Math 280A Fall 2009

1. Axioms of Set Theory

Let V be the collection of all sets and $\in$ be a membership relation.

We consider $(V, \in)$ as a mathematical structure. Analogy: A group is a mathematical structure $(G, \cdot, ^{-1}, 1)$.

As the properties of group operations are given by group axioms, in the case of set theory, the properties of $\in$ will be given by axioms of set theory. These axioms will mimic an intuitive understanding of membership. Among other things, the axioms of set theory give us rules how to construct sets. This requires some care.

1.1 Example: Russell's Paradox

$E = \{x | x \notin x\}$. Then E is not a set. Why? If E is a set, then $E \in E$ or $E \notin E$. We get a contradiction under either assumption. More exactly:

Either: $E \in E$. In this case E is an element of $\{x | x \notin x\}$. So E must satisfy $x \notin x$. Hence $E \notin E$. Contradiction.

Or else: $E \notin E$. In this case E is not an element of $\{x | x \notin x\}$. So E does not satisfy the property $x \notin x$. Therefore $E \in E$.

To avoid paradoxes of these kinds, we have to set some rigorous framework as a background. The framework is the rigorous language. Our language will *not* be contained in V.

1.2 Definition: Language of Set Theory

Language consists of *logical symbols*:

Variables $v_0, v_1, ...$ of which there are countably many.

Logical conjunctives: $\neg$ - negation, $\wedge$ - conjunction (Professor Zeman writes &).

Quantifier: $\exists$ - existential quantifier.

Symbol for equality: $=$; Parentheses: (,) .

Special symbol: $\in$ denotes membership relation.

Strictly speaking, we should differ between symbols and relations. e.g. we should write $=, \in$ for actual relations (abstractly), and $\dot{=}, \dot{\in}$ for symbols denoting the relations (the everyday language). In practice, we "abuse notation" and do not formally distinguish between those– it will be clear from the context whether we consider the relation or the symbol.

1.3 Definition: Formula

The notion of formula is defined inductively:

—The expressions $v_i \in v_j$, $v_i = v_j$ (e.g. should be $\dot{\in}, \dot{=}$) are formulae. These are called atomic formulae.

—If φ is a formula, then also $\neg\varphi$ is a formula.
—If φ, ψ are formulae, then $(\varphi \wedge \psi)$ is a formula.
—If φ is a formula then $(\exists v_j)\varphi$ is a formula. And nothing else is a formula.

The remaining obvious symbols are viewed as abbreviations:
$\varphi \vee \psi : \neg(\neg\varphi \wedge \neg\psi)$
$\varphi \rightarrow \psi : \neg\varphi \vee \psi$
$\varphi \leftrightarrow \psi : (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$
$(\forall v_j)\varphi : \neg(\exists v_j)\neg\varphi$
So any formula can be viewed as a sequence of symbols.

1.4 Definition: Free occurrence

If φ is a formula and $k <$length(φ) we define what it means that k is a free occurrence of v_i in φ.

(i) k is an *occurrence* of v_i in φ iff the k^{th} element of φ is v_i (we start numbering from 0).
Example:

$v_i \in v_j$
0 1 2

(ii) If φ is atomic then any occurrence of v_i in φ is free. If φ is of the form $\neg\varphi'$ and k is a free occurrence of v_i in φ', then $k+1$ is a free ccurence of v_i in φ.

(iii) if φ is of the form $\varphi' \wedge \varphi''$ and

—k is a free occurrence of v_i in φ', then $k+1$ is a free occurrence of v_i in φ (parentheses count).
—k is a free occurrence of v_i in φ'' then $k+$length$(\varphi')+2$ is free occurrence of v_i in φ.
Example:
$v_i \in v_j$
0 1 2
$v_i = v_l$
0 1 2
$(v_i \in v_j \wedge v_i = v_l)$
0 1 2 3 4 5 6 7 8

(iv) If φ is of the form $(\exists v_j)\varphi'$, and $i \neq j$ and k is a free occurrence of v_i in φ' then $k+4$ is a free occurrence of v_i in φ. No other occurrence of v_i in φ is free.

We say that v_i has a *free occurrence* in φ iff there is some $k <$length(φ) that is a free occurrence of v_i in φ. (Has an *occurrence* k of v_i in φ iff it is not free).

Remarks

(a) v_i may have more free occurrences in φ.

(b) Intuitively, k is a free occurrence of v_i in φ if that occurrence of v_i in φ is not under influence of any quantifier.

(c) v_i may have both free and bound occurrence in φ. Example:
$\varphi : v_i = v_j \wedge(\exists v_i)(v_i = v_k)$. It is free in its first appearance, and bound in the last two.

1.6 Definition: Sentence

A formula σ is a *sentence* iff no variable has a free occurrence in σ.

1.7 Remark

If $\varphi(v_0, ..., v_n)$ is a formula and $v_0, ..., v_n$ have free occurrences in φ, then the truth of φ depends on what objects we plug in for these variables. Example:

$v_0 \in v_1$: If $v_0 \mapsto \varnothing$, $v_1 \mapsto \{\varnothing\}$ then the formula is true. But if $v_0 \mapsto \varnothing$ and $v_1 \mapsto \varnothing$ then the formula is false.

But if we had a sentence σ, then we don't need any evaluation of variables to determine the truth of σ: $(\forall v_1)(\exists v_0)(v_0 \in v_1)$ is false, while $(\forall v_1)(\exists v_0)(v_1 \in v_0)$ is true (after we have defined our axioms!).

The point of introducing rigorous language of set theory: it puts restrictions on what we can express, and enables us to classify the "complexity" of notions. This way it makes sense that we avoid all known paradoxes.

1.8 Zermelo-Fraenkel Axioms of Set Theory:

0. Existence

It tells us that the universe V is nonempty. (Actually, this axiom is superfluous, but it is added for completeness). It is a logical axiom- something which is always true:

$$(\exists x)(x = x).$$

1. Extensionality

Expresses the basic property of sets: two sets are equal iff they contain the same elements.

$$(\forall x)(\forall y)[x = y \leftrightarrow (\forall z)(z \in x \leftrightarrow z \in y)]$$

2. Foundation

Intuitively, it guarantees that the following situations never happen: $x \in x$, $x \in y \in x$, $x \in y_0 \in ... \in y_n \in x$ etc. or even $x_0 \ni x_1 \ni ... \ni x_n \ni ...$

The way we express all of this in our language: we postulate that every nonempty set has an $\in$-minimal element. So if A is a set and $a \in A$, then a is an $\in$-minimal element of A iff no elements of a are in A. i.e. $a \cap A = \varnothing$.

If $x \in x$ then the set $\{x\}$ has no $\in$-minimal element. If $x \in y \in x$ then the set $\{x, y\}$ has no $\in$-minimal element and so on. In the rigorous language:

$$(\forall x)[(\exists y)(y \in x) \longrightarrow (\exists y)(y \in x \wedge (\forall z)(z \in y \rightarrow z \notin x))]$$

3. Pairing

For any two sets x, y there is a set whose elements are just x, y and nothing else. In mathematics, we denote this set by $\{x, y\}$. Rigorously:

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

4. Union

This says: if x is a collection of sets then there is a set whose elements are precisely all elements of sets in x. We call this set the union of x and denote it by $\bigcup x$. Example: if $x = \{u, v\}$, then $\bigcup x = u \cup v$. Now in our language :

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

Without the following, the above axioms would be equivalent to arithmetic.

5. Infinity

There is an infinite set. We have to express this using a finite sentence. How: we introduce the operation S that to each set assigns the set $S(x) = x \cup \{x\}$. Also: $\varnothing \in V$ (we will see later). $y = \varnothing$ iff $(\forall z)(z \in y \rightarrow z \neq z)$. Now, the axiom:

$$(\exists x)[\varnothing \in x \wedge (\forall z)(z \in x \rightarrow S(z) \in x)]$$

So x contains: $\varnothing$, $\varnothing \cup \{\varnothing\} = \{\varnothing\}$, $\{\varnothing \cup \{\varnothing\}\} = \{\varnothing, \{\varnothing\}\}$,...

6. Separation Schema

It says: if a is a set and $P(v)$ is a property expressable in our language, then there is a set $\{x \in a | P(x)\}$. The important point: recall Russell's Paradox: $\{x | x \notin x\}$ is not a set. But: if a is any set, then $\{x \in a | x \notin x\}$ is a set.

It defines the following: if $\varphi(x, y_1, ..., y_n)$ is a formula and $a, a_1, ..., a_n$ are sets then $\{z \in a | \varphi(z, y_1, ..., y_n)\}$ is a set. Rigorously, the schema consists of all formulae of the form:

$$(*)\ (\forall x)(\forall y_1)...(\forall y_n)(\exists u)(\forall z)[z \in u \longleftrightarrow (z \in x \wedge \varphi(z, y_1, ..., y_n))]$$

Note: must state this syntactically- state variables instead of sets.

So for each formula φ, the schema contains the formula $(*)$. So the separation schema consists of *infinite* lists of formulae. Note: we may not quantify formulae. This is a limitation of the language. This forces us to list infinitely many things. Hence the descriptive word "schema."

Example: If we let $\varphi(z)$ be the formula $z \neq z$ then the Separation Schema says:
$(\forall x)(\exists u)(\forall z)(z \in u \longleftrightarrow (z \in x \wedge z \neq z))$.
(So this gives us a set u and we will write it as mathematicians usually do: $\{z \in x | z \neq z\} (= \varnothing)$).
In practice, the Separation Schema is used informally, so the above example would look as follows:

By the existence axiom there is some set a. By separation, $\{z \in a | z \neq z\}$ is a set. If we picked another set b, then again $\{z \in b | z \neq z\}$ is a set. By extensionality: $\varnothing_a = \{z \in a | z \neq z\} = \{z \in b | z \neq z\} = \varnothing_b$. (Just for fun: to see that $\varnothing_a = \varnothing_b$ we have to show $(\forall z)(z \in \varnothing_a \leftrightarrow z \in \varnothing_b)$. The inside two statements are trivially false so the inside is trivially true). Hence, this procedure *uniquely* determines a set $\varnothing$ which we call the *empty set*. It's uniquely determined by the property:

$x \in \varnothing \leftrightarrow (\forall z)(z \notin x) \leftrightarrow (\forall z)(z \in x \rightarrow z \neq z)$.

(So we can use this $\varnothing$ in Axiom 5)

7. Replacement Schema

Informally: this schema expresses the following: if $F : a \rightarrow V$ is a function defined by a formula and a is a set, then there is a set b such that all values of F are in b. Toward rigorous formulation: the quantifier of the form $\exists! x$ means "there is exactly one x." $(\exists! x)\varphi(x)$ is an abbreviation for $(\exists x)\varphi(x) \wedge (\forall y)(\forall z)(\varphi(y) \wedge \varphi(z) \rightarrow y = z)$. The Replacement Schema consists of all formulae of the form:

$$(\forall u)(\exists v)(\forall y_1)...(\forall y_n)[(\forall x)(x \in u \rightarrow (\exists! y)\varphi(x, y, y_1, ..., y_n)) \longrightarrow \\ (\forall x)(\exists y)(x \in u \rightarrow y \in v \wedge \varphi(x, y, y_1, ..., y_n))]$$

The first part tells us that φ defines a function that to each x assigns y and this assignment depends on parameters $y_1, ..., y_n$. The last part says that for each x the corresponding value y is in v.

8. Power Set

The axiom expresses that to each set x, the collection of all subsets of x is contained in a set. $z \subseteq x$ is an abbreviation for $(\forall u)(u \in z \rightarrow u \in x)$.

$$(\forall x)(\exists y)(\forall z)(z \subseteq x \rightarrow z \in y)$$

9. Axiom of Choice (AC)

It says that if a is a collection of nonempty mutually disjoint sets, then we can find a set C that has exactly one element in common with every set from a. So this gives us a "choice function:" $A \in a \mapsto$ the unique object in $A \cap C$

(picture)

$(\forall x)[(\exists u)(u \in x) \wedge (\forall u, v)(u, v \in x \rightarrow \neg(\exists z)(z \in u \wedge z \in v)) \wedge (\forall u)(u \in x \rightarrow (\exists z)(z \in u))] \longrightarrow$
$(\exists y)(\forall u)(u \in x \rightarrow (\exists! z)(z \in u \wedge z \in y))$

1.9 Remarks

(a) Axioms 0-6, and 8 constitute the so-called Zermelo axiomatic system, Z.
(b) Z + Replacement is called the Zermelo–Fraenkel axiomatic system and is denoted by ZF.
(c) ZC=Z+AC and ZFC= ZF+AC

1.10 Definition: Class

Given a formula $\varphi(x)$ with the only free variable x, the collection of all $a \in V$ such that $\varphi(a)$ holds is informally denoted by $\{x|\varphi(x)\}$

Similarly, we can make this collection dependent on parameters. Say $\varphi(x, y_1, ..., y_n)$ is a formula with only free variables $x, y_1, ..., y_n$. If $p_1, ..., p_n \in V$, then the collection of all $a \in V$ is denoted by $\{x|\varphi(x, p_1, ..., p_n)\}$

Collections of this kind are called *classes*. Every set is a class: if $a \in V$ then $a = \{x|x \in a\}$ ($x \in a$ being the $\varphi(x, y)$ mentioned above).

A class that is not a set is called a *proper class*.

1.11 Remark

Proper classes exist:

$E = \{x|x \notin x\}$ is obviously a class, but the argument for Russell's Paradox shows that E is not a set. Similarly, V is a proper class:

$V = \{x|x = x\}$

But if V were a set, then by the Separation Schema, also $\{x \in V|x \notin x\} = E$ would be a set. Side note: the Axiom of Foundation hints at $E = V$.

(picture)

Intuitively: proper classes are collections that are very large, so that we would not consider them sets.

Sets are elements of classes. However, proper classes are *not* elements of any classes. Proper classes may be identified with formulae that describe them.

These formulae are *not* part of our language. This puts limitations on manipulators with proper classes.

1.12 Basic Constructions

We show how to simulate usual mathematical constructions in $(V, \in)$

(a) We know already that unions are given by the Union Axiom. We define the *intersection* of set a as follows:

$$\bigcap a = \{x|(\forall y \in a)(x \in y)\}$$

Here, $(\forall y \in a)$ is an abbreviation for $(\forall y)(y \in a \to ...)$.

So: we view a as a family of sets. Then $\bigcap a$ is the set that consists of all elements that are common to all sets in a. Informally, $\bigcap\{x, y, z\} = x \cap y \cap z$, although we don't know what any of that notation means yet.

If $a = \varnothing$ then $\bigcap a = V$. This is because every $x \in V$ satisfies the implication $(\forall y)(y \in a \to x \in y)$. If $a \neq \varnothing$, then $\bigcap a$ is a set. This is because we can pick $b \in a$ and then $\bigcap a \subseteq b$. So

$\bigcap a = \{x \in b|(\forall y \in a)(x \in y)\}$ i.e. $\bigcap a$ is defined using the Separation Schema.

(b) *Finite tuples*
Pairing axiom tells us that if x_0, x_1 are sets then we have the set $\{x_0, x_1\}$. We can iterate this using Pairing and Union Axioms:

If x_0, x_1, x_2 are sets, then also $\{x_0, x_1\}$ is a set and $\{x_2\}$ is a set. Then $\{\{x_0, x_1\}, \{x_2\}\}$ is a set. Then using the Union Axiom: $\bigcup\{\{x_0, x_1\}, \{x_2\}\}$ is the set whose elements are precisely x_0, x_1, x_2. By continuing this process, for any sets $x_0, ..., x_{n-1}$ we get the set whose elements are precisely $x_0, ..., x_{n-1}$ (i.e. $\{x_0, ..., x_{n-1}\}$).

