forall x \Big(V(x) \longrightarrow \exists^{\leq 1} y \left(V(y) \wedge \phi^V(x, y)\right)\Big).$$

Then the formula $\psi(x, y) := V(x) \wedge V(y) \wedge \phi^V(x, y)$ defines a functional relation in $\mathcal{U}$, and thus, for any set A, there is (by replacement in $\mathcal{U}$) a set B such that

$$y \in B \Longleftrightarrow \exists x \in A,\ \psi(x, y) \Longleftrightarrow \left(\exists x \in A,\ \phi^V(x, y)\right) \wedge V(y)$$

Since B is a subset of V, B actually belongs to V and is the image of A in V of the class function defined by ϕ.

(v): Trivial since $\varnothing$ belongs to V.

(vi): We only have to show that ω belongs to V, but actually we may show that all ordinals belong to V.[a] This is by induction.

First, $0 = \varnothing$ belongs to V.

Now suppose ξ belongs to V for all $\xi < \alpha$. Since $\alpha = \{\xi : \xi < \alpha\}$, α is a subset of V and hence belongs to V, finishing the induction.

Finally, to see that AF holds in V, suppose that a is a non-empty set belonging to V. Let $b \in a$ have minimal rank. Then since a is also a subset of V, b belongs to V too. Moreover, if $c \in b$, then $\operatorname{rk} c < \operatorname{rk} b$ and so $c \notin a$. It follows that $b \cap a = \varnothing$, so AF holds in V.

□

3.2. **Inaccessible Ordinals and Models of ZFC.**

Suppose ZFC holds in our universe $\mathcal{U}$. By induction on ordinals, we see that $\xi \subseteq V_\xi$ for any ordinal ξ.

Lemma 3.2. If κ is a strongly inaccessible cardinal, then $|V_\kappa| = \kappa$ and for any $a \subseteq V_\kappa$, we have $a \in V_\kappa \Longleftrightarrow |a| < \kappa$.

Proof.

Since $\kappa \subseteq V_\kappa$, notice $|V_\kappa| \geq \kappa$. Conversely, by induction we show that for any $\xi < \kappa$, $|V_\xi| < \kappa$, which implies that $|V_\kappa| = \kappa$. For the induction, note that if $|V_\xi| < \kappa$, then $|V_{\xi+1}| = |\mathcal{P}(V_\xi)| = 2^{|V_\xi|} < \kappa$. And if $|V_\xi| < \kappa$ for all $\xi < \lambda < \kappa$, λ limit, then $|V_\lambda| = |\bigcup_{\xi<\lambda} V_\xi| \leq \sup\limits_{\xi<\lambda} |V_\xi| < \kappa$ by regularity of κ.

Thus, if $a \subseteq V_\kappa$, $|a| < \kappa$, then $\operatorname{rk} : a \to \kappa$ cannot be cofinal in κ since κ is regular. So for some $\beta < \kappa$, $a \subseteq V_\beta$, hence $a \in V_{\beta+1} \subseteq V_\kappa$.

□

[a] *TN:* If you've already done the exercise to show that for any ordinal α, $\operatorname{rk} \alpha = \alpha + 1$, then you're already done, as this implies that for any ordinal α, $\alpha \in V_{\alpha+1}$ and hence belongs to V.

Lemma 3.3. If κ is inaccessible, then V_κ satisfies ZFC.

Proof.

One can check extensionality, union, powerset, set existence, axiom of infinity, and AF in the structure $(V_\kappa, \in)$. So, let us check AC and replacement.

If $a \in V_\kappa$ is a family of pairwise disjoint non-empty sets, then we know that there is some T which is a transversal for a, i.e., such that $\forall b \in a$, $T \cap b$ is a singleton, and $T \subseteq \bigcup a$. Since a is a subset of V_κ, so are all of its subsets, hence $T \subseteq V_\kappa$. Moreover, since $a \in V_\kappa$, $|T| = |a| < \kappa$, hence $T \in V_\kappa$. So AC holds in V_κ.

For replacement, suppose $\phi(x, y)$ is a formula with parameters in V_κ defining a class function in V_κ, i.e.

$$\forall x \in V_\kappa \, \exists^{\leq 1} y \in V_\kappa, \ \ \phi^{V_\kappa}(x, y),$$

and that $a \in V_\kappa$.

Then $\psi(x, y)$ given by $\psi(x, y) := x \in V_\kappa \wedge y \in V_\kappa \wedge \phi^{V_\kappa}(x, y)$ defines a function f with domain contained in V_κ. So $f[a]$ is a subset of V_κ of size $< \kappa$, hence $f[a] \in V_\kappa$, verifying replacement.

$\square$

Theorem 3.4. If ZFC is consistent, then so is ZFC + "there are no strongly inaccessible cardinals".

Remark:
We note that this is a statement of the metatheory, that is, not a statement of the first order language of set theory. We are claiming that if there is no way of obtaining a contradiction from ZFC, then the same holds for ZFC + "there are no strongly inaccessible cardinals".

Proof.

Suppose $\mathcal{U}$ is a universe of set theory satisfying ZFC. If there are no inaccessible cardinals in $\mathcal{U}$, we are done. So, suppose there are and let κ be the minimal of these. We will show that ("there are no strongly inaccessible cardinals")$^{V_\kappa}$ holds.
Note first that since AF holds, a set α is an ordinal if and only if α is transitive and totally ordered by $\in$, i.e.

$$(\star) \qquad \forall x, y \in \alpha \, (x \in y \vee y \in x \vee x = y) \wedge \forall x \, (x \in \alpha \longrightarrow x \subseteq \alpha)$$

Claim: The ordinals in V_κ are simply the ordinals below κ, i.e. $\mathrm{Ord}^{V_\kappa} = \kappa$.

Proof.

Note that $\kappa \subseteq V_\kappa$. So if $\alpha < \kappa$, then $\alpha \in V_\kappa$ and $(\star)$ holds for α. But then also $(\star)^{V_\kappa}$ holds, hence α is an an ordinal in V_κ, i.e., α belongs to Ord^{V_κ}.

Conversely, if α belongs to Ord^{V_κ}, then $\alpha \in V_\kappa$ and is transitive and totally ordered by $\in$, hence α is an ordinal, $|\alpha| < \kappa$, i.e. $\alpha \in \kappa$.

$\blacksquare$

Claim: The cardinals in V_κ are the cardinals below κ.

Proof.

If λ is a cardinal in V_κ, then it is an ordinal in V_κ and thus $\lambda \in \kappa$. To see that λ is also an actual cardinal, note that if $f : \lambda \longleftrightarrow \alpha$ were a bijection of λ with an ordinal $\alpha < \lambda$, then $f \subseteq V_\kappa$, $|f| < \kappa$, and thus $f \in V_\kappa$. It follows that f would be a bijection in V_κ between λ and a smaller ordinal, contradicting that λ is a cardinal in V_κ.

And, similarly, if $\lambda < \kappa$ is a cardinal, then $\lambda \in V_\kappa$ and λ is an ordinal in V_κ. Moreover, any bijection in V_κ between λ and a smaller ordinal would also be a bijection in $\mathcal{U}$ between λ and a smaller ordinal, contradicting that λ is a cardinal.

■

Now that we've established that the cardinals in V_κ are exactly the cardinals in $\mathcal{U}$ that are less than κ, we finally prove that no cardinal $\lambda \in V_\kappa$ is strongly inaccessible in V_κ.

If $\lambda \in V_\kappa$ is a cardinal in V_κ, then $\lambda < \kappa$ is also a cardinal in $\mathcal{U}$, but not strongly inaccessible. So either

(i) $\lambda \leq \omega$, hence also $(\lambda \leq \omega)^{V_\kappa}$,
(ii) $\lambda \leq 2^\xi$ for some cardinal $\xi < \lambda < \kappa$, hence $\xi, 2^\xi \in V_\kappa$ and also $(\xi < \lambda \leq 2^\xi)^{V_\kappa}$, or
(iii) there is a cofinal function $f : \alpha \to \lambda$ from an ordinal $\alpha < \lambda$, hence again $\alpha, f \in V_\kappa$ and thus λ is singular in V_κ.

So no cardinal in V_κ is inaccessible.

□

3.3. **The Reflection Scheme.**

Definition 3.5. Suppose C is a class and $\phi(\vec{x})$ is a formula all of whose parameters belong to C. We say that $\phi(\vec{x})$ is <u>absolute</u> for C if for all $\vec{a}$ in C, $\phi(\vec{a}) \iff \phi^C(\vec{a})$. I.e., if and only if

$$\forall \vec{x} \left(C(\vec{x}) \longrightarrow \left(\phi(\vec{x}) \longleftrightarrow \phi^C(\vec{x})\right)\right).$$

Since the relativization ϕ^C of a quantifier-free formula is ϕ itself, any quantifier-free ϕ is absolute for C.

Definition 3.6. A formula $\phi(\vec{x})$ is said to be in <u>prenex-form</u> if $\phi = Q_1 y_1 Q_2 y_2 ... Q_n y_n \psi$ where each Q_i is a quantifier and ψ is quantifier free.

<u>Observation</u>:

The class of formulas absolute for C is closed under logical equivalence and Boolean combinations. That is, if $\vdash \psi(\vec{x}) \longleftrightarrow \phi(\vec{x})$, then ψ is absolute for C if and only if ϕ is absolute for C. This follows from

$$\vdash (\psi \longrightarrow \phi) \implies \vdash (\psi^C \longrightarrow \phi^C),$$

which can be proven by induction on proofs or by model-theoretic considerations.
Since every formula is logically equivalent to one in prenex-form, when dealing with absoluteness, it suffices to consider formulas in prenex-form.

Lemma 3.7. Suppose $\phi(\vec{x})$ is a formula without parameters in prenex-form and that $(X_u)_{u\in\omega}$ is an increasing sequence of sets. If ϕ and all its subformulas are absolute for every X_n, then ϕ and all of its subformulas are absolute for $X = \bigcup_{n\in\omega} X_n$.

Proof.

The result is proved by induction on the length of the prefix of ϕ.

If ϕ is quantifier-free, then ϕ is absolute for any class or set, so the result is trivial.

Suppose now that the result is true for $\psi(y, x_1, ..., x_n)$ and let $\phi(\vec{x}) = \exists y\, \psi(y, \vec{x})$. Then for any $\vec{c} \in X$, choose $k < \omega$ such that $\vec{c} \in X_k$. Now, if $\phi(\vec{c})$ holds, then since ϕ is absolute for X_k, also $\phi^{X_k}(\vec{c})$ holds, so for some $b \in X_k$, $\psi^{X_k}(b, \vec{c})$ holds. As ψ is absolute for X_k, we get that $\psi(b, \vec{c})$, and as ψ is absolute for X, we get that $\psi^X(b, \vec{c})$. Thus, finally, $\exists b \in X\ \psi^X(b, \vec{c})$, i.e. $\phi^X(\vec{c})$.

Conversely, if $\phi^X(\vec{c})$, then for some $b \in X$, $\psi^X(b, \vec{c})$. Since ψ is absolute for X, also $\psi(b, \vec{c})$ and so $\exists y \psi(y, \vec{c})$, i.e. $\phi(\vec{c})$.

Universal quantification is proved similarly.

□

Theorem 3.8 (The Reflection Scheme (ZF))**.** Suppose $\phi(\vec{x})$ is a formula without parameters. Then for every α, there is a limit ordinal $\beta > \alpha$ such that ϕ is absolute for V_β.

Proof.

Without loss of generality, we can suppose that ϕ is in prenex-form. We show by induction on the length of the quantifier-prefix of ϕ that:

$$\forall \alpha\ \exists \beta > \alpha \text{ a limit (every subformula of } \phi \text{ is absolute for } V_\beta).$$

The base case when ϕ is quantifier-free is trivial, since ϕ is absolute for $V_{\alpha+\omega}$.

Now, suppose that the induction hypothesis holds for $\psi(y, \vec{x})$ and let $\phi(\vec{x}) = \exists y \psi(y, \vec{x})$. Then, by the induction hypothesis, for any α there is $\beta > \alpha$ limit such that ψ and all its subformulas are absolute for V_β. Fix α.

We define a class function $F(\vec{x}) = z$ by "$z = F(\vec{x})$ is the set of all y of minimal rank such that $\psi(y, \vec{x})$." Thus, $\vec{x}$ belongs to the domain of F if and only if $\exists y\ \big(\psi(y, \vec{x}) \wedge y \in F(\vec{x})\big)$.

We now define a strictly increasing sequence of odrinals $(\beta_n)_{n\in\omega}$ as follows:

- $\beta_0 = \alpha$.
- β_{2n+1} = smallest ordinal $> \beta_{2n}$ such that $F(\vec{c}) \in V_{\beta_{2n+1}}$ for every tuple $\vec{c} = (c_1, ..., c_k)$ in the domain of F with $c_1, ..., c_k \in V_{\beta_{2n}}$.
- β_{2n+2} = smallest ordinal $> \beta_{2n+1}$ such that ψ and all its subformulas are absolute for $V_{\beta_{2n+2}}$.

Now set $\beta = \sup\limits_{n<\omega} \beta_n$, which is a limit ordinal $> \alpha$. Also, since $V_\beta = \bigcup_{n<\omega} V_{\beta_{2n+2}}$, the previous lemma implies that ψ and all its subformulas are absolute for V_β. To finish the proof of the induction step, it thus suffices to prove that also ϕ is absolute for V_β.

We fix $c_1, ..., c_k \in V_\beta$, say $c_1, ..., c_k \in V_{\beta_{2u+1}}$. First, if $\phi^{V_\beta}(\vec{c})$, then there is $b \in V_\beta$ such that $\psi^{V_\beta}(b, \vec{c})$. Since ψ is absolute for V_β, also $\psi(b, \vec{c})$, hence $\exists y \psi(y, \vec{c})$, i.e. $\phi(\vec{c})$.

Conversely, if $\phi(\vec{c})$, then there is some b of minimal rank such that $\psi(b, \vec{c})$, hence $b \in F(\vec{c})$. It follows that $F(\vec{c}) \in V_{\beta_{2u+2}}$, and so also $b \in F(\vec{c}) \subseteq V_{\beta_{2u+2}} \subseteq V_\beta$. Thus, as ψ is absolute for V_β, we have $\psi^{V_\beta}(b, \vec{c})$ and hence $\exists y \in V_\beta$, $\psi^{V_\beta}(y, \vec{c})$, i.e. $\phi^{V_\beta}(\vec{c})$.

The case of universal quantifiers is similar. Alternatively, by using $\forall = \neg\exists\neg$, one can reduce it to existential quantifiers.

□

Corollary 3.9. For any true sentence σ without parameters, there are arbitrarily large limit ordinals β such that σ^{V_β} holds.

Using the preceding arguments, one can prove a more general statement:

Theorem 3.10 (ZF$^-$)**.** Suppose $W : \mathrm{Ord} \to \mathcal{U}$ is a class function such that

- $\alpha < \beta \implies W_\alpha \subseteq W_\beta$ (increasing)
- λ limit $\implies W_\lambda = \bigcup_{\beta<\lambda} W_\beta$.

Let W be the class $\bigcup_{\beta\in\mathrm{Ord}} W_\beta$. Then for any formula $\phi(\vec{x})$ without parameters and any ordinal α, there is a limit ordinal $\beta > \alpha$ such that

$$\forall x_1, ..., x_n \in W_\beta \; (\phi^{W_\beta}(\vec{x}) \longleftrightarrow \phi^W(\vec{x})).$$

4. Formalizing Logic in $\mathcal{U}$

Our universe of sets $\mathcal{U}$ should be a place for all mathematics to be done. That is, all groups, manifolds, function spaces, etc. can be constructed as elements of $\mathcal{U}$ and all reasoning about these objects should ultimately hark back to an underlying reasoning based on ZFC. Thus, in many ways, the set theoretical language is our machine language, while concepts such as fiber-bundles, C^∞-maps, solution spaces of partial differential equations are special kinds of sets defined by more or less involved definitions upon definitions.