(c) *Finite boolean operators*
Now we can let $x_0 \cup x_1 \cup \cdots \cup x_{n-1} = \bigcup\{x_0, ..., x_{n-1}\}$ and $x_0 \cap x_1 \cap \cdots \cap x_{n-1} = \bigcap\{x_0, ..., x_{n-1}\}$ (these sets exist by (b)).

We also define:
Set difference: $x - y = \{z \in x | z \notin y\}$. This is a set by Separation.
Symmetric difference: $x\Delta y = (x - y) \cup (y - x)$.
$x\Delta y$ consists of those objects on which x, y disagree.

(d) *Ordered pairs*
For each pair of sets x, y we define the set $\langle x, y\rangle$ that would simulate what mathematicians understand under ordered pair. It has to satisfy the following property:
$\langle x, y\rangle = \langle x', y'\rangle \iff (x = x' \wedge y = y')$.
We let $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$. This is in V by the above.

Claim: $\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\} \iff (x = x' \wedge y = y')$. The $\Leftarrow$ direction is obvious. As for $\Rightarrow$: we use the Axiom of Extensionality to check that $x = x'$ and $y = y'$. By the Axiom, we know that the elements of the two sets are equal. Then we use the Axiom of Extensionality again to check the separate elements of our first set, getting that their elements are equal.

(e) *Ordered tuples*
This is defined inductively: $\langle x_0, x_1, x_2\rangle = \langle\langle x_0, x_1\rangle, x_2\rangle$
$\cdots$
$\langle x_0, ..., x_{n-1}\rangle = \langle\langle x_0, ..., x_{n-2}\rangle, x_{n-1}\rangle$.

(f) A *binary relation* is a class of ordered pairs. A function F is a binary relation that simulates the assignment, i.e., the one that satisfies
$(\forall x)(\forall y)(\forall y')(\langle x, y\rangle \in F \wedge \langle x, y'\rangle \in F \rightarrow y = y')$.
We write $y = F(x)$ instead of $\langle x, y\rangle \in F$.

(g) *Cartesian products*
If A, B are classes then the Cartesian product $A \times B$ is defined:
$\{\langle x, y\rangle | x \in A \wedge y \in B\}$
This is a class because if $\varphi(x)$ defines A and $\psi(y)$ defines B then $\theta(x, y) = \varphi(x) \wedge \psi(y)$ defines $A \times B$.

Now: if A, B are sets then also $A \times B$ is a set. Rigorously: if $A, B \in V$, then $A \times B \in V$. This is not obvious and requires either Replacement or Power set. (It can be proved that without these two, Cartesian products may not be possible with some sets).

$A \times B$ is a set:

Proof. Argument 1- Using Replacement:

Fix $x \in A$. Then $(\forall y \in B)(\exists! z)(z = \langle x, y\rangle)$. In other words, we have a formula $\varphi(x, y, z) \equiv z = \langle x, y\rangle$ that defines a function $y \mapsto \langle x, y\rangle$. By Replacement there is a set C such that the range of this function is contained in C, i.e., for each $y \in B$ we get the tuple $\langle x, y\rangle \in C$. We then use the Separation Schema to separate all these tuples: we let

$B_x = \{\langle x, y\rangle | \langle x, y\rangle \in C\} =$ (by construction) $\{z \in C | (\exists y \in B)(\langle x, y\rangle = z)\}$

This is a valid definition using Separation. This tells us that for each $x \in A$, $B_x = \{\langle x, y\rangle | y \in B\}$ is a set.

(picture)

So: we showed using Replacement Schema + Separation that for each $x \in A$, $B_x = \{\langle x, y\rangle | y \in B\}$ is a set. Moreover, the assignment/function $x \mapsto B_x$ is a class because we have a description for this assignment in the language of set theory:

$$u = B_x \Longleftrightarrow (\forall z)[z \in u \longleftrightarrow (\exists y)(y \in B \wedge \langle x, y\rangle = z)]$$

So: for each $x \in A$ there is exactly one u such that $\varphi(x, u)$. We apply Replacement again to conclude that there is some set D such that: for each $x \in A$ there is $u \in D$ such that $\varphi(x, u)$. Hence all "sections" B_x are elements of the set D. But D may possibly contain other elements that are not of our interest. So we use Separation to get rid of these:

$\{u \in D | (\exists x)(x \in A \wedge u = B_x)\}$ is a set by Separation and it is exactly $\{B_x | x \in A\}$. Finally, by Union Axiom: $A \times B = \bigcup\{B_x | x \in A\} = \bigcup\{\langle x, y\rangle | y \in B\}$ (sloppy notation for the second one!)

Now let's use the Power axiom instead of Replacement:

Argument 2- Power axiom

Again recall $A \times B = \{\langle x, y\rangle | x \in A \wedge y \in B\}$. We said that $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$. Now:

$$x \in A \Rightarrow x \in A \cup B \Rightarrow \{x\} \subseteq A \cup B \Rightarrow \{x\} \in \mathcal{P}(A \cup B).$$
$$x \in A, y \in B \Rightarrow x, y \in A \cup B \Rightarrow \{x, y\} \subseteq A \cup B \Rightarrow \{x, y\} \in \mathcal{P}(A \cup B).$$

Those two $\Rightarrow$

$\{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(A \cup B) \Rightarrow \{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$. So

$A \times B = \{z \in \mathcal{P}(\mathcal{P}(A \cup B)) | (\exists x)(\exists y)(x \in A \wedge y \in B \wedge [z = \{\{x\}, \{x, y\}\}])\}$ This is a set by Separation.

□

(h) Generalization of construction from (g),(a)

If F is a class function, say $y = F(x) \Leftrightarrow \varphi(y, x, p_1, ..., p_n)$ and A is a set, then the restriction of F to A:

$F \restriction A = \{\langle x, F(x)\rangle | x \in A\}$ is a set. Hint: First use replacement to find a set B such that the pointwise image of A $(F[A]) = \{F(x)|x \in A\} \subseteq B$ then use Separation to conclude that $F \restriction A = \{\langle x, y\rangle \in A \times B | y = F(x)\}$ is a set.

(i) Operations on Classes:
We can perform finite Boolean operations: $A_1 \cap ... \cap A_n$, $V - A$,...

For instance, if $\varphi_i(x)$ is a formula that describes the class A_i , $i = 1, ..., n$, then $\varphi_1(x) \wedge ... \wedge \varphi_n(x)$ describes $A_1 \cap ... \cap A_n$. Similarly, the complement $V - A$ is described by $\neg\varphi_1(x)$.

We *cannot* form pairs: if A, B are classes then $\{A, B\}$ is not a class... we don't have a description in our language for that.

There is a way of forming infinite intersection/unions of classes but this requires just some way of coding families of classes through one class.

Example: $\{A, B\}$ could be coded as: $\{\langle\varnothing, x\rangle | x \in A\} \cup \{\langle\{\varnothing\}, y\rangle | y \in B\}$.

2. Natural Numbers

Goals of this section:
—Show how to represent natural numbers in the structure $(V, \in)$.
—Present a simple example of recursion.
In section 2 we work in the theory "ZF without Axiom of Foundation." This will be important for some things that we will do.

2.1 Definition: Inductive set
A set X is inductive iff:
(i) $\varnothing \in X$
(ii) for each $x \in X$, also $S(x) = x \cup \{x\} \in X$.

So inductive sets contain all elements of the form $\varnothing$, $\{\varnothing\}$, $\{\varnothing, \{\varnothing\}\}$,...
Look back at the Axiom of Infinity: it says that there is an inductive set.

Let IND= the class of all inductive sets. This is a class because IND=$\{x|x$ is inductive $\}$. And by the Axiom of Infinity: IND$\neq \varnothing$. So by the remarks at the end of section 1: $\bigcap$IND is a set.

2.2 Definition: ω
$\omega = \bigcap$IND=smallest inductive set. ω will represent the natural numbers in this structure.

2.3 Definition: Transitive set
A set x is *transitive* iff every element of x is a subset of x, i.e., $(\forall z)(z \in x \rightarrow z \subseteq x)$. Equivalently: $u \in z \wedge z \in x \longrightarrow u \in x$. This tell us that x is a family of sets that correctly identifies all elements of sets in x.

In other words, if $z \in x$ it is enough to know x to be able to recover all elements of z.

(picture)

Equivalently: x is closed under "going backwards" in terms of x. Notice: elements of ω are transitive.

10/07/09

Recall: a set x is transitive iff $z \in x \to z \subseteq x$ for all z. So if x is transitive and $z, z' \in x$ then the information that x provides is sufficient to decide whether $z = z'$. So this x behaves like a little universe of sets. I.e., the structure $(x, \in)$ satisfies the Axiom of Extensionality.

2.4 Proposition:

(a) Every element of ω is a transitive set.

(b) ω itself is a transitive set.

Proof. This is mathematical induction simulated inside the universe $(V, \in)$.

(a) Let $A = \{x \in \omega | x \text{ is a transitive set}\}$. A is a set by Separation. Obviously $A \subseteq \omega$. We show that A is inductive. Because ω is the smallest inductive set, we must have $\omega \subseteq A$. Hence $A = \omega$. To see that A is inductive:

—$\varnothing \in A$ this is true since $\varnothing$ is obviously transitive.

—$x \in A \Rightarrow x \cup \{x\} \in A$:

Assume x is transitive. Now if $z \in x \cup \{x\}$ then either $z \in x$, but then $z \subseteq x \subseteq x \cup \{x\}$ (since x is transitive)(this is the induction hypothesis) or else $z = x$ but then $z \subseteq x \cup \{x\}$.

(b) We let $B = \{x \in \omega | x \subseteq \omega\}$. Again we show that B is inductive. As in A, this will give $B = \omega$.

—$\varnothing \in B$ (because $\varnothing \subseteq \omega$)

—If $x \in B$ then $x \cup \{x\} \subseteq \omega$ (by the induction hypothesis: $x \subseteq \omega$ as $x \in B$). □

Now one of the most important definitions in set theory:

2.5 Definition: Well-founded

A binary relation R is *well-founded* iff every nonempty set has an *R-minimal element*. That is: if A is a set and $A \neq \varnothing$ then there is some $a \in A$ such that

$x \in A \longrightarrow \langle x, a \rangle \notin R$ for all x.

(picture)

This includes the option that $\langle x, a \rangle \notin R$ for all x.

2.6 Remark:

The Axiom of Foundation asserts that the membership relation is well-founded.

Convention: if R is a binary relation, we often write xRy instead of $\langle x, y \rangle \in R$ (we don't want to be writing $\langle x, y \rangle \in\in$. This would mean $\langle x, y \rangle \in \dot{\in}$)

2.7 Notation:

If R is a binary relation and A is a class, then the *restriction of R to A* is the binary relation $R \cap (A \times A) = \{\langle x, y \rangle \in R | x, y \in A\}$.

2.8 Proposition:

The restriction of the relation $\in$ to ω is well-founded. The important point here is that we can prove this without the Axiom of Foundation (hence we are not assuming it).

(important example of induction)

Proof. This boils down to proving the following: If $A \subseteq \omega$ is nonempty, then A has an $\in$-minimal element.

Notice: $x \in \omega$ is an $\in$-minimal element of A iff $x \in A$ and $x \cap A = \varnothing$ (this says that $z \in x \rightarrow z \notin A$— check with definition 2.5). We prove the contraposition: If $A \subseteq \omega$ has no $\in$-minimal element, then $A = \varnothing$.

Let $B = \{x \in \omega | x \cap A = \varnothing\}$. We show that $B = \omega$. This will tell us that $x \cap A = \varnothing$ for all $x \in \omega$. So in particular: if $y \in \omega$ then also $y \cup \{y\} \in \omega$ and $(y \cup \{y\}) \cap A = \varnothing$, hence $y \notin A$. This shows: $y \in \omega \Rightarrow y \notin A$. Since $A \subseteq \omega$, we have $A = \varnothing$.

So we now prove that $B = \{x \in \omega | x \cap A = \varnothing\}$ is equal to ω. Again, it suffices to prove that B is inductive.

—$\varnothing \in B$ because $\varnothing \cap A = \varnothing$.

—If $x \in B$ then $(x \cup \{x\}) \cap A = \varnothing$, as otherwise $(x \cup \{x\}) \cap A = \{x\}$ since $x \in B$, i.e., $x \cap A = \varnothing$.

But then $x \in A$ and $x \cap A = \varnothing$ so x is an $\in$-minimal element of A. But we were assuming that A has no $\in$-minimal element. □

Summary: this is essentially a rigorous version of the naive inductive proof that every nonempty $A \subseteq \mathbb{N}$ has a least element. However, here we are not assuming that $\in$ is a linear ordering on ω.

2.9 Definition: Ordering

A binary relation R is a *partial ordering* on A iff
R is reflexive on A, i.e., xRx for all $x \in A$
R is antisymmetric, i.e., $(xRy \wedge yRx) \Rightarrow y = x$
R is transitive, i.e., $(xRy \wedge yRz) \Rightarrow xRz$

R is a *strict partial ordering* iff
R is irreflexive, i.e., $\neg xRx$ for all x.
R is transitive.

There is an obvious relationship between the two: If R is a non-strict partial ordering, then $(xRy \wedge x \neq y)$ is a strict one; we call it *the strict part of R*.

If R is a strict partial ordering then $(xRy \vee x = y)$ is a non-strict one, and its strict part is R (check this).

A strict partial ordering R is *linear* on A iff it is *trichotomic*:
xRy or $x = y$ or yRx for all $x, y \in A$.

2.10 Definition: Well-ordering

A strict linear ordering on A is a *well-ordering* on A iff it is well-founded.

2.11 Remark:

(a) If R is a linear ordering $\prec$ on A and $X \subseteq A$ is nonempty, then if $x \in X$ is a minimal element of X with respect to ($\prec$) , because any two elements are $\prec$-comparable.

(b) If we have a strict partial ordering $\prec$ on A with the property that each nonempty $X \subseteq A$ has a $\prec$-least element then $\prec$ is automatically linear, hence $\prec$ is a well-ordering on A. (Think about this).

Recall:

• A binary relation R is well-founded iff every set $B \neq \varnothing$ has an R-minimal element, i.e. some $b \in B$ such that $\langle z, b\rangle \notin R$ for all $z \in B$.

• A strict ordering is:

—A *strict linear ordering* $\prec$ on A iff it satisfies the trichotomicity:

$x \prec y$ or $x = y$ or $y \prec x$ for all $x, y \in A$.

—A *well-ordering* on (of) A iff it is linear and well-founded. Equivalenty: iff every $\varnothing \neq X \subseteq A$ has an $\prec$-least element.

2.12 Remark:

Any well-founded relation is irreflexive. If R is well-founded, then for each x we have $\langle x, x\rangle \notin R$ because otherwise $\{x\}$ would be a set without an R-minimal element. However, R need not be necessarily transitive if it is irreflexive.

2.13 Lemma:

For each $x, y \in \omega$ we have:

$x \in y \rightarrow (S(x) \in y \vee S(x) = y)$

Analogy: $x < y \rightarrow (x + 1 < y \vee x + 1 = y)$ for $x, y \in \mathbb{N}$. We also have the following parallel: $x + 1 = S(x)$.

Proof. Induction on y:

Let $A_x = \{y \in \omega | x \in y \rightarrow (S(x) \in y \vee S(x) = y)\}$

We show that A_x is inductive.

—$\varnothing \in A_x$ trivially, as "$x \in \varnothing$" is always false.

—Now assume $y \in A_x$. We prove that $S(y) \in A_x$.

Assume $x \in S(y) = y \cup \{y\}$. Then

Either $x \in y$, in which case $S(x) \in y$ or $S(x) = y$ by the inductive hypothesis. In either case, $S(x) \in S(y)$.

Or else $x = y$ but then trivially $S(x) = S(y)$. □

2.14 Proposition:

$\in$ restricted to ω is a strict linear ordering on ω. Because we already proved that this restriction is well-founded, it follows that it is a well-ordering on ω.