As all other mathematical topics, logic also admits a formalization in $\mathcal{U}$, in such a way that formulas, proofs, and models simply are objects within $\mathcal{U}$. We shall give a cursory treatment of this.

Definition 4.1. Let $\dot\vee$, $\dot\neg$, $\dot\exists$, $\dot\in$, $\dot=$ be distinct sets in $\mathcal{U}$, e.g. $0, 1, 2, 3, 4$, and let $\mathcal{V}$ be a disjoint countable set, say $\mathcal{V} = \big\{n < \omega : n \geq 5\big\}$, called the set of $\underline{\mathcal{U}\text{-variables}}$.[a]

By induction on $n < \omega$, define a function $n \longmapsto \mathcal{F}_n$ with domain ω by:

- $\mathcal{F}_0 = \big\{(\dot\in, x, y),\ (\dot=, x, y) \mid x, y \in \mathcal{V}\big\}$
- $\mathcal{F}_{n+1} = \mathcal{F}_n \cup \big\{(\dot\neg, f),\ (\dot\vee, f, g),\ (\dot\exists, x, f) \mid f, g \in \mathcal{F}_n,\ x \in \mathcal{V}\big\}$.

Finally, $\mathcal{F} = \bigcup\limits_{n<\omega} \mathcal{F}_n$. Elements of $\mathcal{F}_0$ are called $\underline{\text{atomic } \mathcal{U}\text{-formulas}}$, while the elements of $\mathcal{F}$ are simply $\underline{\mathcal{U}\text{-formulas}}$. For $f \in \mathcal{F}$, $l(f) = \text{length}(f) = \text{minimal } n < \omega$ such that $f \in \mathcal{F}_n$.

Lemma 4.2 (Unique Readability[b])**.** For any $\mathcal{U}$-formula $f \in \mathcal{F}$, exactly one of the following holds:

(i) f is an atomic $\mathcal{U}$-formula
(ii) $f = (\dot\neg, g)$ for some unique $g \in \mathcal{F}$
(iii) $f = (\dot\vee, g, h)$ for some unique $g, h \in \mathcal{F}$
(iv) $f = (\dot\exists, x, g)$ for some unique $g \in \mathcal{F}$, $x \in \mathcal{V}$.

Moreover, in each of these cases, $l(g), l(h) < l(f)$.

<u>Notation</u>:
For simplicity of notation, we shall write

$$(x \mathbin{\dot\in} y),\ (x \mathbin{\dot=} y),\ (\dot\neg f),\ (f \mathbin{\dot\vee} g),\ \dot\exists x(f)$$

for the $\mathcal{U}$-formulas

$$(\dot\in, x, y),\ (\dot=, x, y),\ (\dot\neg, f),\ (\dot\vee, f, g),\ (\dot\exists, x, f).$$

Similarly, the $\mathcal{U}$-formulas

$$\big((\dot\neg f) \mathbin{\dot\vee} g\big),\ \Big(\dot\neg\ \big((\dot\neg f) \mathbin{\dot\vee} (\dot\neg g)\big)\Big),\ \big(\dot\neg\ \dot\exists x(\dot\neg f)\big)$$

are written

$$(f \mathbin{\dot\longrightarrow} g),\ (f \mathbin{\dot\wedge} g),\ \big(\dot\forall x(f)\big)$$

[a] *TN:* In class, we struggled to come up with a way to represent the same logical symbols we know and love, while making sure that there is little confusion about which are meta-logical symbols, and which are logical symbols that are really sets in $\mathcal{U}$. We settled on the convention of dotting the logical symbols inside our universe, language, etc.

[b] *TN:* Professor Tserunyan finds the inclusion of such a lemma to be useless. The truth of this lemma ought to be painfully obvious from the construction.

By induction on $l(f)$, we define for any $f \in \mathcal{F}$ the set $var(f)$ of free variables in f by

- if f is $(x \mathrel{\dot\in} y)$ or $(x \mathrel{\dot=} y)$, then $var(f) = \{x, y\}$
- $var(\dot\neg f) = var(f)$
- $var(f \mathrel{\dot\vee} g) = var(f) \cup var(g)$
- $var(\dot\exists x(f)) = var(f) \setminus \{x\}$.

Also $f \in \mathcal{F}$ is said to be a $\mathcal{U}$-sentence if $var(f) = \varnothing$.

Note:

For any formula $\phi(\vec{x})$ of set theory, there is a corresponding $\mathcal{U}$-formula f which we will denote by $\ulcorner\phi\urcorner$. Thus, while ϕ is an object of our metalanguage, $\ulcorner\phi\urcorner$ is a set belonging to our universe $\mathcal{U}$.

So, for example, it makes sense to quantify over $\mathcal{U}$-formulas in the language of set theory, which is not the case for true formulas of the metalanguage.

Also, note that if our universe $\mathcal{U}$ contains non-standard natural numbers, then there may be non-standard $\mathcal{U}$-formulas, i.e. $\mathcal{U}$-formulas f not of the form $\ulcorner\phi\urcorner$ for some ϕ of the language of set theory.

4.1. **Model Theory for $\mathcal{U}$-formulas.** By induction on the length of $f \in \mathcal{F}$, we define for every non-empty set X, a set $Val(f, X)$ by:

(i) $Val\big((x \mathrel{\dot\in} y),\, X\big) = \big\{\delta \in X^{\{x,y\}} \,\big|\, \delta(x) \in \delta(y)\big\}$[a]
(ii) $Val\big((x \mathrel{\dot=} y),\, X\big) = \big\{\delta \in X^{\{x,y\}} \,\big|\, \delta(x) = \delta(y)\big\}$
(iii) $Val\big((\dot\neg f),\, X\big) = X^{var(f)} \setminus Val(f, X)$
(iv) $Val\big((f \mathrel{\dot\vee} g),\, X\big) = \big\{\delta \in X^{var(f \dot\vee g)} \,\big|\, \delta\restriction_{var(f)} \in Val(f,\, X) \text{ or } \delta\restriction_{var(g)} \in Val(g,\, X)\big\}$
(v) $Val\big(\dot\exists x(f),\, X\big) = \big\{\delta \in X^{var(f) \setminus \{x\}} \,\big|\, \exists \tilde\delta \in Val(f,\, X),\ \tilde\delta\restriction_{var(f)\setminus\{x\}} = \delta\big\}$.

Note:

For any formula $\phi(x_1, ..., x_n)$ of our metalanguage, we have (modulo changing variables)

$$Val(\ulcorner\phi\urcorner,\, X) = \big\{\delta : \{x_1, ..., x_n\} \to X \,\big|\, \phi^X(\delta x_1, ..., \delta x_n) \text{ holds}\big\}.$$

So we can see $Val(\ulcorner\phi\urcorner, X)$ as the set

$$\big\{(a_1, ..., a_n) \in X^n \mid \phi^X(a_1, ...a_n) \text{ holds}\big\}.$$

Suppose f is a $\mathcal{U}$-formula with free variables among $x_1, ..., x_n$, written $f(x_1, ..., x_n)$. Assume also that γ is a function from a subset of $var(f)$ into a set X. Then we say that (f, γ) is a $\mathcal{U}$-formula with parameters in X.

For simplicity of notation, if $f(x_1, ..., x_n, y_1, ..., y_k)$ is given with $x_1, ..., x_n \in var(f)$ and $\gamma : \{x_1, ..., x_n\} \to X$ with $\gamma(x_i) = a_i$, we write $f(a_1, ..., a_n, y_1, ..., y_k)$ or just $f(\vec{a}, \vec{y})$ for (f, γ). In this case, $var(f(\vec{a}, \vec{y})) = var(f) \setminus \{x_1, ..., x_n\}$.

A $\mathcal{U}$-formula f (possibly with parameters) is said to be a $\mathcal{U}$-sentence if $var(f) = \varnothing$. Also,

$$Val\big((f, \gamma),\, X\big) = \big\{\delta \in X^{var(f,\gamma)} \mid \delta \cup \gamma \in Val(f, X)\big\}$$

If f is a $\mathcal{U}$-sentence whose parameters belong to a set X, then $Val(f, X)$ is a subset of $X^{\varnothing} = \{\varnothing\}$. If $Val(f, X) = \{\varnothing\} = 1$, we say that f is true in X, written $X \models f$. Similarly, if $Val(f, X) = \varnothing = 0$,

[a] *TN:* The notation $X^{\{x,y\}}$ hearkens back to when we defined exponentiation for cardinals. This just means "the set of functions from $\{x, y\}$ to X".

we say that f is <u>false</u> in X.

Theorem 4.3 (Löwenheim-Skolem (AC))**.** Suppose $P \subseteq X$ are sets. Then there is a subset $Y \subseteq X$ containing P, $|Y| \leq |P| \oplus \aleph_0$, such that for any $\mathcal{U}$-sentence f with parameters in Y,

$$X \models f \qquad \Longleftrightarrow \qquad Y \models f.$$

Proof.

Fix a choice function $\pi : \mathcal{P}(X) \smallsetminus \{\varnothing\} \to X$, i.e. such that $\pi(A) \in A$ for any non-empty subset A of X. We define inductively an increasing sequence $(P_n)_{n<\omega}$ of subsets of X as follows:

- $P_0 = P$
- Given P_n, let $R_n = \big\{g(\vec{a}, x) \,|\, g(\vec{a}, x)$ is a $\mathcal{U}$-formula with parameters $\vec{a}$ in X and $X \models \dot{\exists}\, x\, g(\vec{a}, x)\big\}$
- For any $g(\vec{a}, x) \in R_n$, let $b_{g(\vec{a},x)} = \pi\big(\{b \in X \,\big|\, X \models g(\vec{a}, b)\}\big)$
- Define $P_{n+1} := P_n \cup \{b_{g(\vec{a},x)} \,|\, g(\vec{a}, x) \in R_n\}$

Since there are only countably many $\mathcal{U}$-formulas, $|P_{n+1}| \leq |P_n| \oplus \aleph_0$, so by induction, $|P_n| \leq |P| \oplus \aleph_0$.

Set $Y := \bigcup_{n<w} P_n$. We show by induction on the length of a formula that if f is a $\mathcal{U}$-sentence with parameters in Y, then $X \models f \iff Y \models f$. This is trivial if f is atomic, and the induction steps for $\dot{\neg}$ and $\dot{\vee}$ are easy.

So suppose instead that $f = f(\vec{a}) = \dot{\exists}\, x\, g(\vec{a}, x)$ where the induction hypothesis holds for g. If $Y \models f(\vec{a})$, then there is $b \in Y$ such that $Y \models g(\vec{a}, b)$, hence also $X \models g(\vec{a}, b)$ and so $X \models f(\vec{a})$.

Conversely, if $X \models f(\vec{a})$, find n large enough such that the parameters $\vec{a} = (a_1, ..., a_k)$ all belong to P_n. It follows that $g(\vec{a}, x) \in R_n$, and so $b_{g(\vec{a},x)} \in P_{n+1} \subseteq Y$. Since $X \models g(\vec{a}, b_{g(\vec{a},x)})$, we have by the induction hypothesis that $Y \models g(\vec{a}, b_{g(\vec{a},x)})$, and so $Y \models f(\vec{a})$.

$\square$

Relativization:

As for formulas of set theory, when given a $\mathcal{U}$-formula $f(\vec{a}, \vec{x})$ with parameters in a set X, we can define the relativized formula $f^X(\vec{a}, \vec{x})$ by induction on the length of f. We then have the following easy fact:

Theorem 4.4. Suppose $X \subseteq Y$ are sets with $X \in Y$ and $f(\vec{a}, \vec{x})$ is a $\mathcal{U}$-formula with parameters in X. Then

$$Val(f(\vec{a}, \vec{x}),\, X) = Val(f^X(\vec{a}, \vec{x}),\, Y) \cap X^{var(f(\vec{a},\vec{x}))}.$$

5. Ordinal Definability and Inner Models of ZFC

To simplify notation, if $f(\vec{a}, x)$ is a $\mathcal{U}$-formula with one free variable x and parameters in a set X, $Val(f(\vec{a}, x), X)$ is a subset of $X^{\{x\}}$, which we canonically can identify with a subset of X.

Definition 5.1 (ZF)**.** Let *OD* be the class of ordinal definable sets given by:

$$OD(a) \iff \begin{cases} \text{there are ordinals } \alpha_1, ..., \alpha_k,\ k < \omega \text{ and } \beta \text{ and a } \mathcal{U}\text{-formula} \\ f(x, \alpha_1, ..., \alpha_k) \text{ with one free variable such that } \alpha_1, ..., \alpha_k < \beta, \\ a \in V_\beta \text{ and } Val(f(x, \alpha_1, ..., \alpha_k),\, V_\beta) = \{a\}. \end{cases}$$

Observation:

Note that every ordinal is ordinal definable by using itself as a parameter.

Proposition 5.2. Suppose $\phi(x, \alpha_1, ..., \alpha_n)$ is a formula of the language of set theory with ordinal parameters $\alpha_1, ..., \alpha_n$ and suppose that a is the unique set satisfying $\phi(x, \alpha_1, ..., \alpha_n)$. Then a is ordinal definable, i.e. belongs to *OD*.

Proof.

By the reflection scheme, we can find an ordinal β satisfying

- $\alpha_1, ..., \alpha_n < \beta$
- $a \in V_\beta$
- $\phi(x, y_1, ..., y_n)$ is absolute for V_β.

In particular, a is the unique element of V_β satisfying $\phi^{V_\beta}(x, \alpha_1, ..., \alpha_n)$ and hence

$$Val(\ulcorner \phi(x, \alpha_1, ..., \alpha_n) \urcorner,\, V_\beta) = \{a\}.$$

So a is ordinal definable.

□

Proposition 5.3. There is a formula $\psi(x, y)$ of the language of set theory such that for any set a,

$$OD(a) \begin{array}{l} \iff \exists \gamma \text{ ordinal } \forall x \left(\psi(x, \gamma) \longleftrightarrow x = a\right) \\ \iff \exists \gamma \text{ ordinal } \psi(a, \gamma). \end{array}$$

Thus, this formula ψ provides a uniform characterization of ordinal definability.

Proof.

Let $\mathcal{S}$ be the class of all ordinal valued functions $s : n \to \mathrm{Ord}$, with $\mathrm{dom}(s) = n$ a finite ordinal. That is, $\mathcal{S}$ is the class of finite sequences of ordinals. For s, t in $\mathcal{S}$, we put

$$s <' t \iff \begin{cases} \sup(s) < \sup(t) \text{ or} \\ \sup(s) = \sup(t) \wedge \mathrm{dom}(s) < \mathrm{dom}(t) \text{ or} \\ \sup(s) = \sup(t) \wedge \mathrm{dom}(s) = \mathrm{dom}(t) \wedge s <_{lex} t. \end{cases}$$

Then one can check that $<'$ defines a well-ordering of $\mathcal{S}$ whose proper initial segments are sets. It follows that there is an order preserving class function $\text{ꙏ} : \mathrm{Ord} \to \mathcal{S}$.[a]

[a] *TN:* The character used for this function is a backwards cyrillic Б. The handwritten notes were a bit hard to distinguish here, leading some in the class to speculate as to which character was used, and this was the funniest answer, and was kept.

Also, in ZF we can construct a bijection $\mathcal{K} : \omega \to V_\omega$. So, in particular, since any $\mathcal{U}$-formula is an object in V_ω, $\mathcal{K}$ maps onto the set of $\mathcal{U}$-formulas $\mathcal{F}$. We can now define $\psi(x,y)$ as follows:

$$\psi(x,y) \Longleftrightarrow \begin{cases} y \text{ is an ordinal and } \text{Ɑ}(y) = (n, \beta, \alpha_1, ..., \alpha_k) \text{ is a finite} \\ \text{sequence of ordinals such that } \alpha_1, ..., \alpha_k < \beta,\ n < \omega \\ x \in V_\beta,\ \text{ and } \mathcal{K}(n) = f(z, y_1, ..., y_k) \text{ is a } \mathcal{U}\text{-formula with} \\ Val(f(z, \alpha_1, ..., \alpha_k), V_\beta) = \{x\}. \end{cases}$$

Then, we have

$$OD(a) \begin{array}{cc} \Longleftrightarrow & \exists\gamma \text{ ordinal } \forall x \left(\psi(x,\gamma) \longleftrightarrow x = a\right) \\ \Longleftrightarrow & \exists\gamma \text{ ordinal } \psi(a,\gamma). \end{array}$$

□

Definition 5.4. The class HOD of hereditarily ordinal definable sets is given by

$$HOD(a) \Longleftrightarrow OD(a) \wedge \mathrm{cl}(a) \text{ is a subset of } OD.$$

Lemma 5.5. $HOD(a) \Longleftrightarrow OD(a) \wedge \forall x \in a,\ HOD(x)$.

This follows from the fact that

$$\mathrm{cl}(a) = a \cup \bigcup_{x \in a} \mathrm{cl}(x).$$

Theorem 5.6. Suppose $\mathcal{U}$ is a model of ZF. Then HOD is a model of ZFC.[a,b]

Proof.

Note first that HOD is a transitive class (i.e., $a \in b \wedge HOD(b) \implies HOD(a)$) containing all ordinals.

Extensionality: Follows from transitivity of HOD plus extensionality in $\mathcal{U}$. ■

Union: Note that if $HOD(a)$ and $b = \bigcup_{x \in a} x$, then b is a subset of HOD. So to see that b belongs to HOD, we only need to see that b is ordinal definable. So pick α such that $\psi(a,\alpha)$ holds. Then b is the unique object x satisfying

$$\phi(x,a) := \forall y \left(y \in x \longleftrightarrow \exists z \left(z \in a \wedge y \in z\right)\right).$$

Hence, the formula $\exists v \left(\psi(v,\alpha) \wedge \phi(x,v)\right)$ defines b. Thus, $OD(b)$ and $HOD(b)$. ■

Powerset: Suppose $HOD(a)$ and let $b = \mathcal{P}(a) \cap HOD$ be the set of all hereditarily ordinal definable subsets of a. Then b is a subset of HOD and can be seen to be ordinal definable by methods as above. Thus $b = \mathcal{P}^{HOD}(a)$. ■

Replacement: Suppose $\sigma(x, y, a_1, ..., a_k)$ is a formula with parameters $a_1, ..., a_k$ from HOD that defines a class function in HOD, i.e.

$$\forall x \left(HOD(x) \longrightarrow \exists^{\leq 1} y \left(HOD(y) \wedge \sigma^{HOD}(x, y, \vec{a})\right)\right).$$

[a] *TN:* In other words, if ZF is consistent, then so is ZFC.

[b] *TN:* HOD is said to be an inner model of set theory. That is, a substructure of some model of a theory T that also contains all the ordinals of the original model.

Suppose X is in HOD and Y is the set of images in HOD of elements of X by this class function. Notice, Y is a subset of HOD, so we need only show that Y is in OD.

So fix ordinals $\beta, \alpha_1, ..., \alpha_k$ such that $\psi(X, \beta), \psi(a_1, \alpha_1), ..., \psi(a_k, \alpha_k)$. Then Y is the unique object satisfying

$$\psi(x, X, a_1, ..., a_k) := \forall z \Big(z \in x \longleftrightarrow \exists u \left(u \in X \wedge HOD(z) \wedge \sigma^{HOD}(u, z, \vec{a}) \right) \Big).$$

Hence, Y is defined by

$$\exists v \exists y_1 ... \exists y_k \left(\psi(v, \beta) \wedge \psi(y, \alpha_1 \wedge ... \wedge \psi(y_k, \alpha_k) \wedge \phi(x, v, y_1, ..., y_k) \right).$$

$\blacksquare$

Infinity: Since $HOD(\omega)$, HOD satisfies the axiom of infinity. $\blacksquare$

Foundation: If $HOD(a)$ and $a \neq \varnothing$, let $b \in a$ have minimal rank. Then $b \cap a = \varnothing$ and $HOD(b)$. $\blacksquare$

Choice: Note first that we have a well-ordering of HOD by

$$a <' b \Longleftrightarrow \exists \alpha \left(\psi(a, \alpha) \wedge \forall \beta \left(\psi(b, \beta) \longrightarrow \alpha < \beta \right) \right).$$

That is, we well-order elements of HOD according to the minimal ordinal defining them via ψ. So, if $X \neq \varnothing$ belongs to HOD, then

$$R = \left\{ (a, b) \in X^2 \,|\, a <' b \right\}$$

is a well-ordering of X, and thus a subset of HOD, that is ordinal definable. So R belongs to HOD and hence X can be well-ordered in HOD.

$\square$

5.1. **The Principle of Choice.**

Recall that we have a formula without parameters in the language of set theory, $\psi(x, y)$, such that for any set a:

$$OD(a) \begin{array}{ll} \Longleftrightarrow & \exists \gamma \text{ ordinal } \forall x \left(\psi(x, \gamma) \longleftrightarrow x = a \right) \\ \Longleftrightarrow & \exists \gamma \text{ ordinal } \psi(a, \gamma). \end{array}$$

Using ψ, we can define a well-ordering of the class OD by

$$a \prec b \Longleftrightarrow OD(a) \wedge OD(b) \wedge \exists \alpha \left(\psi(a, \alpha) \wedge \forall \beta \left(\psi(b, \beta) \longrightarrow \alpha < \beta \right) \right).$$

Definition 5.7. The principle of choice is the statement that there is a first order formula $\phi(x, y)$ without parameters defining a well-ordering of $\mathcal{U}$.

Note:
Since this is really a disjunction over all formulas of set theory, the principle of choice is not even an axiom scheme and even less a first order axiom. However, on the basis of ZF, we shall see that it is first order.

Proposition 5.8. $\mathcal{U}$ satisfies the principle of choice if and only if $\forall x\, OD(x)$ holds.

Proof.

Note that in $\mathcal{U}$, $\prec$ defines a well-ordering of OD. Thus, if "$\forall x\, OD(x)$", i.e. $\mathcal{U} = OD$, then $\prec$ is a well-ordering of $\mathcal{U}$.

Conversely, suppose $\phi(x, y)$ is a formula without parameters defining a well-ordering of $\mathcal{U}$. Then there is a unique class function $F : \mathrm{Ord} \to \mathcal{U}$ such that $\alpha < \beta \iff \phi(F(\alpha), F(\beta))$. It follows that for every α, $F(\alpha)$ is ordinal definable with parameters α, and hence, $\forall x\, OD(x)$.

□

Thus: principle of choice $\iff \mathcal{U} = OD \iff \mathcal{U} = HOD$.

6. Constructibility d'apres K. Gödel

Definition 6.1 (ZF)**.** Suppose A is a set and $X \subseteq A$ is a subset. We say that X is definable with parameters in A if

$$\exists k < \omega, \exists f(x, y_1, ..., y_k) \text{ a } \mathcal{U}\text{-formula}, \exists a_1, ..., a_k \in A,\ X = Val\big(f(x, a_1, ..., a_k), A\big).$$

Again, this is a first order property in X and A, so we can define the set

$$\mathcal{D}(A) = \big\{X \subseteq A \,|\, X \text{ is definable with parameters in } A\big\}.$$

Example:

Suppose $\phi(x, y_1, ..., y_n)$ is a formula in the language of set theory and $a_1, ..., a_n \in A$ are such that $X = \{x \in A \,|\, A \models \phi(x, \vec{a})\}$. Then

$$X = Val\big(\ulcorner\phi(x, \vec{a})\urcorner, A\big) \in \mathcal{D}(A).$$

Remark: (ZFC)

Suppose $|A| \geq \aleph_0$. Then $|\mathcal{D}(A)| = |A|$.

To see this, note that the set of $\mathcal{U}$-formulas with parameters in A has size $|A|$ itself, and thus also $|\mathcal{D}(A)| = |A|$. In particular, for infinite A, $|\mathcal{D}(A)| < |\mathcal{P}(A)|$.

Also, $A \in \mathcal{D}(A)$ for any A. But if $A \subseteq B$ is not definable, then $A \notin \mathcal{D}(B)$ and so $\mathcal{D}(A) \not\subseteq \mathcal{D}(B)$. Nevertheless, we have the following theorem.

Theorem 6.2. If $A \subseteq B$ and $A \in B$, then $\mathcal{D}(A) \subseteq \mathcal{D}(B)$.

Proof.

Suppose $X \in \mathcal{D}(A)$ and that $f(x, a_1, ..., a_k)$ is a $\mathcal{U}$-formula with parameters in A such that $X = Val\big(f(x, \vec{a}), A\big)$. Then also

$$X = Val\big(f^A(x, \vec{a}), B\big) \cap A = Val\big(f^A(x, \vec{a}) \wedge x \in A, B\big)$$

so $X \in \mathcal{D}(B)$.

□

By transfinite induction, we now define a hierarchy of sets by

$$L_\alpha := \bigcup_{\beta<\alpha} \mathcal{D}(L_\beta).$$

So $L_0 = \varnothing$ and $L_\beta \subseteq L_\alpha$ for $\beta < \alpha$. Also, for $\beta < \alpha$, $L_\beta \in \mathcal{D}(L_\beta) \subseteq L_\alpha$, and hence, by the preceding theorem, $\mathcal{D}(L_\beta) \subseteq \mathcal{D}(L_\alpha)$. It follows that the hierarchy can alternatively be described by

$$L_0 = \varnothing\,, \qquad L_{\alpha+1} = \mathcal{D}(L_\alpha)\,, \qquad L_\lambda = \bigcup_{\xi<\lambda} L_\xi \text{ for } \lambda \text{ limit.}$$

Moreover, since $\mathcal{D}(A) \subseteq \mathcal{P}(A)$ for every set A, we see by induction on α that $L_\alpha \subseteq V_\alpha$.
Let L be the class of constructible sets defined by $L := \bigcup_{\alpha\in\mathrm{Ord}} L_\alpha$. So $L \subseteq V$.

Definition 6.3. The axiom of constructibility is the statement $\mathcal{U} = L$, which, assuming AF, is just $V = L$, i.e. $\forall x\, \exists\alpha,\, x \in L_\alpha$.

Lemma 6.4. L is a transitive class, that is, if A is constructible, then so is every element of A. Also, if $A \in L_\alpha$, then $A \subseteq L_\beta$ for some $\beta < \alpha$.

Proof.

Note that if $A \in L_{\xi+1} = \mathcal{D}(L_\xi)$, then $A \subseteq L_\xi$.

□

By the proof, we see that any L_α is a transitive set.

Definition 6.5. For any constructible set X, we let order$(X) = \min(\alpha \,|\, X \in L_\alpha)$.

Theorem 6.6. Ord $\subseteq L$ and Ord$\cap L_\alpha = \alpha$, $\forall \alpha$. So any ordinal α is constructible with order $\alpha + 1$.

Proof.

By induction on α, we prove Ord $\cap L_\alpha = \alpha$. So suppose this holds for all α less than some ordinal β.
If β is a limit, then

$$\text{Ord} \cap L_\beta = \bigcup_{\alpha<\beta} \text{Ord} \cap L_\alpha = \bigcup_{\alpha<\beta} \alpha = \sup_{\alpha<\beta} \alpha = \beta.$$

On the other hand, if $\beta = \alpha + 1$ for some α, then, by assumption, Ord $\cap L_\alpha = \alpha$ and so $\alpha \subseteq L_\alpha$. On the other hand, since $\alpha \notin L_\alpha$, $\gamma \nsubseteq L_\alpha$ for any $\gamma \geq \beta = \alpha \cup \{\alpha\}$. Thus, to see that Ord $\cap L_\beta = \beta$, we need only to show that α is a definable subset of L_α, i.e. $\alpha \in \mathcal{D}(L_\alpha)$.
For this, consider the formula

$$\begin{aligned}\phi(x) := \forall u \forall v \Big((u \in x \wedge v \in x) \longrightarrow \big((u \in v) \vee (v \in u) \vee (u = v) \big) \Big) \\ \wedge \forall u \forall v \, (u \in v \in x \longrightarrow u \in x),\end{aligned}$$

which, given AF, states that x is an ordinal. Also, for any transitive class C containing a set x, the formulas $\phi(x)$ and $\phi^C(x)$ are equivalent, so, in particular,

$$\alpha = Val(\ulcorner \phi(x) \urcorner, L_\alpha) \in \mathcal{D}(L_\alpha) = L_\beta.$$

□

6.1. Σ_1 **Formulas and Absoluteness.**

We shall now consider a subclass of formulas without parameters of the language of set theory obtained by restricting quantification.

Definition 6.7. The class of Σ_1-formulas is the smallest class of first order formulas of the language $\{\in\}$ such that

(i) any quantifier-free formula is Σ_1
(ii) if ϕ, ψ are Σ_1, then so are $\phi \wedge \psi$ and $\phi \vee \psi$
(iii) if ϕ is Σ_1, then so is $\exists x \, \phi$
(iv) if ϕ is Σ_1, then so is $\forall x \in y \, \phi$, i.e. $\forall x \, (x \in y \longrightarrow \phi)$.

A class, class function, or class relation is said to be Σ_1 if it is defined in the universe $\mathcal{U}$ by some Σ_1-formula. So a class function is Σ_1 if it has a Σ_1 graph.

Remark:
The subclass of Δ_0-formulas is obtained by replacing (iii) by the following two conditions and otherwise changing "Σ_1" to "Δ_0".

(a) if ϕ is Δ_0, then so is $\exists x \in y,\, \phi$
(b) if ϕ is Δ_0, then so is $\neg\phi$.

Theorem 6.8. Suppose $\phi(x_1, ..., x_k)$ is Σ_1 and $M \subseteq N$ are classes with M transitive. Then for any $a_1, ..., a_k$ in M:

$$(\star) \qquad \phi^M(a_1, ..., a_k) \implies \phi^N(a_1, ..., a_k).$$

Remark:
Before proving this, we should acknowledge that most formulas are not Σ_1, but only equivalent to Σ_1-formulas in some background theory. However, suppose T is a theory and $\phi(\vec{x})$, $\psi(\vec{x})$ are formulas with ϕ Σ_1 such that $T \models (\phi(\vec{x}) \longrightarrow \psi(\vec{x}))$. Assume also that M, N satisfy T. Then $(\star)$ implies that also $\psi^M(\vec{a}) \implies \psi^N(\vec{a})$ for any $\vec{a}$ in M.

Proof.
We prove the result by induction on the construction of ϕ by the principles (i)–(iv). (i) and (ii) are trivial, so suppose that $\psi(x, x_1, ..., x_k)$ satisfies the induction hypothesis and that $a_1, ..., a_k$ belong to M.

Consider first case (iii):
If $\big(\exists x\, \psi(x, a_1, ..., a_k)\big)^M$ holds, then there is some b in $M \subseteq N$ such that $\psi^M(b, \vec{a})$, hence, by the induction hypothesis, also $\psi^N(b, \vec{a})$ and so $\big(\exists x\, \psi(x, \vec{a})\big)^N$.

For case (iv):
Suppose b is in M and that $\big(\forall x \in b\, \psi(x, \vec{a})\big)^M$, i.e. $\forall x\, \big(M(x) \wedge x \in b \longrightarrow \psi^M(x, \vec{a})\big)$. By the induction hypothesis for ψ, we have for any x in M,

$$\psi^M(x, \vec{a}) \implies \psi^N(x, \vec{a}).$$

So $\forall x\, \big((M(x) \wedge x \in b) \longrightarrow \psi^N(x, \vec{a})\big)$. But as M is transitive, $b \subseteq M \subseteq N$, so

$$\forall x\, \big((N(x) \wedge x \in b) \longrightarrow \psi^N(x, \vec{a})\big)$$

that is, $\big(\forall x \in b\ \psi(x, \vec{a})\big)^N$, which shows the induction hypothesis for the formulas $\exists x\, \psi$ and $\forall x \in y\, \psi$.