Proof. Irreflexivity: We want $x \notin x$ for all $x \in \omega$. Let $A = \{x \in \omega | x \notin x\}$. We show that A is inductive:

—$\varnothing \in A$ trivially

—Assume that $x \notin x$. Show that $S(x) \notin S(x)$. What if $S(x) \in S(x) = x \cup \{x\}$?

Case 1: $S(x) \in x$. But we have proved that x is transitive, so $S(x) \subseteq x$. But then $x \in x$— a contradiction.

Case 2: $S(x) = x$. In this case, trivially $x \in x$. Contradiction again.

Transitivity: This follows from our previous result that all elements of ω are transitive sets: if $x \in y$, then $y \in z$, because z is a transitive set: $y \subseteq z \rightarrow x \in z$.

Trichotomicity: Want to prove: $x \in y$ or $x = y$ or $y \in x$ for all $x, y \in \omega$. This is proved by induction on x:

Let $A = \{x \in \omega | (\forall x \in \omega)(x \in y \vee x = y \vee y \in x)\}$. Prove A is inductive:

—$\varnothing \in A$: $\varnothing \in y \vee \varnothing = y \vee y \in \varnothing$ for all $y \in \omega$. This is proved by induction on y. Easy–exercise.

—$x \in A \rightarrow S(x) \in A$. So assume $(x \in y \vee x = y \vee y \in x)$. If $x = y$ or $y = x$ then $y \in S(x) = x \cup \{x\}$. This is clear.

If $x \in y$ then we get $S(x) \in y$ or $S(x) = y$. Also easy to check. □

The point of the previous work:
$(\omega, \in)$ behaves exactly as our intuitive $(\mathbb{N}, <)$.

Recursion on ω (A warm up for later.)

Intuitively, we want to construct objects $a_0, a_1, ..., a_n, ...$ for $n \in \mathbb{N}$ *recursively*, i.e. we will have some "recipe" how to construct a_n if we already know $a_0, ..., a_{n-1}$.

The "recipe" can be viewed as some function G, so $a_n = G(\langle a_0, ..., a_{n-1} \rangle)$. We do not require that objects $a_0, a_1, ...$ are elements of a set fixed in advance. However, we would like that the sequence $\langle a_0, a_1, ... \rangle$ constructed at the end is a set, i.e. it is an element of V. This will require the use of Replacement.

2.15 Theorem: (Construction by Recursion on ω)

Let $G : V \rightarrow V$ be a class function. Then, there is a unique function $F : \omega \rightarrow V$ such that for all $n \in \omega$ we have

$$(*) f(n) = G(f \upharpoonright n).$$

Recall $n = \{0, ..., n-1\}$.

Proof. Uniqueness:

Induction on n: Assume f, f' satisfy $(*)$. Now we let $A = \{n \in \omega | (\forall i < n) f(i) = f'(i)\}$. We show that A is inductive. This will tell us that $f(n) = f'(n)$ for all $n \in \omega$.

—$\varnothing \in A$: trivial.

—$n \in A \longrightarrow S(n) \in A$: $n \in A$ means $f \upharpoonright n = f' \upharpoonright n$. To see that $S(n) \in A$, it is enough to prove that $f'(n) = f(n)$. So: $f'(n) = G(f' \upharpoonright n) = G(f \upharpoonright n) = f(n)$. Next: existence. We will look at all finite approximations to this function and show that these finite approximations are coherent, then we take the union. We will then use Replacement to show that the collection of all these objects is a set.

We are proving the theorem on construction by Recursion. We proved the uniqueness part, now we prove the existence.

Existence. Recall: given is a class function $G : V \rightarrow V$ which tells us what to do at the next step in the recursion. We want to construct $f : \omega \rightarrow V$ such that

(i) $(\forall u \in \omega)(f(u) = G(f \upharpoonright u))$

(ii) f is a set.

Strategy: we define a class (F) of all approximations to f. Then we show that $\mathcal{F}$ is a set, i.e. $\mathcal{F} \in V$. Then finally we "glue" all approximations together.

We let $\mathcal{F} =$ the class of all functions p such that $\operatorname{dom}(p) \in \omega$ and for all $i \in \operatorname{dom}(p)$ we have $p(i) = G(p \restriction i)$.

Here $\operatorname{dom}(p) =$ the set of all i such that $p(i)$ is defined. In general, if A is a class we let $\operatorname{dom}(A) = \{x | (\exists y)(\langle x, y\rangle \in A\})$ and $\operatorname{rng}(A) = \{y | (\exists x)(\langle x, y\rangle \in A\})$

Easy to see: if A is a set then so are $\operatorname{dom}(A)$ and $\operatorname{rng}(A)$ (exercise).

So elements of $\mathcal{F}$ are functions p where $\operatorname{dom}(p) = \{0, 1, ..., n-1\} = n$ for some $n \in \omega$ and $p(i) = G(\langle p(0), p(1), ..., p(n-1)\rangle (= p \restriction n))$ – intuitively.

$\mathcal{F}$ is a class, because it has a description in the language of set theory:

$$p \in \mathcal{F} \iff p \text{ is a function} \wedge \ \operatorname{dom}(p) \in \omega \wedge (\forall i)[i \in \operatorname{dom}(p) \to (\exists z)(z = p \restriction i) \wedge \langle i, G(z)\rangle \in p]$$

(tells me $p(i) = G(p \restriction i)$). It is easy to check that this can be turned into a formula in the language of set theory. (exercise).

Claim 1: $\mathcal{F}$ is a *coherent class* of functions, i.e. if $p, q \in \mathcal{F}$ then for all $i \in \omega$:

$i \in \operatorname{dom}(p) \cap \operatorname{dom}(q) \Rightarrow p(i) = q(i)$. In particular, if $p, q \in \mathcal{F}$ and $\operatorname{dom}(p) = \operatorname{dom}(q)$ then $p = q$.

Proof: This is very much like the proof of uniqueness we had last time. Given $p, q \in \mathcal{F}$, show that

$$A = \{n \in \omega | (\forall i \in n)[(i \in \operatorname{dom}(p) \cap \operatorname{dom}(q)) \to p(i) = q(i)]\}$$

is inductive. (exercise).

Notation— If we have a coherent class of functions $\mathcal{H}$, i.e. $p(x) = q(x)$ whenever $p, q \in \mathcal{H}$ and $x \in \operatorname{dom}(p) \cap \operatorname{dom}(q)$, we can "glue" $\mathcal{H}$ together into one function h defined as follows:

$$\operatorname{dom}(h) = \bigcup\{\operatorname{dom}(p) | p \in \mathcal{H}\} \text{ and } h(x) = p(x) \text{ for any } p \in \mathcal{H} \text{ s.t. } x \in \operatorname{dom}(p).$$

This gives us a function, because the value $p(x)$ is always the same no matter how we pick $p \in \mathcal{H}$. Because we view functions in $\mathcal{H}$ as sets of ordered pairs:

$$h = \bigcup \mathcal{H}.$$

By Claim 1: $(\forall n)(\exists! p)(p \in \mathcal{H} \wedge \operatorname{dom}(p) = n)$.

Claim 1 tells us that there exists at most one p for each n. On the other hand, by induction on n we can prove that for each $n \in \omega$ there is at least one $p \in \mathcal{H}$ such that $\operatorname{dom}(p) = n$. Simply show that

$$A = \{n \in \omega | (\exists p)(p \in \mathcal{F} \wedge \operatorname{dom}(p) = n\})$$

is inductive. The induction step: if $n \in \mathcal{F}$, we have $p \in \mathcal{F}$ with $\operatorname{dom}(p) = n$. Then $p' = p \cup \{\langle n, G(p)\rangle\} \in \mathcal{F}$ and $\operatorname{dom}(p) = S(n)$.

By Replacement and Separation (Was a homework problem) $\mathcal{F}$ is a set, i.e. $\mathcal{F} \in V$. But by the above remark on notation: $f := \bigcup \mathcal{F}$ is a function. Moreover, $\operatorname{dom}(f) = \omega$ because if $n \in \omega$ then $S(n) \in \omega$ so we have some $p \in \mathcal{F}$ such that $\operatorname{dom}(p) = S(n)$, i.e. $n \in \operatorname{dom}(p)$.

Finally: for each $p \in \mathcal{F}$ and $i \in \operatorname{dom}(p)$, $f(i) = p(i)$ by definition of f. Hence if $i \in \omega$ pick some $p \in \mathcal{F}$ such that $i \in \operatorname{dom}(p)$. Then $f(i) = p(i)$ by the above $(= G(p \restriction i)$ since $p \in \mathcal{F} = G(f \restriction i))$ (by the above, or just by the fact that $p \in \mathcal{F}$ and the definition of $\mathcal{F}$). □

Example 1

To each $x \in \omega$ there is exactly one function $f_x \in V$ such that $f_x : \omega \to \omega$ and the following is true:

$$f_x(0) = x \ (x + 0 = x)$$
$$f_x(S(y)) = S(f_x(y)) \ (x + (y + 1) = (x + y) + 1)$$

This is guaranteed by the theorem on construction by recursion. What would be the function G in this case? $G(u) = x$ if $x = \varnothing$ $G(u) = S(\bigcup \operatorname{rng}(u))$ otherwise. $(u = \langle x + 0, ..., x + n\rangle)$.

So f_x simulates adding y to x. Now again: to each $x \in \omega$ there is a unique f such that f satisfies the conditions we assigned to it. By construction by recursion we have a function F that assigns $x \mapsto f_x$ when f_x is as in above.

So we have $F : \omega \to V$ such that $f(x) = f_x$. Hence we can define a function $f : \omega \times \omega \to \omega$ by $f(x, y) = f_x(y) = F(x)(y)$.

Hence $f \in V$ and f simulates addition on ω.

Similarly we can define functions simulating multiplication and exponentiation: following the recursive rules:

$x \cdot 0 = 0$

$x \cdot S(y) = x \cdot y + x$

and also $x^0 = 1$

and $x^{n+1} = x^n \cdot x$.

From now on we will write $x + 1$ instead of $S(x)$.

In mathematics, we start from natural numbers $\mathbb{N}$ – assume they are given. Then we construct $\mathbb{Z}$ as a ring: elements of $\mathbb{Z}$ are represented as equivalence classes of pairs (m, n) where the equivalence relation is given by

$(m, n) \sim (m', n') \Leftrightarrow m + n' = m' + n$

So $(m, n) \sim (m', n')$ represents the difference $m - n$. The operations are defined the usual way as on quotients. We may do this by taking the cross product, using the power set axiom, etc...

Then we construct $\mathbb{Q}$ in a similar way: pairs (p, q) where $p \in \mathbb{Z}$, $q \in \mathbb{Z}^+$ represent the ratio $\frac{p}{q}$; the equivalence relation is $(p, q) \sim (p', q')$ iff $pq' = p'q$. The operations are again defined the usual way as on quotients.

Then we construct $\mathbb{R}$. One possibility is to follow Dedekind:

If $(A, <)$ is a strict linear ordering on A, an *initial segment* of $(A, <)$ is that is a subset $(A', <)$ that is *downward* closed: i.e. if $x \in A'$, $y \in A$, and $y < x$ then $y \in A'$.

Example: $\mathbb{Q} \cap (-\infty, \sqrt{2})$ cannot be described if we want to refer only to finitely many elements of $\mathbb{Q}$. But it is an initial segment of $(\mathbb{Q}, <)$.

Let us say that the initial segment $(A', <)$ of $(A, <)$ is induced by $X \subseteq A$ iff

—$X \subseteq A'$

—For every $a \in A'$ there is some $x \in X$ such that $a \leq x$.

So A' is the "downward closure" of X.

For instance: the downward closure of $\{1 - \frac{1}{n} | n \in \mathbb{N}\}$ is $(-\infty, 1)$.

Let $\mathbb{R}$ = the set of all initial segments of $(I, <)$ of $(\mathbb{Q}, <)$ such that $I \notin \{\varnothing, \mathbb{Q}\}$.

We think of I as representing $\sup(I)$. Here we seriously use the Power Set Axiom, as each I is in $\mathcal{P}(\mathbb{Q})$. So we let $\mathbb{R}^+$ = the set of all $I \in \mathbb{R}$ such that I contains some positive rational number.

We define the operations:

$\oplus$ on $\mathbb{R}$: $I \oplus J =$ the initial segment of $\mathbb{Q}$ induced by $\{x+y | x \in I \wedge y \in J\}$. (Actually this is an initial segment of $(\mathbb{Q}, <)$).

$\otimes$: For $I, J \in \mathbb{R}^+$ we let $I \otimes J =$ the initial segment of $\mathbb{Q}$ induced by $\{xy | x \in I \cap \mathbb{Q}^+ \wedge y \in J \cap \mathbb{Q}^+\}$. This operation can then be extended to $\mathbb{R} \times \mathbb{R}$ in the straightforward way. With a little work, we can prove that $(\mathbb{R}, \oplus, \otimes, (-\infty, 0), (-\infty, 1))$ is a ring; $\mathbb{Q}$ is dense in $\mathbb{R}$ and $\mathbb{R}$ has suprema:

If $\mathcal{F} \subseteq \mathbb{R}$ then $\bigcup \mathcal{F}$ is an initial segment of $\mathbb{Q}$ so if $\bigcup \mathcal{F} \neq \mathbb{Q}$ then $\bigcup \mathcal{F} \in \mathbb{R}$ and one can show that this operation $\bigcup$ is really the "sup" operation.

Once we have $\mathbb{R}$: We can construct all objects relevant for maths: Function spaces, measures, algebraic structures: these involve several applications of the Power Set Axiom.

Say the collection of all function $f : \mathbb{R} \to \mathbb{R}$ is viewed as a subset of $\mathcal{P}(\mathbb{R} \times \mathbb{R})$.

Summary: once we have a faithful representation of natural numbers in $(V, \in)$ we can faithfully represent all other objects of mathematical interest.

Important/popular Zeman example:

It is possible to show: there is a continuous function $f : [0,1] \to [0,1] \times [0,1]$. Such a function is called *Peano function*. So this function is represented by a pair of functions (f_1, f_2), i.e.

$f(x) = (f_1(x), f_2(x))$

One can show: for each $x \in (0,1)$ both f_1, f_2 cannot have a derivative.

Question: is it possible that for every $x \in (0,1)$ at least one of f_1, f_2 has a derivative?

Answer: No way to decide. This means the following– we can find a universe of sets $(V, \in)$ where this question has positive answer, and a universe of sets $(V', \in')$ where this question has negative answer.

From now on, we treat natural numbers informally: will often write $n < n'$ instead of $n \in n'$ and $n+1$ instead of $S(n)$.

Example 2:

Using recursion we can construct the following sets:

$$V_0 = \varnothing$$
$$V_{n+1} = \mathcal{P}(V_n)$$
$$V_\omega = \textstyle\bigcup_{n \in \omega} V_n.$$

One can show that $V_n \subseteq V_{n+1}$. Also, all V_n are transitive.

One can show that all axioms of ZFC without the axiom of infinity are true in the structure $(V_\omega, \in)$. On the contrary, $(V_\omega, \in)$ satisfies the statement "there is no infinite set."

One can show that if we replace the axiom of infinity by the axiom "there is no infinite set" (negation of axiom of infinity) we get a theory of the same strength as arithmetic (the theory of finite sets).

What is the function G in the theorem on recursion?

Recursively, we are constructing sequences: $g_0 = \varnothing$, $g_1 = \langle \varnothing \rangle$, $g_2 = \langle \varnothing, \mathcal{P}(\varnothing) \rangle$, ... $g_n = \langle \varnothing, \mathcal{P}(\varnothing), ..., \mathcal{P}^{n-1}(\varnothing) \rangle$.