□

Lemma 6.9. Suppose $\psi(y, \vec{z})$ is a Σ_1-formula and $y = F(\vec{x})$ is a Σ_1 class function. Then $\psi\big(F(\vec{x}), \vec{z}\big)$ defines a Σ_1 class relation.

Proof.
Just note that $\psi\big(F(\vec{x}), \vec{z}\big)$ is given by

$$\psi\big(F(\vec{x}), \vec{z}\big) \iff \exists y\, \big(F(\vec{x}) = y \wedge \psi(y, \vec{z})\big).$$

□

Lemma 6.10. Suppose a is a set defined by a Σ_1-formula $\phi(x)$, i.e. $x = a \iff \phi(x)$, and that $\psi(y, \vec{z})$ is Σ_1. Then also $\psi(a, \vec{z})$ is Σ_1.

Proof.

$\psi(a, \vec{z}) \iff \exists x\, (\phi(x) \wedge \psi(x, \vec{z}))$.

□

Fact: (ZF)

Ord is a Δ_0 class and ω is defined by a Σ_1-formula.

Proof.

Note that

$$\mathrm{Ord}(x) \iff \forall y \in x\, \forall z \in x\, (y \in z \vee z \in y \vee z = y) \wedge \forall y \in x\, \forall z \in y,\ z \in x.$$

For the definition of ω, we check that the following formulas are Σ_1.

$\underline{x = \varnothing}$: $\forall y \in x,\ y \neq y$
$\underline{y = x \cup \{x\}}$: $\forall z \in y\, (z \in x \vee z = x) \wedge \forall z \in x,\ z \in y \wedge x \in y$
$\underline{x = \omega}$: $\varnothing \in x \wedge \forall y \in x,\ y \cup \{y\} \in x \wedge \forall y \in x\, (y = \varnothing \vee \exists z\, y = z \cup \{z\})$

□

Proposition 6.11 (ZF)**.** Suppose $H(-)$ is a Σ_1 class function of one variable defined on all functions with ordinal domain. Then the unique class function F defined on all ordinals and satisfying $F(\alpha) = H(F{\restriction_\alpha})$ is Σ_1 too.

Proof.

Again, we successively verify that certain objects, classes, and class relations are Σ_1:

$\underline{z = \{x, y\}}$: $x \in z \wedge y \in z \wedge \forall u \in z\, (x = u \vee y = u)$
$\underline{z = (x, y)}$: $z = \{\{x\}, \{x, y\}\}$
$\underline{y \subseteq x}$: $\forall z \in y,\ z \in x$
$\underline{z = x \cup y}$: $\forall u \in z\, (u \in x \wedge u \in y) \wedge \forall u \in x,\ u \in z \wedge \forall u \in y,\ u \in z$
$\underline{z = x \cap y}$: $\forall u \in z\, (u \in x \wedge u \in y) \wedge \forall u \in x\, (u \in y \longrightarrow u \in z)$
$\underline{z = x \smallsetminus y}$: $\forall u \in z\, (u \in x \wedge u \notin y) \wedge \forall u \in x\, (u \notin y \longrightarrow u \in z)$
$\underline{z \subseteq x \times y}$: $\forall u \in z\, \exists a\, \exists b\, (a \in x \wedge b \in y \wedge u = (a, b))$
$\underline{z \supseteq x \times y}$: $\forall a \in x\, \forall b \in y\, ((a, b) \in z)$
$\underline{f \text{ is a function from } x \text{ to } y}$:

$$f \subseteq x \times y \wedge \forall z \in x\, \exists v\, (v \in y \wedge (z, v) \in f) \ \wedge$$
$$\forall z \in x\, \forall v \in y\, \forall u \in y\, ((z, v) \notin f \vee (z, u) \notin f \vee v = u)$$

[Here, $(z, v) \notin f \iff \exists a\, (a \notin f \wedge a = (z, v))$ is Σ_1 in the variables z, v, f]
$\underline{f \text{ is a function with domain } x}$: $\exists y\, f : x \to y$
$\underline{g = f{\restriction_x}}$: $\exists a\, \exists b\, (x \subseteq a \wedge f : a \to b \wedge g : x \to b \wedge g \subseteq f)$
$\underline{f \text{ is a function}}$:$\exists a\, \exists b\, f : a \to b$
$\underline{f \text{ is a function and } f(x) = y}$: f is a function and $(x, y) \in f$.

So finally, we can write $F(\alpha) = y$ by

$$\mathrm{Ord}(\alpha) \wedge \exists f\, (f \text{ is a function with domain } \alpha \wedge \forall \beta \in \alpha\, (f(\beta) = H(f{\restriction_\beta})) \wedge y = H(f))$$

□

Fact:
The following are Σ_1.

f is an injection from x to y:

$$(f : x \to y) \wedge \forall a \in x \, \forall b \in x \, \forall c \in y \left(\Big((a, c) \notin f \Big) \vee \Big((b, c) \notin f \Big) \vee (a = b) \right)$$

f is a surjection from x to y: $(f : x \to y) \wedge \forall b \in y \, \exists a \, (a \in x \wedge f(a) = b)$

$h = f \circ g$:

$$\exists x \, \exists y \, \exists z \Bigg((g : x \to y) \wedge (f : y \to z) \wedge (h : x \to z) \ \wedge$$

$$\forall a \in x \, \exists b \, \exists c \Big(\big(f(a) = b \big) \wedge \big(g(b) = c \big) \wedge \big(h(a) = c \big) \Big) \Bigg)$$

Proposition 6.12. $B = \mathcal{D}(A)$ is Σ_1 in the variables A, B.

The proof, as before, proceeds by successively verifying that the following classes and properties are Σ_1.[a]

Proof.

$z = \mathcal{F}_0$: I.e., z is the set of atomic formulas. For this, note that "$x = 0$" and "$x = y \cup \{y\}$" are Σ_1 and so also "$x = 1$", "$x = 2$", "$x = 3$", and "$x = 4$" are Σ_1. Since $\dot\neg$, $\dot\vee$, $\dot\exists$, $\dot\in$, $\doteq$ are $0, 1, 2, 3, 4$, and $\mathcal{V} = \omega \setminus \{0, 1, 2, 3, 4\}$ is also Σ_1, we have that

$$\mathcal{F}_0 = \big(\{ \dot\in \} \times \mathcal{V} \times \mathcal{V} \big) \cup \big(\{ \doteq \} \times \mathcal{V} \times \mathcal{V} \big)$$

is Σ_1 too.

$k < \omega \wedge z = \mathcal{F}_k$ is Σ_1 in the variables k and z:

$$\exists f \, \big[(f \text{ is a set with } \mathrm{dom}(f) = \omega) \wedge (f(0) = \mathcal{F}_0) \wedge$$

$$\forall u \in \omega \, \big[f(u+1) = f(u) \cup \big(\{ \dot\neg \} \times f(u) \big) \cup \big(\{ \dot\vee \} \times f(u) \times f(u) \big) \cup$$

$$\big(\{ \dot\exists \} \times \mathcal{V} \times f(u) \big) \big] \wedge (k < \omega) \wedge \big(z = f(u) \big) \big]$$

f is a $\mathcal{U}$-formula: $\exists k \, (f \in \mathcal{F}_k)$

$z = \mathcal{F}$: the set of $\mathcal{U}$-formulas

$f \in \mathcal{F} \wedge a = var(f)$ is Σ_1 in the variables f and a.

We now show that the class relation "$Y = X^{var(f)} \wedge f \in \mathcal{F}$" in the three variables X, Y, f is Σ_1. Note, however, that the class relation "$Y = X^a$" is not Σ_1. To see this, observe that if Z is a transitive set, then for $Y, X, a \in Z$,

$$(Y = X^a)^Z \iff Y = \{ g \in Z \,|\, g : a \to X \}.$$

Since we can construct transitive sets Z with elements Y, X, a such that

$$Y = \{ g \in Z \,|\, g : a \to X \} \subsetneq \{ g \,|\, g : a \to X \},$$

being the set of functions from a to X is not preserved from Z to $\mathcal{U}$ and hence cannot be Σ_1.

[a] *TN:* All the FOL symbols, parens, and everything else made this obnoxious to TeX. In making revisions to this pdf, I'm reluctant to redo all the various sizes of parens, brackets, and others to make it readable. I'll do so iff people care, aka send me emails.

$\underset{\sim}{k < \omega \wedge Y = X^k}$ is Σ_1 in k, Y, X:

$$k \in \omega \wedge \exists f \left[(f \text{ is a function with domain } \omega) \wedge \big(f(0) = \{\varnothing\}\big) \wedge \forall n \in \omega \right.$$
$$\big[(\forall g \in f(n+1),\ g : n+1 \to X) \ \wedge$$
$$\left.\big(\forall h \in f(n)\, \forall x \in X,\ h \cup \{(n,x)\} \in f(n+1)\big)\big] \wedge \big(Y = f(k)\big)\right]$$

$\underset{\sim}{f \in \mathcal{F} \wedge Y = X^{var(f)}}$:

$$f \in \mathcal{F} \wedge \exists p\, \exists k < \omega \big[(p : var(f) \to k \text{ is a bijection}) \wedge$$
$$(\forall g \in X^k,\ g \circ p \in Y) \wedge (\forall h \in Y\, \exists g \in X^k,\ g \circ p = h)\big]$$

■

$\underline{f \text{ is a } \mathcal{U}\text{-sentence with parameters in } X}$: $\exists g \in \mathcal{F}\, \exists \delta : var(g) \to X,\ f = (g, \delta)$
$\underline{z = \mathcal{F}_X^0}$ where this latter is the set of $\mathcal{U}$-sentences with parameters in X:

$\forall f \in z\, (f$ is a $\mathcal{U}$-sentence with parameters in $X) \wedge \forall f \in \mathcal{F}\, \forall \delta \in X^{var(f)},\ (f, \delta) \in z$

$\underline{f \text{ is a } \mathcal{U}\text{-formula with parameters in } X \text{ and } x \text{ as a single free variable}}$
$\underline{z = \mathcal{F}_X^x}$ where the latter denotes the set of $\mathcal{U}$-formulas with parameters in X and a single free variable x
$\underline{f \in \mathcal{F}_X^0 \wedge t = Val(f, X)}$ is Σ_1 in variables f, X, t and where t can take the values 0, 1.

To see this, note that $t = Val(f, X)$ if and only if there is a function satisfying Tarski's recursive definition of truth eventually ending up with t.
$\underline{f \in \mathcal{F}_X^x \wedge y = Val(f, X)}$
$\underline{y \in \mathcal{D}(X)}$: $\exists x \in \mathcal{V}\, \exists f \in \mathcal{F}_X^x,\ y = Val(f, X)$
$\underline{z = \mathcal{D}(X)}$: $\forall y \in z\, \big(y \in \mathcal{D}(X)\big) \wedge \forall x \in \mathcal{V}\, \forall f \in \mathcal{F}_X^x\, (Val(f, X) \in z)$

□

Corollary 6.13 (ZF)**.** The class function $L : \mathrm{Ord} \to \mathcal{U}$ is Σ_1.

Proof.

L is a class function defined by transfinite recursion from the Σ_1 class function $\mathcal{D}(-)$

□

Theorem 6.14 (ZF)**.** L satisfies ZF + "$V = L$"

Proof.

Recall that L is a transitive class. Assume that L satisfies ZF. Then, as the class function $\alpha \longmapsto L_\alpha$ is well-defined in any model of ZF, we have that

$$\big[\alpha \longmapsto L_\alpha\big]^L$$

is a class function defined on all ordinals α in L. Moreover, since $y = L_\alpha$ is Σ_1 in the variables y, α, we have for all y, α in L:

$$\big[y = L_\alpha\big]^L \implies y = L_\alpha.$$

Now, suppose x belongs to L and find an ordinal α such that $x \in L_\alpha$ (not necessarily the least such ordinal). Then α belongs to L and, since the class of ordinals is Δ_0 and thus has Σ_1 complement, we have

$$\big[\mathrm{Ord}(\alpha)\big]^L.$$

Now, let y be the set in L such that $[y = L_\alpha]^L$. Then also $y = L_\alpha$ and so $x \in y$, hence also $[x \in y]^L$. So, finally, $[x \in L_\alpha]^L$, showing that

$$\big[\forall x\, \exists \alpha\; x \in L_\alpha\big]^L$$

i.e., $[V = L]^L$.

■

It now remains to show that L satisfies ZF.

Extensionality holds in L since L is transitive.

Union: Suppose a is constructible, say $a \in L_\alpha$. Then $b = \bigcup_{x \in a} x \subseteq L_\alpha$ since L_α is transitive and

$$b = Val(\dot{\exists}\, y\, (y \,\dot{\in}\, a \,\dot{\wedge}\, x \,\dot{\in}\, y),\, L_\alpha) \in L_{\alpha+1}.$$

Powerset: Suppose a is in L and let $b = \{x \subseteq a \,|\, x \text{ is in } L\}$. Find also α sufficiently large such that $b \subseteq L_\alpha$ and thus also $a \in L_\alpha$. But then

$$b = Val(\dot{\forall}\, y\, (y \,\dot{\in}\, x \mathrel{\dot{\longrightarrow}} y \,\dot{\in}\, a),\, L_\alpha) \in L_{\alpha+1}.$$

Infinity: ω belongs to L.

Replacement: Suppose $\phi(x, y, \vec{a})$ is a formula with parameters $a_1, ..., a_k$ in L that defines a class function in L. Suppose c is a constructible set and set

$$b = \big\{y \,|\, L(y) \wedge \exists x \in c\; \phi^L(x, y, \vec{a})\big\}.$$

Then $b \subseteq L_\alpha$ for some α large enough such that $a_1, .., a_k, c \in L_\alpha$. By the reflection scheme, we can find $\beta > \alpha$ a limit such that

$$\forall x, y, \vec{z} \in L_\beta\; \big(\phi^{L_\beta}(x, y, \vec{z}) \longleftrightarrow \phi^L(x, y, \vec{z})\big)$$

and so, in particular,

$$\forall x, y \in L_\beta\; \big(\phi^{L_\beta}(x, y, \vec{a}) \longleftrightarrow \phi^L(x, y, \vec{a})\big).$$

It follows that

$$b = Val(\ulcorner \exists x \in x\; \phi(x, y, \vec{a}) \urcorner,\, L_\beta) \in L_{\beta+1}.$$

Foundation: If $a \neq \varnothing$ belongs to L, pick $b \in a$ of minimal rank. Then b belongs to L and $a \cap b = \varnothing$.

□

Theorem 6.15. $V = L \implies$ principle of choice.
In particular, the principle of choice is consistent. Also,

$$V = L \implies V = L = OD = HOD.$$

Proof.

List the set of $\mathcal{U}$-formulas as $(f_n)_{n<\omega}$. Suppose X is a set well-ordered by a relation $\preceq$. We then define a well-ordering $\preceq'$ of $\mathcal{D}(X)$ as follows:

- The ordering $\preceq$ of X canonically induces a well-ordering $\preceq_1$ of the set $X^{<\omega}$ of finite sequences of elements of X.
- Now, if $A, B \in \mathcal{D}(X)$, put

$$A \preceq_2 B \iff \begin{cases} \text{there is } f_n(x, \vec{y}) \in \mathcal{F} \text{ and } \vec{a} \in X^{<\omega} \text{ with} \\ A = Val\big(f_n(x, \vec{a}), X\big) \text{ and for any } f_m(x, \vec{z}) \in \mathcal{F} \text{ and} \\ b \in X^{<\omega} \text{ with } B = Val\big(f_m(x, \vec{b}), X\big), \text{ either} \\ n < m, \text{ or } n = m \wedge \vec{a} \preceq_1 \vec{b}. \end{cases}$$

- Finally, let for $A, B \in \mathcal{D}(X)$:

$$A \preceq' B \iff \begin{cases} A, B \in X \ \wedge \ A \preceq_1 B \\ A, B \notin X \ \wedge \ A \preceq_2 B \\ A \in X \ \wedge \ B \notin X. \end{cases}$$

Note that then if $\preceq$ is a well-ordering of $X = L_\alpha$, then $\preceq'$ is a well-ordering of $L_{\alpha+1} = \mathcal{D}(L_\alpha)$ in which L_α is an initial segment on which the ordering agrees with $\preceq$.

Now, by transfinite induction, we define

- $\leqslant_0 =$ trivial ordering on $L_0 = \varnothing$
- $\leqslant_{\alpha+1} = \leqslant'_\alpha$
- $\leqslant_\lambda = \bigcup_{\alpha<\lambda} \leqslant_\alpha$ for λ limit.

Then, each $\leqslant_\alpha$ is a well-ordering of L_α, for each $\alpha < \beta$, $\leqslant_\alpha = \leqslant_\beta\restriction_{L_\alpha}$, and L_α is an initial segment of $(L_\beta, \leqslant_\beta)$.

Finally, let $\leqslant = \bigcup_{\alpha \in \mathrm{Ord}} \leqslant_\alpha$, which is a class well-ordering of L in which each initial segment is contained in some L_α.

□

6.2. **The Generalized Continuum Hypothesis in L.**

Theorem 6.16 (ZFC)**.** Suppose $F(-)$ is a Σ_1 class function of one variable. Then for any a in the domain of F, we have

$$\big|F(a)\big| \leq \big|\operatorname{cl}(a)\big| \oplus \aleph_0$$

where $\operatorname{cl}(a)$ is the transitive closure of a.

Proof.

Suppose $\phi(x, y)$ is a Σ_1 formula with $\phi(x, y) \iff F(x) = y$. Suppose a belongs to the domain of F and let α be large enough such that $a, F(a) \in V_\alpha$ and

$$\forall x, y \in V_\alpha \left(\phi^{V_\alpha}(x, y) \longleftrightarrow \phi(x, y)\right)$$

(of course, we can find such an α by reflection applied to ϕ).

Note that $\operatorname{cl}(\{a\}) = \{a\} \cup \operatorname{cl}(a) \subseteq V_\alpha$ since V_α is transitive and that $|\operatorname{cl}(\{a\})| = |\operatorname{cl}(a)| \oplus 1$. Let $\mathcal{A}$ denote the set of $\mathcal{U}$-sentences f with parameters in $\operatorname{cl}(\{a\})$ such that $V_\alpha \models f$. By Löwenheim-Skolem applied to $\operatorname{cl}(\{a\}) \subseteq V_\alpha$, there is a subset $\operatorname{cl}(\{a\}) \subseteq X \subseteq V_\alpha$ with $|X| \leq |\operatorname{cl}(a)| \oplus \aleph_0$ such that $X \models f$ for any $f \in \mathcal{A}$.

Since V_α is transitive, it is extensional and so

$$V_\alpha \models \ulcorner \forall x \, \forall y \left(x = y \longleftrightarrow \forall z \, (z \in x \longleftrightarrow z \in y)\right) \urcorner$$

Thus, V_α satisfies the axiom of extensionality and hence so does X.

Let now Y denote the Mostowski collapse and let $j : X \to Y$ be the corresponding isomorphism. That is, Y is a transitive set and j is a bijection such that

$$\forall x, y \in X \left(x \in y \longleftrightarrow j(x) \in j(y)\right)$$

It follows that Y satisfies every sentence in $\mathcal{A}$ where any parameter $x \in \operatorname{cl}(\{a\})$ is replaced by $j(x)$. However, since $\operatorname{cl}(\{a\})$ is already transitive, the collapsing map j is the identity on $\operatorname{cl}(\{a\})$, and so $\operatorname{cl}(\{a\}) \subseteq Y$ and $Y \models f$ for any $f \in \mathcal{A}$.

Now, $\ulcorner \exists y\, \phi(a,y) \urcorner \in \mathcal{A}$ and so $Y \models \ulcorner \exists y\, \phi(a,y) \urcorner$, hence for some $b \in Y$, we have $\phi^Y(a,b)$. Since ϕ is Σ_1 and Y is transitive, it follows that also $\phi(a,b)$, hence $F(a) = b \in Y$. Again, as Y is transitive, $F(a) \subseteq Y$ and so

$$|F(a)| \leq |Y| = |X| \leq |\operatorname{cl}(a)| \oplus \aleph_0.$$

□

Remark:

Since $|\mathcal{P}(\omega)| > \aleph_0 = |\operatorname{cl}(\omega)| \oplus \aleph_0$, we see that the class function $\mathcal{P}(-)$ cannot be Σ_1.

Remark:

If a is a Σ_1 definable set, i.e. the statement $x = a$ is Σ_1 in the variable x, then $|a| \leq \aleph_0$. For in this case, $a = F(\varnothing)$ for a Σ_1 class function.

Theorem 6.17 (ZFC)**.**

(i) $|L_\alpha| = |\alpha|$ for any ordinal $\alpha \geq \omega$.

(ii) for any constructible set a, $|\operatorname{order}(a)| \leq |\operatorname{cl}(a)| \oplus \aleph_0$

Proof.

(i): Since $\alpha \subseteq L_\alpha$, we have $|\alpha| \leq |L_\alpha|$. Conversely, note that $\alpha \longmapsto L_\alpha$ is Σ_1, so, by the previous theorem, $|L_\alpha| \leq |\operatorname{cl}(\alpha)| \oplus \aleph_0 = |\aleph|$.

(ii): Again, note that order$(-)$ is a Σ_1 class function since

$$\operatorname{order}(a) = \alpha \iff a \in L_\alpha \wedge \forall \beta \in \alpha\; a \notin L_\beta.$$

So the result follows from the preceding theorem.

□

Theorem 6.18 (ZFC)**.** If $V = L$ holds, then the Generalized Continuum Hypothesis (GCH) holds, i.e., for any infinite cardinal κ, $2^\kappa = \kappa^+$.[a]

Proof.

Suppose $a \subseteq \kappa$, then $|\operatorname{order}(a)| \leq |\operatorname{cl}(a)| \oplus \aleph_0 \leq \kappa$ and so $a \in L_\alpha$ for some $\alpha < \kappa^+$. So, $\mathcal{P}(\kappa) \subseteq L_{\kappa^+}$ and so $|\mathcal{P}(\kappa)| = 2^\kappa \leq |L_{\kappa^+}| = \kappa^+$.

□

Definition 6.19. A sentence is arithmetical if all quantifiers are of the form

$$\exists x \in V_\omega \qquad \text{or} \qquad \forall x \in V_\omega.$$

For example, since Peano Arithmetic is definable in V_ω, any statement in Peano Arithmetic is an arithmetical statement.[b]

---

[a] *TN:* Recall that we mentioned earlier that we would be proving the independence of CH. This is technically the first part of that proof. We are now showing that ZFC + "CH" is consistent if ZFC is, and we will later show, using methods of forcing, that ZFC + "¬CH" is also consistent if ZFC is.

[b] *TN:* This is one of the rare times that definitions work out very nicely.

Theorem 6.20 (ZF)**.** If an arithmetical statement ϕ is provable from ZFC+"$V{=}L$"+GCH, then ϕ is provable from ZF.

Proof.

$V_\omega = L_\omega \subseteq L$, ϕ is Σ_1 and so from ZF, we get that $\phi^L \implies \phi$. Now, suppose $\psi_1, \psi_2, ..., \psi_n, \phi$ is a proof of ϕ from the axioms of ZFC+"$V{=}L$". If ψ_i is an axiom, then also ψ_i^L holds (as can be prove only supposing ZF i $\mathcal{U}$), so $\psi_1^L, \psi_2^L, ..., \psi_n^L, \phi^L, \phi$ is a proof of ϕ only using axioms of ZF. $\square$

7. Forcing

Whereas Gödel's construction of L provided us with a model of ZFC+"$V{=}L$"+GCH, we shall now present Paul Cohen's method of forcing giving us a model of ZFC+$\neg$CH.

Main Idea:
If $\mathcal{U}$ is a model of ZFC and M is a countable, transitive set in $\mathcal{U}$, then forcing is a method for adjoining a new set x to M, assumed to be somehow generic, to obtain a new countable set $M[x]$ still satisfying ZF.

We also have tools for studying this adjoining of x to M and to control the properties of $M[x]$ in terms of M and x. First, let us see how we can obtain countable transitive set models of ZFC.

Theorem 7.1. Suppose T is a theory in the language of set theory extending ZFC and let m be a new constant symbol. Then, if T is consistent, so is the theory $T^* := T + T^m +$ "m is a countable, transitive set$''$.

Proof.

Suppose towards a contradiction that T^* is inconsistent. Then there is a finite fragment of T^* that is inconsistent,[a] and so there are sentences $\phi_1, ..., \phi_n \in T$ such that

$$T + \bigwedge_{i=1}^{n} \phi_i^m + \text{“}m \text{ is a countable, transitive set}''$$

is inconsistent.

Now, since T is consistent, let $\mathcal{U}$ be a model of T. By the reflection scheme, find an ordinal α such that $\left(\bigwedge_{i=1}^{n} \phi_i\right)^{V_\alpha}$ holds. Also, by Löwenheim-Skolem, there is a countable set $X \subseteq V_\alpha$ such that for any $\mathcal{U}$-formula f, we have $X \models f \iff V_\alpha \models f$. In particular, since V_α satisfies the axiom of extensionality, so does X, and as $V_\alpha \models \ulcorner \phi_i \urcorner$ for all $i \leq n$, we have $X \models \ulcorner \phi_i \urcorner$ for all $i \leq n$ as well.

Let $j : X \to Y$ be the canonical map from X onto its Mostowski collapse. Then Y is a countable transitive set and $Y \models \ulcorner \phi_i \urcorner$ for all $i \leq n$. We can therefore expand $\mathcal{U}$ to a model of $T + \bigwedge_{i=1}^{n} \phi_i^m +$ "m is a countable, transitive set$''$ by interpreting m as Y, contradicting our assumption.

$\square$

7.1. **Generic Extensions.**

In the following, suppose M is a countable transitive set satisfying ZF in a universe $\mathcal{U}$ satisfying ZFC. Assume also that $(\mathbb{P}, \leq) \in M$ is a poset (partially ordered set) in M, $\mathbb{P} \neq \varnothing$.

[a] *TN:* Isn't it just so useful that proofs are finite?

Definition 7.2.

- Elements of $\mathbb{P}$ are called forcing conditions, and if $p \leq q$, we say that p is stronger than q.[a]
- Two forcing conditions p, q are said to be compatible if $\exists r \in \mathbb{P}\,(r \leq p \wedge r \leq q)$. Otherwise, p and q are incompatible, written $p \perp q$.
- A subset $D \subseteq \mathbb{P}$ is dense if $\forall p \in \mathbb{P}\,\exists q \in D\ (q \leq p)$
- D is said to be saturated if $\forall p \in D\,\forall q \leq p\ (q \in D)$.
- D is said to be predense if $\forall p \in \mathbb{P}\,\exists q \in D$ (p and q are compatible).
- For any set $X \subseteq \mathbb{P}$, let
$$\tilde{X} := \{p \in \mathbb{P} \mid \exists q \in X\ p \leq q\}$$
denote the saturation of X. Note that if X is predense, then $\tilde{X}$ is dense.

Definition 7.3 (Generic Extension)**.** Now, suppose $G \subseteq \mathbb{P}$ is a subset, *not* necessarily belonging to M, but only to $\mathcal{U}$. We say that G is $\mathbb{P}$-generic over M if

(i) $\forall p \in G\,\forall q \in \mathbb{P}\,(p \leq q \longrightarrow q \in G)$
(that is, G is closed upwards)
(ii) $\forall p, q \in G$, p and q are compatible
(iii) $\forall D \in M$ (if D is a dense subset of $\mathbb{P}$, then $D \cap G \neq \varnothing$)

Since $\mathbb{P}$-generics are upwards closed, we see that (iii) can be replaced by either

(iii)' $\forall D \in M$ (if D is a predense subset of $\mathbb{P}$, then $D \cap G \neq \varnothing$) or
(iii)" $\forall D \in M$ (if D is a dense and saturated subset of $\mathbb{P}$, then $D \cap G \neq \varnothing$)

Lemma 7.4. Suppose G is $\mathbb{P}$-generic over M. Then $\forall p \in \mathbb{P}\,\big(p \notin G \longleftrightarrow \exists q \in G\ (p \perp q)\big)$.

Proof.

Since any two elements of G are compatible, if $q \in G$ and $p \perp q$, then $p \notin G$.
Conversely, suppose $p \notin G$ and consider the set
$$D = \{q \in \mathbb{P} \mid (q \leq p) \vee (q \perp p)\}.$$
We claim that D is dense. For if $r \in \mathbb{P}$ is given, then either $r \perp p$, in which case $r \in D$, or there is $q \in \mathbb{P}$ such that $q \leq r \wedge q \leq p$, in which case $q \in D$, showing density.

Also, since M satisfies ZF, the construction of D can be done inside M, and so $D \in M$. In other words, $D \in M$ is a dense subset of $\mathbb{P}$. So, as G is $\mathbb{P}$-generic over M, we have $G \cap D \neq \varnothing$.

So, let $q \in G \cap D$ be any element. Note that if $q \leq p$, then as G is closed upwards, also $p \in G$, which is by assumption not the case. So instead we must have $p \perp q$.

$\square$

Lemma 7.5. Suppose G is $\mathbb{P}$-generic over M. Then $\forall p, q \in G\,\exists r \in G\,(r \leq p \wedge r \leq q)$. That is, any two elements of G have a common minorant.

Proof.

Fix $p, q \in G$ and let
$$D = \big\{r \in \mathbb{P} \mid r \perp p \vee (r \leq p \wedge r \perp q) \vee (r \leq p \wedge r \leq q)\big\}.$$
Again, since M satisfies ZF, the construction of D can be done in M, and so $D \in M$. Moreover, D is dense: for given any $t \in \mathbb{P}$, either $t \perp p$, and so $t \in D$, or there is $s \in \mathbb{P}$ with $s \leq t \wedge s \leq p$. In the latter case, either $s \perp q$, where $s \in D$, or there is some $r \leq s \leq p \wedge r \leq q$, hence $r \in D$.

[a] *TN:* This is not a typo.

So pick some $r \in G \cap D$. Since any two elements of G are compatible, this must mean that $r \leq p \wedge r \leq q$, and so p, q have a common minorant in G.

□

By induction, we see that

Lemma 7.6. Any finite subset of G has a common minorant in G. □

Definition 7.7. Suppose $D \subseteq \mathbb{P}$ and $p \in \mathbb{P}$. We say that D is dense below p if

$$\forall q \in \mathbb{P} \left(q \leq p \longrightarrow \exists r \in D \; (r \leq q)\right).$$

Lemma 7.8. Suppose G is $\mathbb{P}$-generic over M and assume $D \in M$ is dense below some $p \in G$. Then $G \cap D \neq \varnothing$.

Proof.

Note that $E = D \cup \{q \in \mathbb{P} \,|\, q \perp p\} \in M$ and E is dense in $\mathbb{P}$. For if $q \in \mathbb{P}$ and q has no minorant in D, then q and p cannot have any common minorant, hence $p \perp q$ and thus also $q \in D$.
It follows that $G \cap E \neq \varnothing$ and so, as any two elements of G are compatible, also $G \cap D \neq \varnothing$.