$g_{n+1} = G(g_n)$ by the theorem, so: G has the following definition:

$$G(u) = \varnothing \text{ if } u = \varnothing.$$
$$G(u) = \mathcal{P}(\bigcup \operatorname{rng}(n)) \text{ otherwise.}$$

This is also defined for u that is not a function, but we do not care about the values of G in that case.

Example 3:
To each set x there is a transitive set x' such that $x \subseteq x'$. In fact, there is a smallest among all such sets; we call this set the *transitive closure of x.*.

So: $x' = \operatorname{trcl}(x)$ iff

—x' is transitive and $x \subseteq x'$

—For all y: if y is transitive and $x \subseteq y \Rightarrow x' \subseteq y$.

How to get x': By recursion:

$$x_0 = x$$
$$x_1 = \bigcup x$$
$$\ldots$$
$$x_{n+1} = \bigcup x_n$$
$$\ldots$$
$$x' = \bigcup_{n\in\omega} x_n.$$

Check that x' has these properties: The function G from the theorem on recursion: $G(u) = x$ if $u = \varnothing$ and $G(u) = \bigcup\bigcup \operatorname{rng}(u)$.

Example 4:
A binary relation R is *set-like* iff for each x, the class $\operatorname{pred}_R(x) = \{z|zRx\}$ is a set. Here $\operatorname{pred}_R(x)$ stands for "$R-$predecessors of x."

An example of a proper class relation that is set-like:

$\in$. Because if $x \in V$ then $\{z|z \in x\} = x$. We say that a class A is *transitive with respect to* R iff for every $x \in A$ and every z: $zRx \Rightarrow z \in A$. In other words, $\operatorname{pred}_R(x) \subseteq A$.

If B is a class, then the *transitive closure of* B with respect to R is the smallest (with respect to inclusion) class B^* such that B^* is transitive with respect to R and $B \subseteq B^*$. If R is $\in$, then we get *Example 3*.

———

Recall: R is a set-like relation, i.e., $\operatorname{pred}_R(x) = \{z|zRx\}$ is a set.

If A is a class: the transitive closure of A under R is the smallest class A' that is transitive with respect to R and contains A as a subclass.

First notice: For each $x \in V$: $\operatorname{trcl}_R(x)$ is a set. Why: Because we can define sets x_n by recursion–

$$x_0 = \bigcup\{\operatorname{pred}_R(z)|z \in x\}$$
$$x_{n+1} = \bigcup\{\operatorname{pred}_R(z)|z \in x_n\}.$$

Then $x_\omega = \bigcup_{n\in\omega} x_n$. Notice that x_ω is a set and $x_\omega = \operatorname{trcl}_R(x)$.

Notice: if x is a set then the function $f_x : x \to V$ defined by $f_x(z) = \operatorname{pred}_R(z)$ is a set because we assume that R is set-like; the conclusion then follows by Replacement + Separation. But then $x_0 = \bigcup \operatorname{rng}(f_x)$ is a set. So the function G in the theorem on recursion is the following:

$$G(u) = \begin{cases} \bigcup\{ \text{ pred}_R(z) | z \in x\} & \text{if} \qquad u = \varnothing \\ \bigcup \text{rng}(u) & \text{otherwise} \end{cases}$$

Hence x_ω is a set. Easy to check: $x_\omega = \text{trcl}_R(x)$. This tells us that for any set x, $\text{trcl}_R(x)$ is a set. Moreover, the fact "$y = \text{trcl}_R(x)$" can be expressed by a formula in the language of set theory– this formula obviously depends on the definition of R, i.e., on the formula that defines R.

Now if A is a class, then $\text{trcl}_R(A) = \bigcup \text{trcl}_R(\{x\})$ – check this.

So we can write: $z \in \text{trcl}_R(A) \leftrightarrow (\exists x)(x \in A \wedge z \in \text{trcl}_R(\{x\}))$. ($A$ can be described by a formula, as A is a class). And $z \in \text{trcl}_R(\{x\})$ means $(\exists y)(y = \text{trcl}_R(\{x\}) \wedge z \in y)$.

So the formula on the right defines $\text{trcl}_R(A)$, which means that $\text{trcl}_R(A)$ is a class.

——————End of Section Two————End of Most Formalisms——————

3. Ordinals

Goals:
- Present basic facts about well-orderings.
- Present Von Neumann's construction of ordinals.

Recall: a pair $(A, <)$ is a *well-ordered set* iff $<$ is a well-ordering on A, i.e. $<$ is a linear ordering on A such that every nonempty $X \subseteq A$ has a least element with respect to A.

The immediate Purpose of well-orderings: they offer a generalization of constructions by recursion that go beyond ω.

(picture of dots and numbers and the ωth element ...)

Notice: $0, 1, 2, ..., \omega, S(\omega) = \omega \cup \{\omega\}, S(S(\omega)), ...$

3.1 Definition: Bijection

Let (A, R) , (B, S) be structures with binary relations, i.e., $R \subseteq A \times A$ and $S \subseteq B \times B$. We say that a bijection $f : A \to B$ is an isomorphism between (A, R) and (S, B) iff for all $x, y \in A$ we have:

$xRy \Leftrightarrow f(x)Sf(y)$

Recall: $f : A \to B$ is a bijection iff f is injective (i.e. $x \neq y \to f(x) \neq f(y)$). and surjective (i.e. $\text{rng}(f) = B$, equivalently $(\forall z \in B)(\exists x \in A)(f(x) = z)$).

Remark: The equivalence "$\Leftrightarrow$" cannot be replaced by "$\Rightarrow$;" for instance: the identity map $id : \mathbb{N}^+ \to \mathbb{N}^+$ defined by $id(x) = x$ satisfies "$\Rightarrow$" but not "$\Leftrightarrow$" if we let $R = |$ (the divisibility relation and $S = \leq$).

3.2 Fact:

If $(A, <)$ and $(A', <')$ are linear orderings then a bijection $f : (A, <) \to (A', <')$ is an isomorphism iff for all $x, y \in A$

$x < y \Rightarrow f(x) <' f(y)$ (exercise)

3.3 Proposition:

Let $(A, <)$, $(A' <')$ be well-orderings and let $f : (A, <) \to (A', <')$ be an isomorphism. Then for each $x \in A$:

$$(*) f(x) = \text{the least element in } A' - f[(\leftarrow, x)] \text{ with respect to } <'$$

where $(\leftarrow, x) = \{z \in A | z < x\}$ is the initial segment below x in A and for $X \subseteq A$:

$$f[X] = \{f(z)|z \in X\} = \text{the point-wise image of } X \text{ under } f. \text{ (picture)}$$

So in particular, the isomorphism is *unique*. Also, this points out towards recursion: if we view f as defined recursively according to $<$ then $f(x) =$ the least element of A' under $<'$ that has not been used so far.

Proof. Assume not. So we have some $x \in A$ that violates $(*)$. Since $<$ is a well-ordering on A, there must be a least such. So assume x is the least such. Then:
—If $z < x$ then $f(z) <' f(x)$
—If $x < u$ then $f(x) <' f(u)$
(picture)
So in particular: $A' - f[(\leftarrow, x)] \ni f(x)$
Since we are assuming $f(x)$ is not the $<'$ −least element in $A' - f[(\leftarrow, x)] \ni f(x)$, we can find some $y \in A' - \delta[(\leftarrow, x)]$ such that $y <' f(x)$. Then $y \notin \text{rng}(f)$. This is because clearly $y \notin f[(\leftarrow, x)]$ by definition, and if $u \geq x$ then $f(u) \geq' f(x) >' y$. So $y \notin \text{rng}(f)$. Contradiction as f must be surjective. □

3.4 Example:
The conclusion in 3.3 heavily depends on the fact that we worked with well-orderings.
For instance, $(\mathbb{Q}, <)$ with natural ordering is not a well-ordering (least elements might not even exist). There are many isomorphisms $f : (\mathbb{Q}, <) \to (\mathbb{Q}, <)$ ($x \mapsto x$, $x \mapsto 2x$, etc.)

3.5 Proposition:
If $(A, \leq)$ is a well-ordered set and $A' \subseteq A$ then $<$ is a well-ordering on A', i.e. $(A', < \cap (A' \times A'))$ is a well-ordered set.
Proof: Exercise.
Remark: In future I will write briefly $(A', <)$ instead of $(A', < \cap (A' \times A'))$.

3.6 Proposition:
(a) If $(A, \leq)$ is a well-ordering and $f : A \to A$ is an order-preserving map, i.e.:

$$x < y \to f(x) < f(y) \implies f(x) \geq x \text{ for all } x \in A.$$

(b) If $(A, \leq)$ is a well-ordering then $(A, <)$ is not isomorphic to any of its *proper* initial segments.
Proof: Exercise. (b) is a direct consequence of (a)

3.7 Proposition:
Assume $(A, \leq)$ and $(A', \leq')$ be two well-orderings. Then one of them is isomorphic to an initial segment of the other.

Proof: (Naive sketch—compare with construction by recursion on ω) Strategy: we look at all isomorphisms of initial segments of $(A, \leq)$ onto initial segments of $(A', \leq')$. We show that these isomorphisms cohere. Then we "glue" them together.

Claim 1: Assume I, J are initial segments of $(A, \leq)$ and I', J' are initial segments of $(A', \leq')$ and $f : (I, <) \to (I', <')$, $g : (J, <) \to (J', <')$ are isomorphisms. Then f, g cohere, i.e. $f(x) = g(x)$ whenever $x \in I \cap J$.

Proof. By proposition 3.1 notice that $(\leftarrow, x) \subseteq I \cap J$ whenever $x \in I \cap J$.

$$f(x) =<' -\text{least element of } I' - f[(\leftarrow, x)].$$

But that is the same as the $<' -$ least element of $A' - f[(\leftarrow, x)]$ because I' is an initial segment of A'.

Similarly for $g(x)$; $g(x) =<' -$least element of $A' - g[(\leftarrow, x)]$. Since this characterization determines the map uniquely (By 3.1), we have $f(x) = g(x)$. □

In particular: If $f : (I, <) \to (I', <')$ is an isomorphism where I, I' are as in Claim 1, then f is unique.

So let $\bar{A} = \{x \in A$ — there is an isomorphism f of the interval $(\leftarrow, x)$ onto an initial segment of $(A', \leq')\}$ The we have a map $x \mapsto f_x$ which to each x assigns the unique isomorphism f_x of $(\leftarrow, x)$ onto an initial segment of $(A', \leq')$. Then, by Replacement and Separation:

$\mathcal{F} = \{f | (\exists x \in A)(f$ is an isomorphism of $((\leftarrow, x), <)$ onto an initial segment of $(A', \leq')\})$ is a set: By Claim 1, this is a coherent set of maps, so if we let $F = \bigcup \mathcal{F}$ then F is a map, and by definition F is an isomorphism of an initial segment of $(A, \leq)$ onto an initial segment of $(A', \leq')$. (Check this).

Claim 2: If $\operatorname{dom}(F) = A$, then F is an isomorphism of $(A, \leq)$ onto an initial segment of $(A', \leq')$. Obvious.

Claim 3: If $\operatorname{dom}(F) \neq A$ then $\operatorname{rng}(F) = A'$.

(picture)

Proof. Let x be the $< -$least element of $A-$ $\operatorname{dom}(F)$.

If $\operatorname{rng}(F) \neq A'$, let y be the $<' -$least element of $A'-$ $\operatorname{rng}(F)$. Then let $F' = F \cup \{\langle x, y\rangle\}$ —In other words $\operatorname{dom}(F') = \operatorname{dom}(F) \cup \{x\}$ and

$$F'(z) = \begin{cases} F(z) & \text{for} \quad z < x \\ y & \text{for} \quad z = x \end{cases}$$

Notice: $F' \in \mathcal{F}$, because obviously $F \in \mathcal{F}$ and by the definition of F'. (Check this). But this is not possible since $F' \supsetneq F$ and $F = \bigcup \mathcal{F}$.

Claim 4: If $\operatorname{dom}(F) \neq A$. Then $G = F^{-1} = \{\langle y, x\rangle | \langle x, y\rangle \in F\}$ is an isomorphism of the initial segment $(A', <')$ onto an initial segment of $(A, \leq)$. (Check this)

□

3.8 Remark:

Define the following binary relation on well-orderings:

$$(A, \leq) \sim (A', \leq') \text{ iff } (A, \leq), (A', \leq') \text{ are isomorphic.}$$

Easy to see: $\sim$ is an equivalence relation, and $\sim$ is a class because $\sim$ can be described in the language of set theory.

Notice: the equivalence class $[(A, \leq)]_\sim$ is a proper class: For any $a \in V$ the well-ordering $(A \times \{a\}, <_a)$ defined by

$$(\langle z, a\rangle <_a \langle z', a\rangle) \text{ iff } z < z'$$

is isomorphic to $(A, \leq)$.

So if we form the quotient: (Well-orderings/ $\sim$) then the elements of this quotient are proper classes, so this quotient is not a class. Our language of set theory does not have enough power to describe this quotient.

We can also define an "ordering" $<^*$ on this quotient by

$$[(A, \leq)] <^* [(A', \leq')] \text{ iff } (A, \leq) \text{ is isomorphic to a proper initial segment of } (A', \leq').$$

Easy to see that the definition of $<^*$ is correct, i.e. it does not depend on the choice of representatives.

By 3.6 (b): $<^*$ is irreflexive. Easy: $<^*$ transitive. By 3.7: $<^*$ is trichotomic. Hence, $<^*$ is a linear ordering of all equivalence classes $(A, \leq)$. With a little work, we can show that $<^*$ is a well-ordering.

Problem: None of these can be formulated with the language of set theory.
Solution: We show: there is a canonical way how to pick representatives from each equivalence class. These representatives are called *ordinals*.

3.9 Definition: Ordinal

An ordinal is a transitive set well-ordered by the relation $\in$.

Examples: $0, 1, 2, 3, ..., n, ..., \omega$.

3.10 Proposition:

Let x, y be ordinals. If $x \subseteq y$ then $x \in y$ or $x = y$. (Actually, we have "if and only if" here.)

Proof. Assume $x \neq y$. We show $x \in y$. Notice: $y - x \neq \varnothing$. So let $x' =$ the $\in$ least element of $y - x$. We show $x = x'$. This will do the job, as $x' \in y$.

$x' \subseteq x$: Notice $x' \subseteq y$ because $x' \in y$ and y is transitive. Because x' is the $\in$- least element of $y - x : x' \cap (y - x) = \varnothing$.

So $x' \subseteq y - (y - x) = x$.

Now to prove $x \subseteq x'$: Let $z \in x$. Because $x \subseteq y : z \in y$. Because $x' \in y$ and y is linearly ordered by $\in$, we have: $z \in x'$ or $z = x'$ or $x' \in z$.

If $z = x'$ then $x' \in x$. Not possible since $x' \in (y - x)$.

If $x' \in z$ then $x' \in z \in x$. Since x is an ordinal, $z \subseteq x$, so again $x' \in x$. In either case, we get $x' \in x$, which is a contradiction as $x' \in (y - x)$ by definition. So $z \in x'$ must hold, and $x \subseteq x'$. □

3.11 Proposition:

Let x, y be ordinals. Then $x \in y$ or $x = y$ or $y \in x$.

Proof. Notice: $x \cap y$ is an ordinal: $x \cap y$ is transitive, as both x, y are. Also, $x \cap y \subseteq x$ so $\in$ is a well-ordering on $x \cap y$, as it is a well-ordering on x.

We show: $x \cap y = x$ or $x \cap y = y$. This will do the job: If $x \cap y = x$ then $x \subseteq y$, so by 3.10, either $x \in y$ or $x = y$.

Assume $y \neq x \cap y \neq x$. This means that $x - (x \cap y) \neq \varnothing$ (same as $x - y$). So we can let y' be the $\in$-least element in $x - y$ and let

$$x' = \text{the } \in\text{-least element in } y - x.$$

By the proof of 3.10, $y' = x \cap y \subseteq y$ and $x' = x \cap y \subseteq x$. Here, $x \cap y$ plays the role of x' in the proof of 3.10.