□

Definition 7.9. A subset $X \subseteq \mathbb{P}$ is an antichain if $\forall p \in X \, \forall q \in X \, (p \neq q \longrightarrow p \perp q)$.
An antichain is said to be maximal if it is not contained in any larger antichain.

Note:
An antichain is maximal if and only if it is predense in $\mathbb{P}$.
In particular, if $X \in M$ is a maximal antichain and G is $\mathbb{P}$-generic over M, then $G \cap X \neq \varnothing$. (In particular, being a maximal antichain is Δ_0.)

Lemma 7.10 (Assuming M satisfies AC)**.** Suppose $D \in M$ is a dense subset of $\mathbb{P}$. Then there is a maximal antichain $X \subseteq D$, $X \in M$.

Proof.

Work in M:

We order the set of all antichains $X \subseteq D$ by inclusion and note that by Zorn's Lemma, this family has a maximal element X, which then is predense in $(D, \leq)$. So for any $p \in \mathbb{P}$, there is $q \in D$, $q \leq p$, and so , by predensity of X in D, some $r \in X$ compatible with q and thus also with p. So X is predense in $\mathbb{P}$ and hence a maximal antichain.

□

Theorem 7.11 (Assuming M satisfies AC)**.** Assume $G \subseteq \mathbb{P}$. Then G is $\mathbb{P}$-generic over M if and only if

(i) any two elements of G are compatible
(ii) if $X \in M$ is a maximal antichain in $\mathbb{P}$, then $G \cap X \neq \varnothing$.

Proof.

We have already seen that if G is $\mathbb{P}$-generic over M, then (i) and (ii) hold. For the inverse, note that if G intersects any maximal antichain $X \in M$, then G also intersects any dense subset $D \in M$.

Finally, to see that G is closed upwards, suppose $p \in G$, $q \in \mathbb{P}$ and $p \leq q$. We let $D_q := \{r \in \mathbb{P} \,|\, r \perp q\}$ and let $X \subseteq D_q$, $X \in M$ be a maximal antichain of the poset $(D_q, \leq)$. Then also

$X \cup \{q\} \in M$ and is a maximal antichain in $\mathbb{P}$. So $G \cap \big(X \cup \{q\}\big) \neq \varnothing$, and hence $q \in G$, since otherwise G would contain two incompatible elements.

□

The following result tells us that for the purposes of forcing, we can work with $(D, \leq)$ instead of $(\mathbb{P}, \leq)$ for any dense subset $D \in M$.

Theorem 7.12. Suppose $D \in M$ is a dense subset of $\mathbb{P}$. Then if G is $\mathbb{P}$-generic over M, also $G \cap D$ is D-generic over M. Conversely, for any $H \subseteq D$ which is D-generic over M, there is a unique $G \subseteq \mathbb{P}$ which is $\mathbb{P}$-generic over M and such that $H = G \cap D$. In fact, $G = \{p \in \mathbb{P} \mid \exists q \in H,\ q \leq p\}$.

Proof.

Suppose G is $\mathbb{P}$-generic over M. Then $G \cap D$ is D-generic over M. For if $E \subseteq D$, $E \in M$, is dense in D, i.e. $\forall p \in D\, \exists q \in E,\ q \leq p$, then E is also dense in $\mathbb{P}$, hence $G \cap E = (G \cap D) \cap E \neq \varnothing$.

Also, if $p, q \in G \cap D$, then p, q are compatible in $\mathbb{P}$, hence $r \leq p$, $r \leq q$ for some $r \in \mathbb{P}$. Since D is dense in $\mathbb{P}$, there is $s \in D$, $s \leq r$, hence $s \leq p$, $s \leq q$, and so p, q are compatible in G. Finally, $G \cap D$ is closed upwards in D.

Conversely, suppose $H \subseteq D$ is D-generic over M and let $G = \{p \in \mathbb{P} \mid \exists q \in H,\ q \leq p\}$ be the upwards closure of H in $\mathbb{P}$. Notice, since any $\mathbb{P}$-generic over M is upwards closed, so G is contained in any $F \supseteq H$ that is $\mathbb{P}$-generic over M. Also, if $F \supseteq H$ is $\mathbb{P}$-generic over M, $F \cap D = H$ and $p \in F$, let $X = \{q \in D \mid q \leq p\}$. Then obviously $X \in M$ is dense below p in $\mathbb{P}$ and so $F \cap X \neq \varnothing$. It follows that there is $q \in F \cap D = H$, $q \leq p$, hence $p \in G$. So $F = G$.

So, if G is actually $\mathbb{P}$-generic over M, then G is unique with the property that $G \cap D = H$. For this, we note that G is upwards closed. And, if $p, q \in G$, then there are $r, s \in H$ such that $r \leq p$, $s \leq q$. Since r, s are compatible, so are p, q. Finally, if $E \subseteq \mathbb{P}$, $E \in M$ is a saturated dense set, then $E \cap D \in M$ is dense in D and so $H \cap E \cap D \neq \varnothing$, whereby also $G \cap E \supseteq G \cap D \cap E = H \cap D \cap E \neq \varnothing$.

□

Theorem 7.13 (Rasiowa-Sikorski (ZFC)[a])**.** Suppose M is a countable transitive set satisfying ZF and $(\mathbb{P}, \leq) \in M$ is a poset. Then for any $p \in \mathbb{P}$, there exists $G \subseteq \mathbb{P}$ which is $\mathbb{P}$-generic over M and such that $p \in G$.

Proof.

Since M is countable, so is $\mathcal{P}(\mathbb{P}) \cap M$, and we can therefore find a sequence $(D_n)_{n<\omega}$ enumerating all dense subsets $D \subseteq \mathbb{P}$, $D \in M$. Let also $\pi : \mathcal{P}(\mathbb{P}) \setminus \{\varnothing\} \to \mathbb{P}$ be a choice function on $\mathbb{P}$.
By induction on $n < \omega$, we define a decreasing sequence $(p_n)_{n<\omega}$ of elements of $\mathbb{P}$ as folows:

- $p_0 := p$
- $p_{n+1} := \pi\big(\{q \in D_n \mid q \leq p_n\}\big)$ (It should be noted that since D_n is dense in $\mathbb{P}$, we know that $\{q \in D_n \mid q \leq p_n\} \neq \varnothing$, hence this sequence is well-defined).

Then $p_0 \geq p_1 \geq \ldots$ and $p_{n+1} \in D_n$ for any $n < \omega$. Letting $G = \{q \in \mathbb{P} \mid \exists n < \omega,\ p_n \leq q\}$, we see that G is upwards closed, any two elements of G are compatible, and G intersects any dense $D \subseteq \mathbb{P}$, $D \in M$. Moreover, $p \in G$.

□

[a] *TN:* Wikipedia says that this theorem "is one of the most fundamental facts used in the technique of forcing."

7.2. **Mostowski Collpase of a Well-founded Relation.**

Definition 7.14. Suppose A is a set and $R \subseteq A \times A$ is a binary relation. We say that R is well-founded if there is no sequence $(a_n)_{n<\omega}$ such that $a_{n+1} R a_n$ [or $(a_{n+1}, a_n) \in R$] for all $n < \omega$. Equivalently, R is well-founded if for any non-empty set $X \subseteq A$, there is $a \in X$ such that $\forall b \in X$, $\neg bRa$, ie $[(b,a) \notin R]$.

Suppose R is a binary relation on a set A. A function Φ with domain A is said to be an R-contraction if for any $a \in A$

$$\Phi(a) = \{\Phi(b) \mid b \in A \wedge bRa\}.$$

Theorem 7.15 (ZF$^-$)**.** If R is a well-founded relation on a set A, then there exists a unique R-contraction Φ.

Proof.

Uniqueness: Suppose towards a contradiction that Φ and Ψ are distinct R-contractions and let $X = \big\{a \in A \mid \Phi(a) \neq \Psi(a)\big\} \neq \varnothing$. Since R is well-founded, there is $a \in X$ such that $\forall b \in A\,(bRa \longrightarrow b \notin X)$. But then

$$\begin{aligned}\Phi(a) &= \{\Phi(b) \mid b \in A \wedge bRa\} \\ &= \{\Psi(b) \mid b \in A \wedge bRa\} = \Psi(a)\end{aligned}$$

contradicting $a \in X$.

Existence: Note that if $X \subseteq A$ is downwards R-closed and Φ is an R-contraction on A, then $\Phi\restriction_X$ is an R-contraction on X. Also, if $X, Y \subseteq A$ are both downwards closed, then so is $X \cap Y$. Therefore, if $X, Y \subseteq A$ and Φ, Ψ are R-contractions on X and Y respectively, then $\Phi\restriction_{X\cap Y}$, $\Psi\restriction_{X\cap Y}$ are R-contractions on $X \cap Y$ and must therefore agree. Therefore, if Φ is the union of all R-contractions defined on downwards closed subsets of A, we see that Φ is an R-contraction whose domain is a downwards closed subset of A.

We claim that $A = \mathrm{dom}(\Phi)$. For if not, let $a \in A \setminus \mathrm{dom}(\Phi)$ be such that $\forall b \in A\,\big(bRa \longrightarrow b \in \mathrm{dom}(\Phi)\big)$, which exists since R is well-founded. But then we can define Ψ on $\mathrm{dom}(\Phi) \cup \{a\} \supsetneq \mathrm{dom}(\Phi)$ by $\Psi(b) = \Phi(b)$ for $b \in \mathrm{dom}(\Phi)$, and $\Psi(a) = \{\Phi(b) \mid b \in A \wedge bRa\}$. Then Ψ is an R-contraction defined on a strictly larger downwards closed subset of A, which is absurd.

□

7.3. **Construction of Generic Extensions.**

Suppose M is a countable transitive set model of ZF and $(\mathbb{P}, \leq) \in M$ is a poset. Suppose also that G is $\mathbb{P}$-generic over M. We define the following relation R on M using G:

For $x, y \in M$

$$xRy \Longleftrightarrow \exists p \in G\,\big[(x,p) \in y\big].$$

Since $xRy \Longrightarrow \mathrm{rk}(x) < \mathrm{rk}(y)$, we see that R is well-founded on M.

By the preceding theorem, there is a unique R-contraction, Φ with domain M, and we shall denote the image $\Phi[M]$ by $M[G]$. $M[G]$ is called a generic extension of M. Our goal is now to show that $M[G]$ *itself* is a countable transitive set model of ZF and eventually study the relation between M, $(\mathbb{P}, \leq)$, G, and $M[G]$.

Lemma 7.16. $M \subseteq M[G]$.

Proof.

Working in M, we define a class function denoted $y = \widehat{x}$ by induction on the rank of x:

$$\widehat{x} = \big\{(\widehat{z}, p) \,|\, z \in x \wedge p \in \mathbb{P}\big\} \in M.$$

Let $\widehat{M}$ denote the set $\widehat{M} = \{\widehat{x} \,|\, x \in M\}$, which is a class in M. We claim that $\widehat{M} \subseteq M$ is downwards R-closed. For if $\widehat{x} \in \widehat{M}$ and $uR\,\widehat{x}$, then there is some $p \in G$ such that $(u,p) \in \widehat{x}$, i.e. $u = \widehat{z}$ for some $z \in x \subseteq M$. Thus, $u = \widehat{z} \in \widehat{M}$ too.

Also, the map $x \in M \longmapsto \widehat{x} \in \widehat{M}$ is injective. For if not, let $x \in M$ have minimal rank such that there is $y \in M$, $y \neq x$, $\widehat{y} = \widehat{x}$. Then for any $z, u \in M$, if $z \in x$, we have

$$z = u \Longleftrightarrow \widehat{z} = \widehat{u}.$$

On the other hand, if $z \in x \smallsetminus y$ and $p \in \mathbb{P}$, then $(\widehat{z}, p) \in \widehat{x} = \widehat{y}$, and so $(\widehat{z}, p) = (\widehat{u}, p)$ for some for some $u \in y$, hence $u = z \in y$, which is absurd.

We therefore have that for any $x, y \in M$

$$(\star) \qquad x \in y \Longleftrightarrow \widehat{x} R\,\widehat{y}.$$

There the implication from left to right is direct from the definition of $\widehat{y}$, and conversely, if $\widehat{x} R\,\widehat{y}$, then $(\widehat{x}, p) \in \widehat{y}$ for some $p \in G$, hence $\widehat{x} = \widehat{z}$ for some $z \in M$ such that $z \in y$. As $\widehat{\cdot}$ is injective, also $x = z \in y$. Thus, the inverse map $\widehat{x} \longmapsto x$ is an isomorphism of $(\widehat{M}, R)$ with $(M, \in)$.

Also, by $(\star)$, $x = \{y \in M \mid \widehat{y} R\,\widehat{x}\}$, which shows that $\widehat{x} \longmapsto x$ is the unique R-contraction defined on $\widehat{M}$. It follows that for any $x \in M$, $\Phi(\widehat{x}) = x \in M[G]$. So $M \subseteq M[G]$.

□

We shall reuse the construction of $\widehat{x}$ above, so let us spell it out:
Construction:
For every $x \in M$,

$$\widehat{x} := \big\{(\widehat{y}, p) \,|\, y \in x \wedge p \in \mathbb{P}\big\}$$

is said to be a $\mathbb{P}$-name for x. Then for any $x, y \in M$,

$$x \in y \Longleftrightarrow \widehat{x} R\,\widehat{y} \qquad \text{and} \qquad \Phi(\widehat{x}) = x.$$

Lemma 7.17. $G \in M[G]$.

Proof.

Set $\Gamma := \big\{(\widehat{p}, p) \,|\, p \in \mathbb{P}\big\} \in M$. We claim that $\Phi(\Gamma) = G$, hence $G \in M[G] = \Phi[M]$. To see this, suppose first that $p \in G$. Then $(\widehat{p}, p) \in \Gamma$ and so $\widehat{p} R \Gamma$, hence $p = \Phi(\widehat{p}) \in \Phi(\Gamma)$. So $G \subseteq \Phi(\Gamma)$. Conversely, suppose $y \in \Phi(\Gamma)$. Then there is $z \in M$ such that $\Phi(z) = y$ and $p \in G$ such that $(z, p) \in \Gamma$. It follows that $z = \widehat{p}$, hence $y = \Phi(z) = \Phi(\widehat{p}) = p \in G$. So $\Phi(\Gamma) = G$ and $G \in M[G]$.

□

Remark:

In general, $G \notin M$, and so no object of M is interdefinable with G. On the other hand, R and Φ are defined using G. The $\mathbb{P}$-names $\widehat{x}$ for $x \in M$ are defined within M without using G, but will be mapped onto $x \in M$ for any choice of $G \subseteq \mathbb{P}$ that is $\mathbb{P}$-generic over M.

Lemma 7.18. For any $x \in M$, $\mathrm{rk}(\Phi(x)) \leq \mathrm{rk}(x)$.

Proof.

Recall that $xRy \implies \mathrm{rk}(x) < \mathrm{rk}(y)$. Suppose that $\mathrm{rk}(\Phi(x)) > \mathrm{rk}(x)$ for some $x \in M$ and choose such an x with minimal rank.[a] Then

$$\mathrm{rk}(\Phi(x)) = \mathrm{rk}(\{\Phi(y) \mid yRx\}) = \big[\sup_{yRx} \mathrm{rk}(\Phi(y))\big] + 1$$

$$\leq \big[\sup_{yRx} \mathrm{rk}(y)\big] + 1 \qquad \text{(by the induction hypothesis)}$$

$$\leq \mathrm{rk}(x), \qquad \text{contradicting our assumption.}$$

□

Lemma 7.19. $M \cap \mathrm{Ord} = M[G] \cap \mathrm{Ord}$.

Proof.

The inclusion from left to right follows from $M \subseteq M[G]$. Conversely, if $\alpha \in M[G]$ is an ordinal, then $\alpha = \Phi(x)$ for some $x \in M$, hence $\alpha + 1 = \mathrm{rk}\,\alpha \leq \mathrm{rk}\,x \in M$. Since M is transitive, also $\alpha \in M$.

□

Lemma 7.20. $M[G]$ satisfies the axioms of extensionality, foundation, and infinity.

Proof.

$M[G]$ is a transitive set and $\omega \in M[G]$.

□

Lemma 7.21. $M[G]$ satisfies the union axiom.

Proof.

Since $M[G]$ is the image of M by the R-contraction Φ, we need to show that for any $a \in M$, there is some $b \in M$ such that

$$\Phi(b) = \{x \mid \exists y \in \Phi(a),\, x \in y\} = \bigcup \Phi(a).$$

So, given $a \in M$, let

$$b = \big\{(y, r) \in \mathrm{cl}(a) \times \mathbb{P} \mid \exists p, q \geq r\, \exists x\big((y, p) \in x \wedge (x, q) \in a\big)\big\},$$

which belongs to M.

Assume $c \in \Phi(b)$. Then $c = \Phi(y)$ for some $y \in M$, $r \in G$ such that $(y, r) \in b$. Thus, for some $p, q \geq r$ and $x \in M$, we have $(y, p) \in x$ and $(x, q) \in a$. Since G is upwards closed, $p, q \in G$, so yRx and xRa, hence $c = \Phi(y) \in \Phi(x) \in \Phi(a)$, and so $c \in \bigcup \Phi(a)$. I.e. $\Phi(b) \subseteq \Phi(a)$.

Conversely, assume $c \in \bigcup \Phi(a)$ and let $d \in \Phi(a)$ be such that $c \in d$. Then there is $x \in M$ and $q \in G$ such that $(x, q) \in a$ and $\Phi(x) = d$, and $p, \in G$, $y \in M$ such that $(y, p) \in x$ and $\Phi(y) = c$. Pick $r \in G$ such that $r \leq p$, $r \leq q$. Then $(y, r) \in b$ and so $c = \Phi(y) \in \Phi(b)$, showing $\bigcup \Phi(a) \subseteq \Phi(b)$.

□

[a] *TN:* I know I've already said this, but well-orderings are great.

7.4. **Definition of Forcing (Strong Forcing).**
Suppose M is a countable transitive set model of ZF and $(\mathbb{P}, \leq) \in M$ a poset. We shall write formulas in the language of set theory with the single parameter $(\mathbb{P}, \leq)$ and three free variables p, x, y, where p ranges over $\mathbb{P}$ and x, y over M. These will be denoted

$$p \Vdash \underline{x} \in \underline{y}, \quad p \Vdash \underline{x} \notin \underline{y}, \quad p \Vdash \underline{x} \neq \underline{y}, \quad p \Vdash \underline{x} = \underline{y}$$

and we will ensure that

(1) $p \Vdash \underline{x} \in \underline{y} \iff \exists r \geq p \, \exists z \big((z, r) \in y \wedge p \Vdash \underline{z} = \underline{x}\big)$
(2) $p \Vdash \underline{x} \neq \underline{y} \iff \exists r \geq p \, \exists z \left[\big((z, r) \in y \wedge p \Vdash \underline{z} \notin \underline{x}\big) \vee \big((z, r) \in x \wedge p \Vdash \underline{z} \notin \underline{y}\big)\right]$
(3) $p \Vdash \underline{x} \notin \underline{y} \iff \forall q \leq p \, \big(q \nVdash \underline{x} \in \underline{y}\big)$
(4) $p \Vdash \underline{x} = \underline{y} \iff \forall q \leq p \, \big(q \nVdash \underline{x} \neq \underline{y}\big)$.

The symbol $\Vdash$ is called the forcing relation and, e.g., "$p \Vdash \underline{x} \in \underline{y}$" should be read as "$p$ forces $\underline{x} \in \underline{y}$".

Since the statements (1), (2), (3), and (4) are defined by induction on the rank of x, y, it is intuitively clear that this is well-defined. But to see that they are class relations, we defined a well-founded ordering $\prec$ of $M \times M$ by

$$(x_0, y_0) \prec (x_1, y_1) \iff \begin{cases} \max(\mathrm{rk}(x_0), \mathrm{rk}(y_0)) < \max(\mathrm{rk}(x_1), \mathrm{rk}(y_1)), & \text{or} \\ \max(\mathrm{rk}(x_0), \mathrm{rk}(y_0)) = \max(\mathrm{rk}(x_1), \mathrm{rk}(y_1)) & \text{and} \\ \min(\mathrm{rk}(x_0), \mathrm{rk}(y_0)) < \min(\mathrm{rk}(x_1), \mathrm{rk}(y_1)). & \end{cases}$$

Let $\rho : M \times M \to \mathrm{Ord} \cap M$ be the corresponding rank function (which is a class function in M). Then, by induction on the stratification ρ, there is a unique class function $F : M \times M \to \mathcal{P}(\mathbb{P})^4$ defined in M such that

$$F(x, y) = \big\{(D_1, D_2, D_3, D_4) \in \mathcal{P}(\mathbb{P})^4 \,|\, p \in D_1 \iff p \Vdash \underline{x} \in \underline{y}, \text{ etc}\big\}.$$

So, using F, we see that (1), (2), (3), (4) are class relations in M.

Having defined (1), (2), (3), (4), we extend the forcing relation to arbitrary formulas with parameters by induction on their construction over literals "$x \in y$", "$x \notin y$", "$x = y$", "$x \neq y$".

If $\phi(a_1, ..., a_n), \psi(a_1, ..., a_n), \sigma(y, a_1, ..., a_n)$ are formulas of the language of set theory with parameters $a_1, ..., a_n \in M$, let

(5) $p \Vdash \big(\phi(\underline{\vec{a}}) \vee \psi(\underline{\vec{a}})\big) \iff p \Vdash \phi(\underline{\vec{a}})$ or $p \Vdash \psi(\underline{\vec{a}})$
(6) $p \Vdash \exists y \, \sigma(\underline{\vec{a}}) \iff \exists b, \; p \Vdash \sigma(\underline{b}, \underline{\vec{a}})$
(7) $p \Vdash \neg\phi(\underline{\vec{a}}) \iff \forall q \leq p, \; q \nVdash \phi(\underline{\vec{a}})$.

WARNING

Contrary to classical logic, for the forcing relation, we consider formulas as being built up from literals and NOT from atomic formulas.
E.g.

$$p \Vdash \neg(\underline{x} \notin \underline{y}) \iff \forall q \leq p, \; q \nVdash \underline{x} \notin \underline{y} \iff \forall q \leq p \, \exists r \leq q, \; r \Vdash \underline{x} \in \underline{y},$$

which is weaker than $p \Vdash \underline{x} \in \underline{y}$.

Lemma 7.22. Suppose $p \leq q$. Then

$$q \Vdash \phi(\underline{\vec{a}}) \implies p \Vdash \phi(\underline{\vec{a}}).$$

Proof.

Suppose first that ϕ is one of the formulas $a \in b$, $a \neq b$, $a \notin b$, or $a = b$. Then the result is proved on the stratification ρ. Now for any other formula ϕ, the result is proved by induction on the construction of ϕ using (5), (6), and (7).

□

Formulas involving $\forall$ and $\wedge$ are introduced by defining

$$\forall x\, \phi \iff \neg\exists x\, \neg\phi$$
$$\phi \wedge \psi \iff \neg(\neg\phi \vee \neg\psi).$$

Thus,

(8) $p \Vdash \forall y\, \sigma(y, \underline{\vec{a}}) \iff p \Vdash \neg\exists y\, \neg\sigma(y, \underline{\vec{a}}) \iff \forall q \leq p\, \forall b,\ q \nVdash \neg\sigma(\underline{b}, \underline{\vec{a}})$
$\iff \forall b\, \forall q \leq p\, \exists r \leq q,\ r \Vdash \sigma(\underline{b}, \underline{\vec{a}})$.

(9) $p \Vdash \big(\phi(\underline{\vec{a}}) \wedge \psi(\underline{\vec{a}})\big) \iff \forall q \leq p\, \big(\exists r \leq q,\ r \Vdash \phi(\underline{\vec{a}})\ \wedge\ \exists s \leq q,\ s \Vdash \psi(\underline{\vec{a}})\big)$.

Lemma 7.23. For all $p \in \mathbb{P}$ and $a, b \in M$ with $(a, p) \in b$, $p \Vdash \underline{a} = \underline{a}$ and $p \Vdash \underline{a} \in \underline{b}$.

Proof.

We show this by induction on $\mathrm{rk}(a)$. So suppose $q \Vdash \underline{c} = \underline{c}$ for all $q \in \mathbb{P}$ and $c \in M$ with $\mathrm{rk}(c) < \mathrm{rk}(a)$. Then

$$\Big[p \Vdash \underline{a} = \underline{a}\Big] \iff \Big[\neg\exists q \leq p,\ q \Vdash \underline{a} \neq \underline{a}\Big] \iff \Big[\neg\exists q \leq p\, \exists r \geq q\, \exists c,\ \big((c, r) \in a \wedge q \Vdash \underline{c} \notin \underline{a}\big)\Big]$$

Now, suppose towards a contradiction that $q \leq p$, $r \geq q$, and $(c, r) \in a$ with $q \Vdash \underline{c} \notin \underline{a}$. Then $\mathrm{rk}(c) < \mathrm{rk}(a)$, and so $q \Vdash \underline{c} = \underline{c}$, hence by (1), $q \Vdash \underline{c} \in \underline{a}$. By (3), this contradicts $q \Vdash \underline{c} \notin \underline{a}$. Therefore, $p \Vdash \underline{a} = \underline{a}$. To see that $p \Vdash \underline{a} \in \underline{b}$, we apply (1) directly.

□

7.5. **The Truth Lemma.**

We begin this subsection with a preliminary lemma which is not the truth lemma.

Lemma 7.24. For any formula $\phi(\vec{a})$ with parameters $a_1, ..., a_n \in M$, there is $p \in G$ such that

$$p \Vdash \phi(\underline{\vec{a}}) \qquad \text{or} \qquad p \Vdash \neg\phi(\underline{\vec{a}}).$$

Proof.

Note that by (7), the set

$$D := \big\{p \in \mathbb{P} \mid p \Vdash \phi(\underline{\vec{a}})\ \vee\ p \Vdash \neg\phi(\underline{\vec{a}})\big\}$$

is dense and also belongs to M. So $G \cap D \neq \varnothing$.

□

Lemma 7.25. If $a, b \in M$, we have

$$\Phi(a) \in \Phi(b) \iff \exists p \in G,\ p \Vdash \underline{a} \in \underline{b}$$
$$\Phi(a) \neq \Phi(b) \iff \exists p \in G,\ p \Vdash \underline{a} \neq \underline{b}.$$

Proof.

The proof is by induction on $\rho(a,b)$. So assume the result holds for all $c,d \in M$ with $\rho(c,d) < \rho(a,b)$.

Suppose $\Phi(a) \in \Phi(b)$ and pick $c \in M$, $q \in G$ such that $\Phi(a) = \Phi(c)$, $(c,q) \in b$. Choose also, by the previous lemma, some $p \le q$, $p \in G$, such that either $p \Vdash \underline{c} \neq \underline{a}$ or $p \Vdash \neg(\underline{c} \neq \underline{a})$, i.e., $p \Vdash \underline{c} = \underline{a}$.

Since $\rho(c,a) < \rho(a,b)$, we have

$$p \Vdash \underline{c} \neq \underline{a} \qquad \Longrightarrow \qquad \Phi(c) \neq \Phi(a)$$

and we must therefore have $p \Vdash \underline{c} = \underline{a}$. By (1), it follows that $p \Vdash \underline{a} \in \underline{b}$.

Conversely, if $p \Vdash \underline{a} \in \underline{b}$ for some $p \in G$, then by (1), there is $r \ge p$ and $c \in M$ such that $(c,r) \in b$ and $p \Vdash \underline{c} = \underline{a}$. Again, $\rho(c,a) < \rho(a,b)$, so if $\Phi(c) \neq \Phi(a)$, then by the induction hypothesis, there is $q \le p$, $q \in G$ such that $q \Vdash \underline{c} \neq \underline{a}$, contradicting $p \Vdash \underline{c} = \underline{a}$. Thus, $\Phi(a) = \Phi(c) \in \Phi(b)$.

The proof for the second equivalence is similar.

□

Lemma 7.26 (Truth Lemma)**.** Suppose $\psi(\vec{x})$ is a formula without parameters and $a_1, ..., a_n \in M$. Then

$$\psi^{M[G]}\big(\Phi(a_1), ..., \Phi(a_n)\big) \qquad \Longleftrightarrow \qquad \exists p \in G,\ \ p \Vdash \psi(\underline{\vec{a}})$$ [a]

Proof.

While the previous lemma is proved by induction on ρ, this is proved by induction on the construction of ψ over literals "$x \in y$", "$x \notin y$", "$x = y$", "$x \neq y$".

Now if ψ is either $x_1 \in x_2$ or $x_1 \neq x_2$, then the result is proved by the previous lemma. Suppose ψ is $x_1 \notin x_2$. Then for $a_1, a_2 \in M$,

$$\exists p \in G,\ p \Vdash \underline{a_1} \notin \underline{a_2} \qquad \Longleftrightarrow \qquad \exists p \in G\ \forall q \le p,\ q \not\Vdash \underline{a_1} \in \underline{a_2}.$$

So if $\Phi(a_1) \in \Phi(a_2)$, there is $r \in G$ such that $r \Vdash \underline{a_1} \in \underline{a_2}$, and so for any $p \in G$, there is $q \le r$, $q \le p$ such that $q \Vdash \underline{a_1} \in \underline{a_2}$, contradicting that $\exists p \in G,\ p \Vdash \underline{a_1} \notin \underline{a_2}$.

Thus,

$$\begin{aligned} &\Big[\exists p \in G,\ p \Vdash \underline{a_1} \notin \underline{a_2}\Big] \implies \Big[\Phi(a_1) \notin \Phi(a_2)\Big] \\ \implies &\Big[\forall p \in G,\ p \not\Vdash \underline{a_1} \in \underline{a_2}\Big] \implies \Big[\exists p \in G,\ p \Vdash \underline{a_1} \notin \underline{a_2}\Big] \end{aligned}$$

Similarly for $x_1 = x_2$ and $\psi = \neg\phi$.

Suppose $\psi(\vec{x}) = \exists y\, \sigma(y, \vec{x})$ and the lemma holds for σ. Then for any $a_1, ..., a_n \in M$,

$$\begin{aligned} &\psi^{M[G]}\big(\Phi(a_1), ..., \Phi(a_n)\big) \iff \exists b \in M,\ \sigma^{M[G]}\big(\Phi(b), \Phi(a_1), ..., \Phi(a_n)\big) \\ \iff &\exists b \in M\ \exists p \in G,\ p \Vdash \sigma(\underline{b}, \underline{\vec{a}}) \iff \exists p \in G,\ p \Vdash \exists y\ \sigma(y, \underline{\vec{a}}) \iff \exists p \in G,\ p \Vdash \psi(\underline{\vec{a}}). \end{aligned}$$

The case of $\vee$ is easy.

□

[a] *TN:* If you don't immediately see why this lemma is so beautiful and powerful for us, then stop, go back to the beginning of this section, and reread until you realize how useful this is about to become.

Notation:
Instead of writing $p \Vdash \psi(\underline{\widehat{a_1}}, ..., \underline{\widehat{a_n}}, \underline{b_1}, ..., \underline{b_k})$ for $\vec{a}, \vec{b} \in M$, we simplify notation and write $p \Vdash \psi(a_1, ..., a_n, \underline{b_1}, ..., \underline{b_k})$.