So $x' = y'$. However: $x' \in y - x$ and $y' \in x - y$, so $x' \neq y'$ as $(y - x) \cap (x - y) = \varnothing$. Contradiction. □

3.12 Proposition:

$\in$ is a well-ordering on ordinals. In fact: if $A \neq \varnothing$ is any class of ordinals then A has an $\in$-least element.

Proof. Since we assume $A \neq \varnothing$, there is some ordinal $a \in A$. If $a \cap A = \varnothing$ then a is the $\in$-least element of A, so we are done.

If $a \cap A \neq \varnothing$ then $a \cap A$ is a set, by Separation: intersection of a set with a class is a set. Of course $a \cap A \subseteq a$. Since a is an ordinal, there is an $\in$-least element of $a \cap A$, call it a^*.

Claim: $a^* \cap A = \varnothing$, i.e. a^* is the $\in$-least element of A.

$a^* \cap (a \cap A) = \varnothing$ by the choice of a^*. And $a^* \cap (A - a) = \varnothing$ because $a^* \in a$, so $a^* \subseteq a$ since a is transitive. Now: $A = (a \cap A) \cup (A - a)$, so $a^* \cap A = \varnothing$. This verifies the minimality requirement.

We still have to see that $\in$ is a strict linear ordering on ordinals:

Irreflexivity: $x \notin x$ for each ordinal x, otherwise $\{x\} \subseteq x$ would be a set without an $\in$-least element.

Transitivity: $x \in y \in z$: because z is transitive: $x \in z$.

Linearity (Trichotomicity): Follows from 3.11.

□

3.13 Proposition:

If x is an ordinal and $x' \in x$ then x' is an ordinal.

Proof. $x' \subseteq x$ since x is transitive. Hence, $\in$ is a well-ordering on x'. So it suffices to check that x' is transitive. So let $u \in z \in x'$. We want to see that $u \in x'$. Since $x' \subseteq x$, $u \in z \in x$. Hence $u \in x$. So: $u, x' \in x$. By trichotomicity of $\in$ on x: either $u \in x'$ or $u = x'$ or $x' \in u$. Want to see $u \in x'$. If $u = x'$: $u \in z \in x' = u$. If $x' \in u : u \in z \in x' \in u$. In either case: $u \in u$ because $\in$ is a linear ordering on x. But this is a contradiction, as $\in$ is irreflexive on x.

□

3.14 Definition + Proposition:

We let $O_n = \{x | x \text{ is an ordinal}\}$. Then O_n is a class. (x is an ordinal $\iff$ x is a transitive *set* well-ordered by $\in$).

Proof: exercise.

3.15 Proposition:

O_n is a transitive *class* well-ordered by $\in$.

Proof: This is just P.3.13 and P.3.12.

3.16 Corollary:

O_n is a proper class.

Proof: If O_n were a set then $O_n \in O_n$. Contradiction to the fact that it is well-ordered.

(Picture of O_n in V–straight line in a messy universe V)

3.17 Proposition:

(a) If $x \in O_n$ then $S(x) = x \cup \{x\}$ is also an ordinal and it is the least ordinal larger than x. We will write $x+1$ instead of $S(x)$. We call $S(x) = x+1$ the (ordinal) successor of x.

(b) If A is a set of ordinals then $\bigcup A$ is an ordinal and it is the least upper bound on A. That is:

(i) $x \in \bigcup A$ or $x = \bigcup A$ for all $x \in A$ (i.e. $\bigcup A$ is an upper bound— especially if we think of $\in$ as $<$).

(ii) If $y \in O_n$ and (i) holds for y in place of x then $\bigcup A \in y$ or $\bigcup A = y$. (i.e. $\bigcup A$ is the *least* upper bound).

In the future we will write $\sup(A)$ instead of $\bigcup A$ and call this ordinal the *supremum* of A.

(c) If A is a proper class of ordinals then $\bigcup A = O_n$. Proof: Homework.

3.18 Remark:

This gives us another proof that O_n is a proper class: if O_n were a *set*, i.e. an ordinal, then $\bigcup O_n$ would be a set (an ordinal), as well as $S(\bigcup O_n)$ would be a set (an ordinal), as $S(\bigcup A)$ would be larger than all ordinals. This is a problem.

3.19 Theorem: Recursion on Ordinals, transfinite recursion

Let $G : V \to V$ be a class function. (Recall: G tells us what to do at the "next step" of recursion)

Then there is a unique class function $F : O_n \to V$ such that for all $x \in O_n$:

$$(*) F(x) = G(F \restriction x)$$

Proof. Uniqueness:

Assume both F and F' satisfy $(*)$. We show $F = F'$. Assume not. Then

$A = \{x \in O_n | F(x) \neq F'(x)\} \neq \varnothing$. So A is a nonempty class of ordinals. Therefore it has an $\in$ −least element, call it a. Hence: if $z \in a$ then $z \in O_n$ but $z \notin A$ by the $\in$ −minimality of a. Hence $F(z) = F'(z)$ for all $z \in a$, i.e. $F \restriction a = F' \restriction a$.

Then by $(*)$: $F(a) = G(F \restriction a) = G(F' \restriction a) = F'(a)$ where the first and last equalities are by $(*)$ and the second is because $F(z) = F'(z)$. Contradiction—this proves the uniqueness of F.

Taking the least element was important— this is where the well-ordering of ordinals came in.

Existence:

Let $\mathcal{F} = \{f|f$ is a set function, $\operatorname{dom}(f) \in O_n$ and $(*)$ holds for f in place of $F\}$.

Then $\mathcal{F}$ is a class (exercise). Intuitively, $\mathcal{F}$ is the class of all approximations to F.

Claim 1:

Assume f, $f' \in \mathcal{F}$. Then f, f' cohere, i.e. $f(z) = f'(z)$ for all $z \in \operatorname{dom}(f) \cap \operatorname{dom}(f')$.

Proof: Almost the same as the proof of uniqueness: Assuming that $f(z) \neq f'(z)$ for some z as above, look at the $\in -$ least such z. Also notice: if $z \in \operatorname{dom}(f) \cap \operatorname{dom}(f')$ then $z \subseteq \operatorname{dom}(f) \cap \operatorname{dom}(f')$ as $\operatorname{dom}(f)$, $\operatorname{dom}(f')$ are ordinals. (Exercise).

Recall: we have a class function $G : V \to V$ and want to construct $F : O_n \to V$ such that $F(x) = G(F \upharpoonright x)$. We defined $\mathcal{F} = \{f|f$ is a function and $\operatorname{dom}(f) \in O_n$ and $(\forall x \in \operatorname{dom}(f))(f(x) = G(f \upharpoonright x))\}$

Claim 1: If $f, f' \in \mathcal{F}$ then f, f' cohere, i.e. $f(z) = f'(z)$ for all $z \in \operatorname{dom}(f) \cap \operatorname{dom}(f')$.

In particular, for each $x \in O_n$ there is at most one $f \in \mathcal{F}$ such that $\operatorname{dom}(f) = S(x) = x \cup \{x\}$.

Claim 2: (Restated- this is just a cosmetic change): For each $x \in O_n$ there is an $f \in \mathcal{F}$ with $\operatorname{dom}(f) = S(x)$.

Suppose not; then we have an $\in -$least counterexample, call it a. So for each $x \in a$ there is an $f \in \mathcal{F}$ with $\operatorname{dom}(f) = S(x)$. By the remark after Claim 1, this function is unique. By Replacement + Separation, there is a set $\mathcal{F}_a$ such that

$$\mathcal{F}_a = \{f \in \mathcal{F}|(\exists x \in a)(\operatorname{dom}(f) = S(x))\}.$$

Even though $\mathcal{F}$ is a class, we know this is a set by the two axioms above.

Since $\mathcal{F}_a \subseteq \mathcal{F}$, all functions in $\mathcal{F}_a$ cohere, so $\bar{f} = \bigcup \mathcal{F}_a$ is a function. Moreover: By assumption to each $x \in a$ there is $f \in \mathcal{F}_a$ with $\operatorname{dom}(f) = S(x) = x \cup \{x\}$. Hence: Each $x \in a$ is in the domain of $\bar{f}$, i.e. $a \subseteq \operatorname{dom}(\bar{f})$. Moreover: if $x \in \operatorname{dom}(\bar{f})$ then pick $f \in \mathcal{F}_a$ with $\operatorname{dom}(f) = S(x)$. Then $f \subseteq \bar{f}$. Hence:

$$\bar{f}(x) = f(x) = G(f \upharpoonright x) = G(\bar{f} \upharpoonright x) \ (*)$$

where the first equality follows from $f \subseteq \bar{f}$, second since $f \in \mathcal{F}$, and the last since $f \upharpoonright x = \bar{f} \upharpoonright x$.

Conclusion: $\bar{f} \in \mathcal{F}$.

Now if $a \in \operatorname{dom}(\bar{f})$ we have a contradiction since $\bar{f} \upharpoonright S(a) \in \mathcal{F}$ by $(*)$. This is a contradiction, as we assumed that a was a counterexample, i.e. there is no $h \in \mathcal{F}$ with $\operatorname{dom}(h) = S(a)$. If $a \notin \operatorname{dom}(\bar{f})$. Then $\operatorname{dom}(\bar{f}) = a$ In this case let $f^* = \bar{f} \cup \{\langle a, G(\bar{f})\rangle\}$.

In other words, $f^* : S(a) \to V$ defined by

$$f^*(x) = \begin{cases} \bar{f}(x) & \text{for} \quad x \in a \\ G(\bar{f}) & \text{for} \quad x = a \end{cases}$$

Then $f^* \in \mathcal{F}$ and $\operatorname{dom}(f^*) = S(a)$; contradiction as in before.

Claim 3: To each $x \in O_n$ there is exactly one $f \in \mathcal{F}$ such that $\operatorname{dom}(f) = S(x)$.

Now let $F = \bigcup \mathcal{F}$. By Claims 1 and 2: F is a function and $\operatorname{dom}(F) = O_n$. By definition, F is a class. It remains to see that F satisfies the recursion requirement. This is like in Claim 2.

Let $x \in O_n$. By Claim 2 there is some $f \in \mathcal{F}$ with $\operatorname{dom}(f) = S(x)$. Notice $f \subseteq F$ i.e. $f(x) = F(x)$ for all $x \in \operatorname{dom}(f(x))$.

Then $F(x) = f(x) = G(f \upharpoonright x) = G(F \upharpoonright x)$ $(f \subseteq F; f \in \mathcal{F}; f \subseteq F$, so $f \upharpoonright x = F \upharpoonright x)$.

□

SOME APPLICATIONS OF RECURSION

3.20 Proposition + Definition

If $(A, \leq)$ is a well-ordered set then there is exactly one ordinal α such that $(A, \leq)$ is isomorphic to $(\alpha, \in)$, and the isomorphism is unique. We call this ordinal the *order-type* of $(A, \leq)$ and denote by $\operatorname{otp}((A, \leq))$, or simply by $\operatorname{otp}(A)$ if the ordering is clear from the context.

Proof. The uniqueness of the ordinal follows from P.3.6 which says that no well-ordering is isomorphic to its proper initial segment: if $(A, \leq) \cong (\alpha, \in)$ and $(A, \leq) \cong (\beta, \in)$ and $\alpha \neq \beta$ then $\alpha \in \beta$ or $\beta \in \alpha$. But obviously $(\alpha, \in) \cong (\beta, \in)$. The uniqueness of the isomorphism follows from P.3.3. So we have to verify the existence.

Idea: We attempt to define an isomorphism from some $(\alpha, \in)$ onto $(A, \leq)$ but we don't know in advance what α is. So instead of α we use O_n and we need a tool that will tell us where to stop.

Rigorously: Pick some $s \in A$. s will be a "sentinal." We define

$$f : O_n \to A \cup \{s\}$$

by recursion:

$$f(x) = \begin{cases} \text{the } < -\text{least element of } A - f[x] & \text{for} \quad A - f[x] \neq \varnothing \\ \qquad\qquad\qquad\qquad s & \text{otherwise} \end{cases}$$

Let α= the least ordinal x such that $f(\alpha) = s$. We have to show that α exists, i.e. that $f(x) \neq A$ for some ordinal x. Then we show $f \restriction \alpha : (\alpha, \in) \to (A, \leq)$ is an isomorphism.

10/28/09

Our function G in the theorem on recursion is $G(u)$ =the least element in $A - f[u]$ if this set is nonempty and is s otherwise. This is the last explanation of G that will be given. Also, $G(f \restriction x) = A-$ $\operatorname{rng}(f \restriction x)$.

Claim 1: If $x < y$ are ordinals and $f(x), f(y) \in A$ then $f(x) \neq f(y)$. Why: $f(y)$ is the $< -$least element in $A - f[y]$. But $f(x) \in f[y]$, as $x \in y$. Hence $f(x) \notin A - f[y]$.

Claim 2: There is an $x \in O_n$ such that $f(x) = s$. (Informally: at some point we run out of all elements of A).

Why: if not, then $f : O_n \to A$ and by Claim 1 is injective. By the theorem on recursion: f is a class, hence $A' = \operatorname{rng}(f)$ is a set, since A is a set by Separation. Then $f^{-1} : A \to O_n$ is a class function, so by Replacement $\operatorname{rng}(f^{-1})$ must be a set. However: $\operatorname{rng}(f^{-1}) = O_n$. Contradiction.

Let $a =$ the least ordinal x such that $f(x) = s$. That is, $f(a) = s$ and $f(z) \in A$ for all $z \in a$.

Notice: $f[a] = A$, as otherwise since $f(a) = s$ we have $A - f[a] = \varnothing$. So $f[a] = A$. So far we have: f maps a bijectively onto A. To see that $f \restriction a$ is an isomorphism between $(a, \in)$ and $(A, \leq)$, it suffices to show:

Claim 3: If $x, y \in a$ and $x \in y$ then $f(x) < f(y)$.

Assume $x \in y$. Then $x \subset y$ so $f[x] \subset f[y]$, so $A - f[x] \supset A - f[y]$. Hence:

$f(x) =$ (the $<$ −least element of $A - f[x]$) is smaller than the $<$ −least element of $A - f[y]$ which is $f(y)$.

□

Proposition 3.20 gives us a definable way how to pick representatives of well-orderings. To each well-ordered set $(A, \leq)$ there is a unique ordinal a such that $(A, \leq) \cong (a, \in)$. This ordinal is called the order-type of $(A, \leq)$ and is denoted by $\operatorname{otp}((A, \leq))$. If the ordering $<$ is clear from the context, we write $\operatorname{otp}(A)$. The function otp: $\{x | x$ is a well-ordered set$\} \rightarrow O_n$ assigning $\operatorname{otp}((A, \leq))$ to each well ordered set $(A, \leq)$ is a class function.

3.21 Convention:

In literature, ordinals are usually denoted by lower case greek letters: $\alpha, \beta, \gamma, ...$
Also, we write $\alpha < \beta$ instead of $\alpha \in \beta$.

3.22 Definition: Successor Ordinals/Limit Ordinals

Ordinals of the form $S(\alpha)$ are called *successor ordinals*. Ordinals that are not of this form are called *limit ordinals*. In particular, 0 is considered a limit ordinal.

3.23 Remark:

Both successor and limit ordinals constitute proper classes.
If α is a limit ordinal and $\bar{\alpha} < \alpha$ then $S(\bar{\alpha}) < \alpha$.

ORDINAL ARITHMETIC

3.24 Definition:

We define an operation $+$ on ordinals by recursion as follows:
$\alpha + 0 = \alpha$
$\alpha + S(\beta) = S(\alpha + \beta)$
$\alpha + \beta = \sup\{\alpha + \bar{\beta} | \bar{\beta} < \beta\}$

In the future, we will frequently do definitions by recursion this way: We specify the definition for 0, successor steps and limit steps separately. The map $+ : O_n \times O_n \rightarrow O_n$ is a class.