So, "un-underlined" a's refer to themselves in the generic extension $M[G]$, while underlined b's refer to $\Phi(b)$ in $M[G]$. So the truth lemma reads:

For any formula $\psi(x_1, ..., x_n, y_1, ..., y_k)$ and any $a_1, ..., a_n,\ b_1, ..., b_k \in M$:

$$\psi^{M[G]}\big(\vec{a}, \Phi(b_1), ..., \Phi(b_k)\big) \qquad \Longleftrightarrow \qquad \exists p \in G,\ p \Vdash \psi(\vec{a}, \underline{\vec{b}}).$$

Theorem 7.27. $M[G]$ satisfies ZF.

Proof.

We need only check the power set axiom and replacement scheme.

For the power set axiom, suppose $\Phi(a) \in M[G]$ for some $a \in M$. We let

$$a' := \big\{(x,p) \in \mathrm{cl}(a) \times \mathbb{P} \bigm| \exists q \geq p,\ (x,q) \in a\big\} \in M$$

and $b = \mathcal{P}^M(a') \times \mathbb{P} \in M$.

We claim that $\Phi(b) = \mathcal{P}^{M[G]}\big(\Phi(a)\big)$. To see this, suppose $\Phi(u) \in \Phi(b)$ and assume without loss of generality that there is $r \in G$ such that $(u,r) \in b$. Then $u \in a'$, and so for any $(x,p) \in u$, there is $q \geq p$ with $(x,p) \in a$. Since G is upwards closed, this means that if $\Phi(x) \in \Phi(u)$, also $\Phi(x) \in \Phi(a)$, hence $\Phi(u) \subseteq \Phi(a)$ and $\Phi(b) \subseteq \mathcal{P}^{M[G]}\big(\Phi(a)\big)$.

Conversely, if $\Phi(u) \subseteq \Phi(a)$, we need to find $v \subseteq a'$ such that $\Phi(v) = \Phi(u)$. It follows then that $(v,r) \in b$ for any $r \in G$, and so $\Phi(u) = \Phi(v) \in \Phi(b)$. Set

$$v := \big\{(x,p) \in a' \mid p \Vdash \underline{x} \in \underline{u}\big\}.$$

Then for $\Phi(v) = \Phi(u)$, suppose that $\Phi(x) \in \Phi(v)$ for some $(x,p) \in a'$, $p \Vdash \underline{x} \in \underline{u}$ and $p \in G$. Then by the truth lemma, also $\Phi(x) \in \Phi(u)$, so $\Phi(v) \subseteq \Phi(u)$. Conversely, if $\Phi(x) \in \Phi(u) \subseteq \Phi(a)$ for some $(x,q) \in a$ with $q \in G$, there is by the truth lemma some $r \in G$ such that $r \Vdash \underline{x} \in \underline{u}$. Now, if $p \in G$, $p \leq r, q$, then $p \Vdash \underline{x} \in \underline{u}$ and so $(x,p) \in v$ and thus also $\Phi(x) \in \Phi(v)$. So $\Phi(u) \subseteq \Phi(v)$. ■

For replacement, suppose $\Phi(a) \in M[G]$ and $\psi\big(x, y, \Phi(a_1), ..., \Phi(a_n)\big)$ is a formula that in $M[G]$ defines a class function of the variable x. Let

$$B := \Big\{\Phi(y) \in M[G] \Bigm| \exists \Phi(x) \in \Phi(a),\ \psi^{M[G]}\big(\Phi(x), \Phi(y), \Phi(a_1), ..., \Phi(a_n)\big)\Big\}.$$

We need to find $b \in M$ such that $\Phi(b) = B$. Now, for all $x \in M$ and $p \in \mathbb{P}$,

$$F(x,p) = \big\{y \in M \bigm| y \text{ has minimal rank such that } p \Vdash \psi(\underline{x}, \underline{y}, \underline{a_1}, ..., \underline{a_n})\big\}.$$

So F is a class function in M and thus, by applying replacement in M, we see that the following is a set in M:

$$b = \big\{(y,p) \in M \times \mathbb{P} \bigm| \exists x\, \exists q \geq p,\ (x,q) \in a \wedge y \in F(x,p)\big\}.$$

Now, if $\Phi(y) \in \Phi(b)$ for some $p \in G$ such that $(y,p) \in b$, we find x and $q \geq p$ such that $(x,q) \in a$ and $y \in F(x,p)$, i.e., $p \Vdash \psi(\underline{x}, \underline{y}, \underline{\vec{a}})$. By the truth lemma, $\psi^{M[G]}\big(\Phi(x), \Phi(y), \Phi(a_1), ..., \Phi(a_n)\big)$ and so $\Phi(y) \in B$.

Conversely, if $\Phi(y) \in B$, find $\Phi(x) \in \Phi(a)$ such that $\psi^M\big(\Phi(x), \Phi(y), \Phi(a_1), ..., \Phi(a_n)\big)$ and assume without loss of generality that $(x,q) \in a$ for some $q \in G$. By the truth lemma, there is

$p \in G$ such that $p \Vdash \psi(\underline{x}, \underline{y}, \vec{\underline{a}})$, and we can assume that $p \leq q$. Pick $y_0 \in F(x, p) \neq \varnothing$. Then $p \Vdash \psi(\underline{x}, \underline{y_0}, \vec{\underline{a}})$ and so, by the truth lemma,[a]

$$\psi^{M[G]}\big(\Phi(x), \Phi(y_0), \Phi(a_1), ..., \Phi(a_n)\big).$$

Since ψ defined a class function in $M[G]$, this implies that $\Phi(y) = \Phi(y_0)$. As also $(y_0, p) \in b$, we have $\Phi(y) = \Phi(y_0) \in \Phi(b)$. Therefore, $B \subseteq \Phi(b)$ and so $B = \Phi(b)$.

□

Theorem 7.28. *M and $M[G]$ contain the same ordinals.*

Proof.

Recall that as $M, M[G]$ satisfy ZF and are transitive, the notion of ordinal is absolute for both. So the result follows from $M \cap \mathrm{Ord} = M[G] \cap \mathrm{Ord}$.

□

Theorem 7.29 (Proof omitted)**.** *$M[G]$ is the smallest transitive set model of ZF such that $M \subseteq M[G]$ and $G \in M[G]$.*

Theorem 7.30. [b] *If M satisfies AC, then so does $M[G]$.*

Proof.

Suppose $a[G] \in M[G]$. It suffices to find in $M[G]$ a surjection from an ordinal onto a superset of $a[G]$. This induces a well-ordering on $a[G]$ by defining an injection from $a[G]$ to this ordinal by mapping $a' \in a[G]$ to the least pre-image of a' under this surjection. And finally, an injection from a set to an ordinal obviously induces a well-ordering on that set.

So by AC in M, let $f : \alpha \twoheadrightarrow \mathrm{cl}(a)$ be a surjection from an ordinal α onto the transitive closure of a. It should be immediately obvious that $a[G] \in \mathrm{cl}(a)[G]$. Using the image of f under our collapse map should work. This is because being a function and being an ordinal are absolute, so the image of f under the collapse map will still be a function with domain an ordinal that maps onto a superset of $a[G]$.

□

7.6. **Consistency of ZFC + ¬CH.**

Recall that not only is CH consistent with ZFC, but GCH is consistent with ZFC. During this final section, we will use the preceding useful lemmas to force a new model which forces $|2^\omega| > \aleph_1$.
Here's the general outline of our steps:

- We look at the poset $\big[\aleph_7 \times \omega \to 2\big]^{<\omega}$ of finite partial functions from $\aleph_7 \times \omega$ to 2 ordered by $a \leq b \iff a \supseteq b$ (really, we could also use $\aleph_{23}$, $\aleph_{35324}$, or $\aleph_2$).
- We present Professor Tserunyan's graph theoretic proof of the Δ-system lemma.
- We show that every antichain in this poset is countable (known as the countable chain condition) using the Δ-system lemma.
- We show that our poset, since it satisfies the countable chain condition, preserves cardinals under the collapse map for a generic subset of the poset.
- We deduce that our collapsed model has an injection from $\aleph_7 \to 2^\omega$.

[a] *TN:* Fun fact, the phrase "by the truth lemma" appears 5 times in this one proof.

[b] *TN:* During the first line of the proof, Rosendal's notes cut off on the pdf on his website. I am indebted to Ms. Jenna Zomback who has provided a hand-written copy of the remainder of the notes as given during the same class I took. I will be using her notes as a basis for the rest of the material, but as her notes are flawless, any glraing errorrs would of course be my sole responsibility.

First, one should verify that the poset we've given is actually a poset and that for some transitive model M satisfying ZFC, that the given poset is genuinely in M.

Also, if $G \subseteq \mathbb{P}$ is a $\mathbb{P}$-generic ultrafilter[a], set $g := \bigcup G$. Then, we want to show that g is a surjective function onto the codomain, and, setting for each $\alpha \in \aleph_7$, $g_\alpha := g(\alpha, -) : \omega \to 2$, viewing g as a function $g' : \aleph_7 \to 2^\omega$ given by $\alpha \longmapsto g_\alpha$, g' is injective. These tasks are left as exercises.

What follows is Professor Tserunyan's graph theoretic proof of the Δ-system lemma.

Definition 7.31. A set (think family of sets) $\mathcal{F}$ is called a Δ-system is there is a set r (called the root) such that for any distinct $a, b \in \mathcal{F}$, $a \cap b = r$.

Definition 7.32. A graph is locally countable if every vertex has at most countably many neighbors.

A quick observation is that a graph is locally countable if and only if each of its connected components is countable.

For a graph G on a set V, a subset $I \subseteq V$ is called independent if no two vertices in it are adjacent.

Lemma 7.33 (Uses AC)**.** A locally countable graph G on an uncountable set V admits an independent set $I \subseteq V$ with $|I| = |V|$.

Proof.

Each connected component is countable, so their collection has cardinality $|V|$. Choosing a point from each connected component gives a desired independent set.

□

The intersection graph on any family V of sets is defined by putting an edge between distinct $u, v \in V$ exactly when $u \cap v \neq \varnothing$. Call a family V of sets point-locally countable if for each $a \in \bigcup V$, the set $V_a := \{v \in V \mid a \in v\}$ is countable.

Lemma 7.34 (Uses Countable-AC)**.** The intersection graph on any point-locally countable family V of countable sets is locally countable.

Proof.

For each $v \in V$, each $a \in v$ is contained in at most countably many other vertices in V. Since v is countable and countable union of countable sets is countable, it intersects at most countably many other sets in V.

□

The previous two lemmas give us the following:

Proposition 7.35 (Uses AC)**.** The intersection graph on a point-locally countable uncountable family V of countable sets admits a pairwise disjoint collection $I \subseteq V$ with $|I| = |V|$.

[a] *TN:* A filter is simply a non-empty subset of a poset that is closed upwards and any two elements are compatible; an ultrafilter would be a maximal filter; thus, we've already been working a lot with filters and ultrafilters through $\mathbb{P}$-generic subsets.

Corollary 7.36 (The Δ-system Lemma, Uses AC)**.** Every uncountable family V of finite sets contains an uncountable Δ-system.

Proof.

By shrinking V using uncountable-to-countable Pigeonhole Principle (involves AC), we may assume that for some $n \in \mathbb{N}$, all sets in V have the same size n and we prove the statement by induction on n. The base case $n = 1$ is trivial, so suppose the statement is true for $n - 1 \geq 1$ and let V be an uncountable family of sets of size n.

If V is point-locally countable, then Proposition 7.35 finishes the proof, so suppose that V is not point-locally countable and let $a \in \bigcup V$ be such that the set V_a is uncountable. Removing a from the sets in V_a, applying induction, and putting a back yields an uncountable Δ-system $V'_a \subseteq V_a$.

□

Lemma 7.37. [a] For sets A, B with $2 \leq |B| \leq \aleph_0$, the poset $\left([A \to B]^{<\omega}, \supseteq\right)$ has the countable chain condition. (c.c.c.)

Proof.

Suppose not. Let F be an uncountable antichain, and let $\mathrm{dom}(F)$[b] $:= \bigcup_{f \in F} \{\mathrm{dom}(f)\}$.

Claim: $|\,\mathrm{dom}(F)| > \omega$.

Proof.

Suppose not. Then for some element $f_0 \in \mathrm{dom}(F)$, there are uncountably many $f \in F$ such that $\mathrm{dom}(f) = f_0$, lest the countably many $\mathrm{dom}(f)$'s have only countably many preimages, which implies that F would be a countable union of countable sets and hence countable. So, take such an f_0.

Recall that f_0 must be finite. Then, since B is countable, we conclude that there are uncountably many functions from a finite set to a countable set, which is a contradiction.

■

From this, we can now apply the Δ-system lemma and let $\Delta \subseteq \mathrm{dom}(F)$ be such a system with $r = a \cap b$ for some $a, b \in \Delta$ where $a \neq b$. Note that $r \neq \varnothing$ (since $\bigcap \mathrm{dom}(F) \neq \varnothing$ because 2 functions with completely different domains *are* compatible).

Now, if $f_a,\ f_b \in F$ have associated $a, b \in \Delta$, then $f_a\restriction_r \neq f_b\restriction_r$. If they were in fact equal, then, since $(a \setminus r) \cap (b \setminus r) = \varnothing$, f_a and f_b would be compatible. So, any $a, b \in \Delta$ with $a \neq b$ implies any associated functions disagree on r which implies we have uncountably many functions from r to B, which is a contradiction.

□

Lemma 7.38. Let M be a transitive model of ZFC. Let $(\mathbb{P}, \leq)$ be a c.c.c. (in M) poset. Then forcing with $\mathbb{P}$ preserves cardinals, i.e. $\forall$ $\mathbb{P}$-generic over M $G \subseteq \mathbb{P}$ and for any ordinal $\kappa \in M$,

$$M \models \kappa \text{ is a cardinal} \iff M[G] \models \kappa \text{ is a cardinal.}$$

Proof.

Suppose towards a contradiction that $\exists \lambda < \kappa$ and $f[G] \in M[G]$ with $f[G] : \lambda \twoheadrightarrow \kappa$ whereas $M \models$ "κ is a cardinal."

---

[a] *TN:* This was originally assigned as a homework exercise, so the following proof is my own.

[b] *TN:* I recognize that this is not a bona fide domain, but this notation was meant to remind myself what I was looking at.

By the truth lemma, there is a $p_0 \in G$ such that $p_0 \Vdash$ "f is a surjection from $\widehat{\lambda} \to \widehat{\kappa}$".
Fix $\alpha \in \lambda$. Then

$$X_\alpha := \left\{\beta \in \kappa \mid \exists p \le p_0,\ p \Vdash \text{“}f(\widehat{\alpha}) = \widehat{\beta}\text{”}\right\}.$$

Claim: $(X_\alpha \text{ is countable})^M$.

Proof.

By AC in M, get $\beta \longmapsto p_\beta$ from X_α to $\mathbb{P}$ such that $p_\beta \subseteq p_0$ and $p_\beta \Vdash \beta \in X_\alpha$. Let A be the image of this function, i.e. $A := \{p_\beta \mid \beta \in X_\alpha\} \in M$. This is an antichain. If p_{β_0} and p_{β_1} are compatible, then they agree on $f(\widehat{\alpha})$ by the truth lemma. Then, by c.c.c., $M \models$ "A is countable".

■

Then, let $X := \{X_\alpha \mid \alpha \in \lambda\}$. Then $|X|^M \le \lambda \otimes \aleph_0 < \kappa$, in particular, $X \subsetneq \kappa$, so f is not surjective (by the truth lemma).

□

By this lemma, we see that $\aleph_7$ in M is $\aleph_7$ in $M[G]$. Thus, from all of this, we conclude that in $M[G]$, we have a function $(g' : \aleph_7 \hookrightarrow 2^\omega)^{M[G]}$, a genuine injection from $\aleph_7$ to 2^ω, which implies $|2^\omega| \ge \aleph_7 > \aleph_1 \implies M[G] \models ZFC + \neg CH$.

Q.E.D.