For each α we define a map $+_\alpha : O_n \rightarrow O_n$ by $+_\alpha(\beta) = \alpha + \beta$. Notice $+_\alpha$ is a proper class.

Note: $1 + \omega = \omega$ but $\omega + 1 > \omega$.

Ordinal addition defines a class function $+ : O_n \times O_n \rightarrow O_n$.

If we let $+_\alpha : O_n \times O_n \rightarrow O_n$ be the function defined by $+_\alpha(\beta) = \alpha + \beta$ then $+_\alpha$ is a proper class. So the collection of all functions $+_\alpha$ is not a class. In order to see that $+$ is a class, we use the method of "truncation." For each $\theta \in O_n$ we define function $+^\theta : \theta \times \theta \rightarrow O_n$ so that the above holds with $+^\theta$ in place of $+$ for all $\alpha, \beta < \theta$. Then $+^\theta$ is a set. This can be proved as in the case of ω:

— By construction by recursion: for each $\alpha < \theta$ there is a unique function $+^\theta_\alpha : \theta \rightarrow O_n$ such that:

$$+^\theta_\alpha = \alpha$$

$$+^{\theta}_{\alpha}(S(\beta)) = S(+^{\theta}_{\alpha}(\beta)) \text{ and}$$

$$+^{\theta}_{\alpha}(\beta) = \sup\{+^{\theta}_{\alpha}(\bar{\beta}|\bar{\beta} < \beta\}).$$

By Replacement and Separation,: Each function $+^{\theta}_{\alpha}$ is a set. Then by Replacement and Separation:

$$A_{\theta} = \{+^{\theta}_{\alpha}|\alpha < \theta\}$$

is a set. Hence $+^{\theta} = \bigcup A_{\theta}$ is a set.

Notice: $+^{\theta}$ is unique satisfying $(*)$ (i.e. if we replace $+$ with $+^{\theta}$ for all $\alpha, \beta < \theta$, as all $+^{\theta}_{\alpha}$ are unique).

Hence: if $\theta < \theta'$ then $+^{\theta} = +^{\theta'} \restriction (\theta \times \theta)$ or in other words, $+^{\theta} \subseteq +^{\theta'}$.

Easy to check: the function $\theta \mapsto +^{\theta}$ is a class (i.e. $\{+^{\theta}|\theta \in O_n\}$ is a class) and
$+ = \bigcup\{+^{\theta}|\theta \in O_n\}$ is a class.

Notice: $+ \restriction \omega \times \omega$ is the usual addition of natural numbers, so one can prove it is commutative. However, $+$ in general is *not* commutative: $1 + \omega = \omega < \omega + 1$.

3.25 Proposition:
$+$ is associative: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for all $\alpha, \beta, \gamma \in O_n$.

Proof. Induction on γ
— $\gamma = 0$: $(\alpha + \beta) + 0 = \alpha + \beta = \alpha + (\beta + 0)$.
— Successors: $(\alpha + \beta) + S(\gamma) = S((\alpha + \beta) + \gamma) = S(\alpha + (\beta + \gamma)) = \alpha + (\beta + S(\gamma))$.
— Limits: $(\alpha+\beta)+\gamma = \sup\{(\alpha+\beta)+\bar{\gamma}|\bar{\gamma} < \gamma\} = \sup\{\alpha+(\beta+\bar{\gamma})|\bar{\gamma} < \gamma\} = \sup\{\alpha+\delta|\delta < \beta+\gamma\} = \alpha + (\beta + \gamma)$. (By $(*)$, I.H., (1), and $(*)$).

To see (1): $\leq$ follows from the fact that $\{\alpha+(\beta+\bar{\gamma})|\bar{\gamma} < \gamma\} \subseteq \{\alpha+\delta|\delta \leq \beta+\gamma\}$, as $\beta+\bar{\gamma} \leq \beta+\gamma$ for all $\bar{\gamma} < \gamma$.

Next: $\geq$: If $\delta < \beta + \gamma = \sup\{\beta + \bar{\gamma}|\bar{\gamma} < \gamma\}$ then there is $\gamma' < \gamma$ such that $\delta \leq \beta + \gamma'$. Hence $\delta < S(\beta+\gamma') = \beta + S(\gamma')$ and if we say that $S(\gamma') = \bar{\gamma}$ then $\bar{\gamma} < \gamma$ because γ is a limit. So for each $\delta < \beta + \gamma$ we can find $\bar{\gamma} < \gamma$ such that $\delta < \beta + \bar{\gamma}$.

□

3.26 Lemma:
For each $\alpha, \beta \in O_n$: $\alpha \leq \alpha + \beta$ and $\beta \leq \alpha + \beta$. Also: if $\beta > 0$ then $\alpha < \alpha + \beta$.
Proof: By induction on β similarly as in the proof of P.3.25.

3.27 Lemma
For each $\alpha, \alpha', \beta, \beta' \in O_n$:
(a) If $\alpha \leq \alpha'$ then there is a unique $\tilde{\alpha}$ such that $\alpha' = \alpha + \tilde{\alpha}$. Moreover, if $\alpha < \alpha'$ then $\tilde{\alpha} > 0$.
(b) If $\beta < \beta'$ then $\alpha + \beta < \alpha + \beta'$.

Proof. (a) We first prove existence. By L.3.26 $\alpha' \leq \alpha + \alpha'$. Let $\tilde{\alpha}$ be the least $\xi \in O_n$ such that $\alpha' \leq \alpha + \xi$. $\tilde{\alpha}$ exists as $\{\xi|\alpha' \leq \alpha + \xi\} \neq \varnothing$ by the above derivation. Then $\alpha' = \alpha + \xi$.

Why: If not, then either $\tilde{\alpha}$ is a successor, say $\tilde{\alpha} = S(\gamma)$. By the definition of $\tilde{\alpha}$: $\alpha + \gamma < \alpha'$. Hence $S(\alpha+\gamma) \leq \alpha'$. But $S(\alpha+\gamma) = \alpha+S(\gamma) = \alpha+\tilde{\alpha}$. So $\alpha+\tilde{\alpha} \leq \alpha'$. By we also know $\alpha' \leq \alpha+\tilde{\alpha}$ for all $\gamma < \tilde{\alpha}$. Contradiction.

Or else: $\tilde{\alpha}$ is a limit. Again by def of $\tilde{\alpha} : \alpha + \gamma < \alpha'$ for all $\gamma < \tilde{\alpha}$. Now $\alpha + \tilde{\alpha} \sup\{\alpha + \gamma | \gamma < \tilde{\alpha}\}$. As before: $\alpha + \tilde{\alpha} = \alpha'$.

This ends the proof of existence.

It is also clear that if $\alpha < \alpha'$ then $\tilde{\alpha} > 0$. (If $\tilde{\alpha} = o$ then $\alpha + \tilde{\alpha} = \alpha + 0 = \alpha < \alpha'$).

Now prove (b): Assume $\beta < \beta'$. By what we have proved: there is $\tilde{\beta} > 0$ such that $\beta' = \beta + \tilde{\beta}$. Hence $\alpha + \beta < (\alpha + \beta) + \tilde{\beta} = \alpha + (\beta + \tilde{\beta}) = \alpha + \beta'$.

Finally, prove uniqueness in (a): Assume we have $\tilde{\alpha_1}, \tilde{\alpha_2}$ such that $\alpha + \tilde{\alpha_1} = \alpha' = \alpha + \tilde{\alpha_2}$. If $\tilde{\alpha_1} \neq \tilde{\alpha_2}$, say $\tilde{\alpha_1} < \tilde{\alpha_2}$, then by (b): $\alpha + \tilde{\alpha_1} < \alpha + \tilde{\alpha_2}$. Contradiction.

□

From what we had:
$S(\alpha) = S(\alpha + 0) = \alpha + S(0)$. But $S(0) = 1 \in \omega$.
So from now on we write "$\alpha + 1$" in place of $S(\alpha)$.

Recall: From now on we write $\alpha + 1$ for $S(\alpha)$.
Then the recursive definition of ordinal addition reads:

$\alpha + 0 = \alpha$
$\alpha + (\beta + 1) = (\alpha + \beta) + 1$
$\alpha + \beta = \sup\{\alpha + \bar{\beta} | \bar{\beta} < \beta\}$

Analogously we introduce ordinal multiplication and exponentiation:

$\alpha \cdot 0 = 0$
$\alpha \cdot (\beta + 1) = \alpha \cdot \beta + \alpha$
$\alpha \cdot \beta = \sup\{\alpha \cdot \bar{\beta} | \bar{\beta} < \beta\}$ and

$\alpha^0 = 1$
$\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$
$\alpha^\beta = \sup\{\alpha^{\bar{\beta}} | \bar{\beta} < \beta\}$

3.28 Proposition:

The following holds of ordinal operations.

(a) Ordinal addition and multiplication are associative, but no commutative. Their restrictions to ω are commutative.

(b) $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$
(c)$\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$
(d)$(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$
(e) $\beta < \beta'$ implies the following:

—$\alpha + \beta < \alpha + \beta'$
—$\alpha \cdot \beta < \alpha \cdot \beta'$ for $\alpha > 0$

—$\alpha^{\beta} < \alpha^{\beta'}$ for $\alpha > 1$

Proof: The proofs of these facts use either methods similar to those we used in the case of addition, or else these facts follow from the rest of them by direct computation.

3.29 Remark:

Ordinal arithmetic can be used to "code" proofs and thereby analyze their complexity. An important ordinal is

$$\epsilon_0 = \sup\{\omega, \omega^{\omega}, ..., \omega^{\omega^{\cdots^{\omega}}}, ...\}$$

which codes all proofs in Peano Arithmetic.

3.30 Proposition:

Let $\beta \geq 2$. Then for every ordinal ξ there are unique ordinals α, ρ such that $\xi > \beta$ and
$\xi = \beta \cdot \alpha + \rho$.

Proof. Existence: Fix ξ. Let $\alpha^* = \min\{\alpha' \in O_n | \beta \cdot \alpha' > \xi\}$. α^* exists as the class on the right side is nonempty. Notice $\beta \cdot \xi \geq \xi$ for all $\xi \in O_n$. (Proof: easy induction on ξ: easy exercise— so $\beta \cdot (\xi + 1) \geq \xi + 1 \geq \xi$).

Then notice: α^* is a successor ordinal. Why: by definition: $\beta \cdot \bar{\alpha} \leq \xi$ for all $\bar{\alpha} < \alpha^*$. So if α^* were a limit then

$$\beta \cdot \alpha^* = \sup\{\beta \cdot \bar{\alpha} | \bar{\alpha} < \alpha^*\} \leq \xi \text{ because } \beta \cdot \bar{\alpha} \leq \xi.$$

But by definition, $\beta \cdot \alpha^* > \xi$. Contradiction.

So α^* is a successor ordinal, say $\alpha^* = \alpha + 1$. We then have:

$$\beta \cdot \alpha \leq \xi < \beta \cdot (\alpha + 1) = \beta \cdot \alpha + \beta.$$

By L.3.27 there is some ρ such that $\xi = \beta \cdot \alpha + \rho$. Now notice that $\rho < \beta$ because otherwise we would have $\rho \geq \beta$ and $\xi = \beta \cdot \alpha + \rho \geq \beta \cdot \alpha + \beta = \beta \cdot (\alpha + 1) = \beta \cdot \alpha^*$. But we know that $\beta \cdot \alpha^* > \xi$. This proves existence.

Uniqueness: Assume $\beta \cdot \alpha + \rho = \xi = \beta \cdot \alpha' + \rho'$ where $\rho, \rho' < \beta$. We show that $\alpha = \alpha'$ and $\rho = \rho'$.

If $\alpha \neq \alpha'$ then WLOG $\alpha < \alpha'$. But then $\xi = \beta \cdot \alpha + \rho < \beta\alpha + \beta = \beta \cdot (\alpha + 1) \leq \beta \cdot \alpha' \leq \beta\dot{\alpha}' + \rho' = \xi$. So we would get $\xi < \xi$. Contradiction. Hence $\alpha = \alpha'$ Similarly for $\rho :$ if $\rho \neq \rho'$ then WLOG $\rho < \rho'$, but then $\xi = \beta \cdot \alpha + \rho < \beta \cdot \alpha + \rho' = \xi$. Contradiction. □

3.31 Proposition: (Cantor normal form)

Let $\beta \geq 2$. Then for every $\xi \in O_n$ there are unique
— $n \in \omega$
— $\epsilon_1 > \epsilon_2 > ... > \epsilon_n$
— $\alpha_1, \alpha_2, ..., \alpha_n$ all $< \beta$ and all $\neq 0$.

such that $\xi = \beta^{\epsilon_1} \cdot \alpha_1 + \beta^{\epsilon_2} \cdot \alpha_2 + ... \beta^{\epsilon_n} \cdot \alpha_n$.

Proof. Claim 1: Assume $\epsilon_1 > \epsilon_2 > ... > \epsilon_n$ and $\alpha_2, \alpha_3, ..., \alpha_n < \beta$. Then

$$\beta^{\epsilon_1} > \beta^{\epsilon_2} \cdot \alpha_2 + \beta^{\epsilon_3} \cdot \alpha_3 + ... + \beta^{\epsilon_n} \cdot \alpha_n.$$

Proof: By induction on ϵ_1.

$\beta^{\epsilon_2} \cdot \alpha_2 + \beta^{\epsilon_3} \cdot \alpha_3 + ... + \beta^{\epsilon_n} \cdot \alpha_n < \beta^{\epsilon_2} \cdot \alpha_2 + \beta^{\epsilon_2} = \beta^{\epsilon_2} \cdot (\alpha_2 + 1) \leq \beta^{\epsilon_2} \cdot \beta = \beta^{\epsilon_2+1} \leq \beta^{\epsilon_2}$. Because $\epsilon_2 < \epsilon_1 \Rightarrow \epsilon_2 + 1 \leq \epsilon_1$.

This ends the proof of Claim 1.

Claim 2: $\xi \leq \beta^{\xi}$ whenever $\beta \geq 2$.

Proof: Easy induction on ξ. Exercise.

Now fix $\xi \in O_n$. Then $\xi < \xi + 1 \leq \beta^{\xi+1}$ Then there is a least ordinal ϵ such that $\xi < \beta^{\epsilon}$.

Using continuity of exponentiation, similarly as in the proof of P.3.30 we show that ϵ is a successor ordinal, say $\epsilon = \epsilon_1 + 1$. So we have $\beta^{\epsilon_1} \leq \xi \leq \beta^{\epsilon_1+1}$.

By P.3.30 there are unique α_1 and $\rho < \beta^{\epsilon_1}$ such that $\xi = \beta^{\epsilon_1} \cdot \alpha_1 + \rho$.

Notice $\alpha_1 < \beta_1$ as otherwise $\xi = \beta^{\epsilon_1} \cdot \alpha_1 + \rho \geq \beta^{\epsilon_1} \cdot \beta + \rho = \beta^{\epsilon_1+1} + \rho \geq \beta^{\epsilon} > \xi$.

Also $\alpha_1 \neq 0$ otherwise $\xi = \rho < \beta^{\epsilon_1}$ which is impossible by the minimality of ϵ.

Now $\rho < \beta^{\epsilon_1} \leq \xi$, so we can apply the induction hypothesis to ρ: There are $\epsilon_2 > ... > \epsilon_n$ and $\alpha_1, ..., \alpha_n < \beta$ and all $\neq 0$ such that

$$\rho = \beta^{\epsilon_2} \cdot \alpha_2 + ... + \beta^{\epsilon_n} \cdot \alpha_n$$

Notice: $\epsilon_2 < \epsilon_1$ since $\rho < \beta^{\epsilon_1}$ By substitution:

$\xi = \beta^{\epsilon_1} \cdot \alpha_1 + \rho = \beta^{\epsilon_1} \cdot \alpha_1 + \beta^{\epsilon_2} \cdot \alpha_2 + ... + \beta^{\epsilon_n} \cdot \alpha_n$ and ϵ_i , α_i are as desired.

11/04/09

We want to prove uniqueness: that is. given $\beta \geq 2$, $n, n' \in \omega$, ordinals $\alpha_1, ..., \alpha_n$, $\alpha'_1, ..., \alpha'_{n'}$ all less than β and all $\neq 0$, and descending sequences $\epsilon_1 > \epsilon_2 > ... > \epsilon_n$ and $\epsilon'_1 > ... > \epsilon'_{n'}$.

We are assuming:

$$(*)\ \beta^{\epsilon_1} \cdot \alpha_1 + \beta^{\epsilon_2} \cdot \alpha_2 + ... + \beta^{\epsilon_n} \cdot \alpha_n = \beta^{\epsilon'_1} \cdot \alpha'_1 + \beta^{\epsilon'_2} \cdot \alpha'_2 + ... + \beta^{\epsilon'_{n'}} \cdot \alpha'_{n'}.$$

Want to show that $n = n'$, $\alpha_i = \alpha'_i$ and $\epsilon_i = \epsilon'_i$ for all $i \in \{1, ..., n'\}$. WLOG: $n \leq n'$. Assume we have some i such that $\langle \epsilon_i, \alpha_i \rangle \neq \langle \epsilon'_i, \alpha'_i \rangle$; assume also that i is the least such. What if $\epsilon_i \neq \epsilon'_i$. WLOG assume $\epsilon_i < \epsilon'_i$.

$\beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot \alpha_i + \beta^{\epsilon_{i+1}} \cdot \alpha_{i+1} + ... + \beta^{\epsilon_n} \cdot \alpha_n < \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot \alpha_i + \beta^{\epsilon_i} = \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot (\alpha_i + 1)$ $\leq \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i+1} \leq \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon'_i} + \beta^{\epsilon'_{i+1}} \cdot \alpha'_{i+1} + ... + \beta^{\epsilon'_{n'}} = \beta^{\epsilon'_1} \cdot \alpha'_1 + ... + \beta^{\epsilon'_{n'}} \cdot \alpha'_{n'}$, since by choice of i: $(\epsilon_j, \alpha_j) = (\epsilon'_j, \alpha'_j)$ for all $j < i$.

Where $\beta^{\epsilon_{i+1}} \cdot \alpha_{i+1} + ... + \beta^{\epsilon_n} \cdot \alpha_n < \beta^{\epsilon_i}$ by Claim 1,

And $\alpha_i + 1 \leq \beta$.

And $\epsilon_i < \epsilon'_i$

This is a contradiction, since the two sums are assumed to be the same. Hence $\epsilon_i = \epsilon'_i$. Hence $\alpha_i \neq \alpha'_i$, so again assume WLOG that $\alpha_i < \alpha'_i$. Similarly we compute

$\beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot \alpha_i + \beta^{\epsilon_{i+1}} \cdot \alpha_{i+1} + ... + \beta^{\epsilon_n} \cdot \alpha_n < \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot \alpha_i + \beta^{\epsilon_i} = \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot (\alpha_i + 1)$ $\leq \beta^{\epsilon_1} \cdot \alpha_1 + ... + \beta^{\epsilon_i} \cdot \alpha'_i + ... + \beta^{\epsilon'_{n'}} \cdot \alpha_{n'} = \beta^{\epsilon'_1} \cdot \alpha'_1 + ... + \beta^{\epsilon'_{n'}} \cdot \alpha'_n$

Contradiction again.

Conclusion: $\epsilon_i = \epsilon'_i$ and $\alpha_i = \alpha'_i$ for all $i \in \{1, ..., n\}$. Now if $n < n'$ then the sum on the right in $(*)$ is a proper extension of the sum on the left. Since we assume that all $\alpha'_i \neq 0$, the sum on the right is larger than the sum on the left. Contradiction. So $n = n'$.

□

3.32 Definition + Proposition:

If $(A, \leq)$ is a linear ordering then the set of all open intervals in $(A, \leq)$ constitutes a base for a topology. This topology is called the *interval topology*; so the open sets are precisely unions of open intervals. All topological notions will be used in the usual sense, e.g. closed sets are complements of open sets; x is a limit point of $B \subseteq A$ iff every open interval containing x has a nonempty intersection with B. Also, B is closed iff it contains all its limit points. The notion of a continuous map is also standard: inverse images of open sets are open.

However: it is not always the case that we have a sequence $\langle b_n | n \in \omega \rangle$ of elements of B converging to x if x is a limit point of B.

3.33 Definition + Proposition:

(a) A class A of ordinals is closed iff for each $\alpha \in O_n$, α a limit ordinal: $A \cap \alpha$ is unbounded in $\alpha \Rightarrow \alpha \in A$. This is true because α is a limit point of A iff $A \cap \alpha$ is unbounded in A. Notice: if $\alpha < \beta$ then $(\alpha, \beta] = (\alpha, \beta + 1)$ is an open interval.

(b) Let α be a limit ordinal. A set $C \subseteq \alpha$ is closed unbounded in α iff C is unbounded in α and is closed in the interval topology of α. (Notice: in this case $\alpha \in \lim(C) - C$). I.e. C contains all its limit points that are $< \alpha$.

(c) Let α be a limit ordinal or $\alpha = O_n$. A class function $f : \alpha \to O_n$ is *normal* iff
(i) f is strictly increasing. I.e. $\xi < \xi' \Rightarrow f(\xi) < f(\xi')$
(ii) f is continuous, i.e. $f(\xi) = \sup\{f(\bar{\xi}) | \bar{\xi} < \xi\}$ whenever ξ is a limit ordinal.

(d) Ordinal addition, multiplication and exponentiation are normal as functions of the second argument.

(e) (i) If α, β are limit ordinals and $f : \alpha \to \beta$ is normal with $\operatorname{rng}(f)$ unbounded in β (we also say f is *cofinal* in β) then $\operatorname{rng}(f)$ is a closed unbounded subset of β.
(ii) If β is a limit ordinal and $C \subseteq \beta$ is a closed unbounded then there are an ordinal α and a normal function $f : \alpha \to \beta$ such that $\operatorname{rng}(f) = C$.
(iii) The obvious analogy is true of class function $f : O_n \to O_n$ and closed proper classes of O_n. (Homework)

The third application of recursion: The hierarchy V_α. This hierarchy is defined as follows.
$V_0 = \varnothing$
$V_{\alpha+1} = \mathcal{P}(V_\alpha)$
$V_\alpha = \bigcup_{\bar{\alpha}<\alpha} V_{\bar{\alpha}}$.
WF= $\bigcup_{\alpha \in O_n} V_\alpha$

Let us stress: We still work in ZF without Axiom of Foundation. By the theorem on recursion: the function $\alpha \mapsto V_\alpha$ is a class, each V_α is a set and also WF is a class. WF is called the *well-founded* core of V.

By easy recursion, we get: $\alpha \subseteq V_\alpha$ for all $\alpha \in O_n$. So $O_n \subseteq$ WF , so WF is a proper class.

By Homework + Recursion, all V_α are transitive. Also: WF is transitive.

By easy recursion, $\alpha < \beta \Rightarrow V_\alpha < V_\beta$ and $V_\alpha \subseteq V_\beta$. (Exercise.)

3.34 Proposition:
All sets in WF are well-founded. (This is why WF is called well-founded core.)

3.35 Theorem (ZF)
WF$=V$. (So Axiom of Foundation implies WF$=V$.)

Proof. We know WF$\subseteq V$. So assume WF$\neq V$.

By Axiom of Foundation, we would like to pick an $\in$ -minimal element in $V-$WF. However, $V-$WF may be a proper class, and Axiom of Foundation only guarantees the existence of $\in$ -minimal elements for sets. So we have to replace $V-$WF by a set.

We do the following: Pick $a \in V-$WF. We know $\operatorname{trcl}(\{a\})$ is a set and $a \in \operatorname{trcl}(\{a\})$ (in general $a \notin \operatorname{trcl}(a)$).

So $\operatorname{trcl}(\{a\})-$WF is a nonempty set because it contains a. By Axiom of Foundation, there is an $\in$ -minimal element of $\operatorname{trcl}(\{a\})-$WF, call it x.

Claim: $x \subseteq$ WF.

Proof: Pick $z \in x$. Because $x \in \operatorname{trcl}(\{a\})$ and $\operatorname{trcl}(\{a\})$ is a transitive set, we have $z \in \operatorname{trcl}(\{a\})$. However: x is an $\in$ -minimal element of $\operatorname{trcl}(\{a\})-$WF, hence $z \notin \operatorname{trcl}(\{a\})-$WF.

So: $z \in \operatorname{trcl}(\{a\})$ and $z \notin \operatorname{trcl}(\{a\})-$WF $\Rightarrow z \in$WF. End proof of claim.

Each $z \in x$ is in some set of the form V_α. So we can define a map $f : x \to O_n$ by

$$f(z) = \text{the least } \alpha \in O_n \text{ such that } z \in V_\alpha.$$

Then f is a class function. Since x is a set, by Replacement: $\operatorname{rng}(f)$ is a set. So we have some ordinal γ such that $\operatorname{rng}(f) \subseteq \gamma$. This means that each $z \in x$ is an element of some $\alpha < \gamma$, so $x \subseteq V_\gamma$. Hence, $x \in V_{\gamma+1} \subseteq$WF, i.e. $x \in$WF. Contradiction. □

3.36 Corollary + Definition: (ZF)
By the above, to each set x we can assign the least ordinal α such that $x \in V_{\alpha+1}$. This ordinal is called the rank of x and we denote it by $\operatorname{rank}(x)$. Examples:

$\operatorname{rank}(\varnothing) = 0$
$\operatorname{rank}(1) = 1$
$\operatorname{rank}(n) = n$ for all $n \in \omega$.
$\operatorname{rank}(\omega) = \omega$.

4. Axiom of Choice (ZF)

4.1 Definition: Selector
Let A be a set. A *selector on A* is a map $f : A \to \bigcup A$ such that

$$f(a) \in a \text{ whenever } a \neq \varnothing.$$

Intuitively: A selector on A selects an element from each nonempty $a \in A$.

4.2 Proposition:

The following are equivalent:

(a) Axiom of Choice (AC)
(b) On every set there is a selector.

Remark: the formulation of AC required $\varnothing \notin A$ and all elements of A are mutually disjoint. In P.4.2(b) we don't require this.

Proof: Homework.

Week 7

Proposition: AC is equivalent to the statement that every set has a selector.
Selector is a function $f : A \to \bigcup A$ such that $f(a) \in a$ for every $a \neq \varnothing$, $a \in A$.

4.3 Definition: Uniformization
Let $R \subseteq A \times A$. A function f *uniformizes* R iff $\operatorname{dom}(f) = \operatorname{dom}(R)$ and $f \subseteq R$.

Consider:
$\mathcal{A} = \langle A_i | i \in I\rangle$ such that $\mathcal{A} : I \to \{A_i | i \in I\}$ defined by $\mathcal{A}(i) = A_i$.

4.4 Definition: Cartesian Product
Let $\langle A_i | i \in I\rangle$ be nonempty indexed sets. Define the *Cartesian Product* $\prod_{i\in I} A_i$, to be the set of all functions f with $\operatorname{dom}(f) = I$ and for each i, $f(i) \in A_i$.

4.5 Proposition:
The following are equivalent:
(a) Axiom of Choice
(b) For every set A, binary relation $R \subseteq A \times A$, there is a function which uniformizes R.
(c) For every system $\langle A_i | i \in I\rangle$ of nonempty sets, $\prod_{i\in I} A_i \neq \varnothing$. Proof: homework.

Let $(P, \leq)$ be a partial order. (poset).
Recall: the above is such that $\leq$ is a binary relation on P which is reflexive, transitive, and antisymmetric. We also define $x < y \Leftrightarrow x \leq y$ and $x \neq y$.

4.6 Definition: Maximal element of a poset
Let $(P, \leq)$ be a poset. $x \in P$ is a *maximal element* if $\forall y \in P$, $y \leq x$.
Remark: here, maximal does not necessarily mean largest.

Example 1: If $(P, \leq)$ is linear, then maximal element is the largest, since any two elements are comparable.

Example 2: Let V be a vector space. $P = \{L \subset V | L \text{ is linearly independent }\}$. Consider $(P, \subseteq)$. In $(P, \subseteq)$, $B \subset V$ is a basis iff B is a maximal element. There may be no largest because some elements in P are not comparable.

Definition— chain:

Let $(P, \leq)$ be a poset.

- $X \subseteq P$ is a *chain* if $(X, \leq)$ is a linear ordering.
- $p \in P$ is an upper bound for $X \subseteq P$ if $(\forall q \in X)(q \leq p)$. p is a *strict upper bound* if $(\forall q \in X)(q < p)$.

4.7 Theorem:

The following are equivalent:

(a) AC

(b) (Zorn's Lemma) if $(P, \leq)$ is a poset such that for each chain $X \subseteq P$, there is an upper bound, then for every $p \in P$, there is a maximal element q such that $p \leq q$.

(c) (Zermelo's Theorem) Every set can be well-ordered.

Proof. (a)$\Rightarrow$ (c) " classical proof "

Let $(P, \leq)$ be a poset such that every chain has an upper bound and let $p \in P$. Let $F : \mathcal{P}(P) \to P$ be a selector. I.e. $(\forall X \subseteq P)$ $(F(X) \in X)$.

Idea: construct $p \in b_0 \subset b_1 \subset ... \subset b_i \subset ...$ an " increasing sequence " of chains. Then take $\bigcup B = b^*$. Show b^* is a chain. Take an upper bound q for b^*. Show q is maximal element.

For every chain $b \subset P$, let A_b be the set of all strict upper bounds. Define B to be the set of chains $b \subset P$ such that

(i) $p = \min(b)$.

(ii) For every initial segment $\bar{b}$ of b, $\min(b - \bar{b})$ exists and is equal to $F(A_{\bar{b}})$.

Recall: $\bar{b}$ is an initial segment of b if $b \ni x \leq y \in \bar{b} \Rightarrow x \in \bar{b}$.

Claim 1: If b_1, $b_2 \in B$, then $b_1 \subset b_2$ or $b_2 \subset b_1$.

Proof: Let $b_1, b_2 \in B$. Let

$$\bar{b} = \bigcup \{c \subseteq b_1 \cap b_2 | c \text{ is an initial segment of } b_1 \text{ and } b_2\}$$

Note that $\bar{b}$ is an initial segment of b_1, b_2 by definition of initial segment. Suppose $\bar{b} \neq b_1$ and $x_1 = \bar{b} \neq b_2$. Then $\min(b_1 - \bar{b}) = F(A_{\bar{b}}) = \min(b_2 - \bar{b}) = x_2$. Contradiction, so $\bar{b} = b_1$ or $\bar{b} = b_2$. End proof of claim 1.

Claim 2: Let $b^* = \bigcup B$. Then $b^* \in B$ and $b^* = \bigcup B$ is a chain and every initial segment $b \subset b^*$ is in B.

Proof: Let $b \in B$. Show b is an initial segment of b^*. Suppose $q \leq r$ such that $q \in b^*$ and $r \in b$. Since $q \in b^*$, let $b_q \in B$ such that $q \in b_q$. By Claim 1, b_q is a proper initial segment of b or b is an initial segment of b_q.

Case 1: $b_q \subseteq b$ so $q \in b$.

Case 2: $b \subseteq b_q$. Then $b_q \ni q \leq r \in b$. So $b = b_q$ by definition. So, every element of B is an initial segment of b^*. Next, we show that $b^* \in B$. So we need to show that $\min(b^*) = p$ and $\bar{b}$ is a proper initial segment of b^*.

Let $q \in b^* - b$. Let $b_q \in B$ such that $q \in b_q$. $\bar{b}$ is an initial segment of b^* and b_q is an initial segment of b^*, then $\bar{b}$ is an initial segment of b_q. Because if $x \leq y \in \bar{b}$ $(b_q \subseteq b^*)$ so $x \in \bar{b}$. And $b_q \in B$. By (ii) for b_q we get $\min(b_q - \bar{b})$ $(=\min(b^* - \bar{b}) = F(A_{\bar{b}}))$.

Subclaim: $\min(b_q - \bar{b}) = \min(b_q - b)$ (let $= r$).

Proof: if $x \in \bar{b}$ then $x \leq r$. If $y \in b^*$ such that $\forall x \in \bar{b}$, $x \leq y$, then let $b_y \in B$ such that $y \in b_y$. b_y is an initial segment of b_q or vice versa. Let $y' \in b_y \cap (b_q - \bar{b})$ such that $y' \leq y$. That implies y' is an upper bound for $\bar{b}$ which implies $r \leq y' \leq y$. So $b^* \in B$.

Let q be an upper bound for $b^* \in B$. b^* is a chain and so it has an upper bound. This leads to:

Claim 3: Let q be an upper bound for b^*, then q is a maximal element.

Otherwise, let $q' \in P$ such that $q < q'$, $q' \in A_{b^*}$. $q < q' \Rightarrow q' \in A_{b^*} \neq \varnothing$ so $b^* \cup \{F(A_{b^*})\}$ is a new chain in B. But $b^* = \bigcup B$ so that is a contradiction.

So (a)⇒(b)

□

Now we do the modern proof:

Proof. Let $b \mapsto A_b$ as above (A_b is a set of strict upper bounds) and let $F : \mathcal{P}(P) \to P$ be a selector. Let $p \in P$ and let $s \notin P$. Define the class function $g : O_n \to P \cup \{s\}$ by

$$g(\alpha) = \begin{cases} p & \text{if} & \alpha = 0 \\ F(A_{g[\alpha]}) & \text{if} & A_{g[\alpha]} \neq \varnothing \\ s & \text{otherwise} & \end{cases}$$

Given $g \restriction \alpha$, if possible we take $g(\alpha) = F$ (chain so far).

Claim 1: If $\alpha < \beta$, $g(\alpha), g(\beta) \in P$ then $g(\alpha) < g(\beta)$.

Proof: $g(\beta) = F(A_{g[\beta]})$ and $A_{g[\beta]} \neq \varnothing$. $g(\beta) \in A_{g[\beta]}$ and $g(\alpha) \in g[\beta] \Rightarrow g(\alpha) < g(\beta)$ by the definition of $b \mapsto A_b$. End proof of claim.

Since P is a set, there is some α such that $g(\alpha) = s$. Otherwise, $g : O_n \to P \cup \{s\}$ is injective and a contradiction ensues. Take the least α such that $g(\alpha) = s$. Then by Claim 1, $g[\alpha]$ is a chain in P. Let q be an upper bound for it. Want to show that q is maximal. Note: $p \leq q$. Otherwise, there is $q' \in P$ such that $q < q'$ but then $q' \in A_{g[\alpha]}$ and $g(\alpha) = s \Rightarrow A_{g[\alpha]} = \varnothing$. Contradiction.

□

Now, we prove (c)⇒(a).

Proof. Let $A \neq \varnothing$ be a set and let $\prec$ be a well ordering of $\bigcup A$. Define F as follows for nonempty $a \in A$:

$$F(a) = \text{the } \prec\text{-minimal element of } a.$$

Then, F is a selector.

□

Week 8:

Recall: Theorem 4.7 stated the equivalence of the Axiom of Choice, Zorn's Lemma and the fact that Every set can be well-ordered. We proved that AC implies Zorn's Lemma and also that the well-ordering of every set implies AC.

We prove (b) $\Rightarrow$ (c) the classical way:

Proof. Given is a set A. We construct a well-ordering on A by looking at the set of all approximations to this well-ordering and applying Zorn's Lemma. We let

$$P = \text{the set of all ordered pairs } (a, r) \text{ where } a \subseteq A \text{ and } r \text{ is a well-ordering on } a.$$

Notice: $r \subseteq a \times a$. P is the set of all approximations to a well-ordering on A.

For (a, r) and $(b, s) \in P$ we let

$$(a, r) \leq (b, s) \text{ iff } (a \subseteq b \text{ and } s \text{ is an end-extension of } r).$$

Recall: (b, s) is an end-extension of (a, r) iff $r \subseteq s$ and if $(x, y) \in s$ and $y \in a$ then $x \in a$ (there are no holes). This says that all elements of $b - a$ are past all elements of a in the ordering s.

Claim 1: $(P, \leq)$ is a partial ordering (in the nonstrict sense).

Proof: exercise.

Goal: First using Zorn's Lemma, we show that there is a maximal element in $(P, \leq)$ and next, we show that any maximal element of $(P, \leq)$ is a well-ordering on A.

Claim 1: The poset $(P, \leq)$ satisfies Zorn's condition.

Proof: Want to show: if B is a chain in $(P, \leq)$, then B has an upper bound. We let

$$a^* = \bigcup\{a \in \mathcal{P}(A) | (\exists r)((a, r) \in B)\}$$
$$r^* = \bigcup\{r \in \mathcal{P}(A \times A) | (\exists a)((a, r) \in B)\}$$

So a^* is the union of all first components. And obviously $a^* \subseteq A$ and $r^* \subseteq a^* \times a^*$.

We show that r^* is a well-ordering on $a^* \times a^*$.

Linearity: Assume $x, y \in a^*$. This means that we can find some $(a, r), (a', r') \in B$ such that $x \in a$ and $y \in a'$. B is a chain so

$$\text{either } (a, r) \leq (a', r') \text{ or } (a', r') \leq (a, r).$$

Assume WLOG $(a, r) \leq (a', r')$. Hence, $a \subseteq a'$, so $x, y \in a'$.

Because r' is a linear ordering on a': $(x, y) \in r'$ or $(y, x) \in r'$. But $r' \subseteq r^*$ by construction, so $(x, y) \in r^*$ or $(y, x) \in r^*$. That proves linearity.

Irreflexivity + Transitivity: Exercise.

So far we have that r^* is a linear ordering on a^*. Now we show that it is a well-ordering:

Let $\varnothing \neq X \subseteq a^*$. We look for a least element of X with respect to r^*.

Consider $x \in X$. Since a^* is the union of all sets a where $(a, r) \in B$, we can find some $(a, r) \in B$ such that $x \in a$. The point is that a^* is an end-extension of a, or equivalently, a is an initial segment of a^*. Because r is a well-ordering on a, there is a least element of $X \cap a$ with respect to r; call it x^*. We show that $(x^*, y) \in r^*$ whenever $y \in a^*$, $y \neq x$.

If $y \in a$: Then $(x^*, y) \in r$ by the definition of x^*, so $r \subseteq r^*$.

If $y \notin a$: Then there is some $(a', r') \in B$ such that $y \in a'$. So actually $y \in a' - a$. But B is a chain, so either $(a, r) \leq (a', r')$ or $(a', r') \leq (a, r)$. Since $y \in a' - a$, the latter cannot hold. Otherwise, we would have $a' \subseteq a$. So necessarily $(a, r) \leq (a', r')$. This means that r' is an end-extension of r. Since $y \in a' - a$: $(x^*, y) \in r'$. So $(x^*, y) \in r^*$, as $r' \subseteq r^*$.

This shows that x^* is the least element of a^* with respect to r^*. This completes the proof that r^* is a well-ordering on a^*, so $(a^*, r^*) \in P$.

Moreover, by construction: we have that $a \subseteq a^*$ and $r \subseteq r^*$ whenever $(a, r) \in B$. So in order to show that $(a, r) \leq (a^*, r^*)$, we have to prove:

$$r^* \text{ is an end-extension of } r.$$

And to see this, it remains to prove: if $x \in a$ and $y \in a^* - a$ then $(x, y) \in r^*$.

By construction, there is some $(a', r') \in B$ such that $y \in a'$. As in the case "$y \notin a$" above, show that $(a, r) \leq (a', r')$, so $(x, y) \in r' \subseteq r^*$.

Summary: $(a, r) \leq (a^*, r^*)$ for all $(a, r) \in B$. I.e. (a^*, r^*) is an upper bound on B. That ends the proof of Claim 1.

By Zorn's Lemma, $(P, \leq)$ has a maximal element, call it $(\hat{a}, \hat{r})$.

Claim 2: $\hat{a} = A$ (Hence $\hat{r}$ is a well-ordering on A).

Proof: If not, $A - \hat{a} \neq \varnothing$. So we pick some $x \in A - \hat{a}$. Let

$$\tilde{a} = \hat{a} \cup \{x\} \text{ so } \tilde{a} \supsetneq \hat{a}.$$
$$\tilde{r} = \hat{r} \cup (\hat{a} \times \{x\}).$$

Then $(\tilde{a}, \tilde{r}) \in P$ and $(\hat{a}, \hat{r}) < (\tilde{a}, \tilde{r})$.

It is sufficient to show that $(\hat{a}, \hat{r}) \leq (\tilde{a}, \tilde{r})$ because $\hat{a} < \tilde{a}$. Proof of these facts: Exercise.

But this contradicts the fact that $(\hat{a}, \hat{r})$ is maximal in $(P, \leq)$. So necessarily $\hat{a} = A$.

This ends the proof of Claim 2, (b)$\Rightarrow$ (c) and of Theorem 4.7. □

Recall that we proved the equivalence of the following three:
(a) AC
(b) Zorn's Lemma
(c) Zermelo's Theorem (Well-ordering principle).

4.8 Theorem:

The following are equivalent for every set A:
(a) There is a selector on $\mathcal{P}(A)$.
(b) A can be well-ordered.

Proof. (b)$\Rightarrow$ (a) same as (c)$\Rightarrow$ (a) in T.4.7. If $<$ is a well-order on A then the function $f : \mathcal{P}(A) \rightarrow A$ defined by $f(x) =$ the $< -$least element of x for $x \neq \varnothing$ is a selector on $\mathcal{P}(A)$.

(a)$\Rightarrow$ (b) homework. Hint: construction by recursion. Keep picking elements of A and put them one after another. That is guided by the selector. Similar to proof that ordinals are isomorphic to some well-ordering.

□

SOME APPLICATIONS OF AC:

4.9 Example:

AC implies that every vector space has a basis.

Proof. We use Zorn's Lemma. Our posets P will consist of all approximations to a basis:

Fix a vector space W over some field F.

$P =$ the set of all linearly independent $X \subseteq W$.

Ordering: $\subseteq$.

Obviously, $(P, \subseteq)$ is a partial ordering. We check that $(P, \subseteq)$ satisfies Zorn's condition. Let B be a chain in $(P, \subseteq)$. Then $\bigcup B$ is an upper bound on B. So clearly $X \subseteq \bigcup B$ whenever $X \in B$, so it suffices to verify that $\bigcup B$ is linearly independent.

Let $v_1, ..., v_n \in \bigcup B$. Then to each n we have some $X_n \in B$ such that $v_n \in X_n$. Since B is a chain for each $i, j \in \{1, ..., n\}$ we have $X_i \subseteq X_j$ or $X_j \subseteq X_i$. Let k be such that X_k is the largest among them, i.e. $X_i \subseteq X_k$ for all $i \in \{1, ..., n\}$. Then $v_1, ..., v_n \in X_k$. Since X_k is linearly independent set of vectors: $v_1, ..., v_n$ are linearly independent. Conclusion: $\bigcup B$ is a linearly independent set of vectors, i.e. $\bigcup B \in P$. It follows that $\bigcup B$ is an upper bound on B in $(P, \subseteq)$.

By Zorn's Lemma: $(P, \subseteq)$ has a maximan element, call it Z. So Z is a linearly independent set of vectors. We show that Z generates the entire W. So pick any $v \in W$. By maximality of Z, we have that $Z \cup \{v\}$ is linearly dependent, so we have finite set of vectors $\{z_1, ..., z_n\} \subseteq Z$ and elements $\alpha_1, ..., \alpha_n \in F$ such that at least one of the alphas is nonzero and $\alpha_1 z_1 = ... + \alpha_n z_n + \alpha v = 0_W$. But $\alpha \neq 0$ otherwise we would have $\alpha_1 z_1 = ... + \alpha_n z_n = 0_W$, which is impossible since $z_1, ..., z_n$ are linearly independent.

It follows that $v = -(\frac{\alpha_1}{\alpha} z_1 + ... + \frac{\alpha_n}{\alpha} z_n)$ □

4.10 Example:

AC implies that every partial ordering can be linearized. This means if (A, R) is such that A is a set and $R \subseteq A \times A$ is a partial ordering on A, then it is possible to find a linear ordering R' on A such that for all $a, b \in A$:

$$\langle a, b \rangle \in R \Rightarrow \langle a, b \rangle \in R'$$

This can be briefly expressed as: $R \subseteq R'$.

Proof. Again we apply Zorn's Lemma to the poset of all approximations to R'.

$P =$ the set of all partial orderings Q on A that extend R, i.e. $R \subseteq Q$.

Ordering: $\subseteq$.

Trivial: $(P, \subseteq)$ is a poset.

Easy to check: $(P, \subseteq)$ satisfies Zorn's condition.

If B is a chain in $(P, \subseteq)$ then $\bigcup B \in P$. By Zorn's Lemma, $(P, \subseteq)$ has a maximal element R'. So R' is a poset. We verify that R' is linear.

What if not: then we have $x, y \in A$ that are not comparable under R'. We show that there is an $R^* \in P$ that properly extends R'. This will contradict the maximality of R'. We let

$$R^* = R' \cup \{\langle a,b\rangle \in A\times A | \langle a,x\rangle \in R' \wedge \langle y,b\rangle \in R'\}$$

We check that R^* is a partial ordering: (picture)

Reflexivity: Trivial, as $R^* \supseteq R' \supseteq R$ and R is reflexive.

Antisymmetry: Need to see: $\langle a,b\rangle \in R^*, \langle b,a\rangle \in R^*$ then $a = b$. Obviously this is true if $\langle a,b\rangle \in R'$, so we focus on the case where $\langle a,x\rangle \in R'$ and $\langle y,b\rangle \in R'$. In this case we show that $\langle b,a\rangle \notin R'$. Hence $\langle b,a\rangle \notin R^*$. If $\langle b,a\rangle \in R'$ we would have:

$\langle y,b\rangle \in R'$ and $\langle a,b\rangle \in R'$ and $\langle a,x\rangle \in R'$ so by transitivity, $\langle y,x\rangle \in R'$. This is a contradiction as x, y are not comparable under R'.

Transitivity: assume $\langle a,b\rangle \in R^*$ and $\langle b,c\rangle \in R^*$. Want to show: $\langle a,c\rangle \in R^*$. Again it suffices to focus on the case where $\langle a,b\rangle \notin R'$ or $\langle b,c\rangle \notin R'$. For instance, assume $\langle a,b\rangle \notin R'$. Then $\langle a,x\rangle \in R'$ and $\langle y,b\rangle \in R'$. Since $\langle b,c\rangle \in R^*$, we must have $\langle b,c\rangle \in R'$, by the argument for antisymmetry above. So we have:

$$\langle a,x\rangle \in R' \text{ and } \langle y,b\rangle \in R' \text{ and } \langle b,c\rangle \in R' \text{ so } \langle y,c\rangle \in R'$$

so $\langle a,c\rangle \in R^*$. The case where $\langle b,c\rangle \notin R'$ is done similarly.

□

4.11 Example: Non-measurable sets

A measure on a set A is a function $\mu : \mathcal{P}(A) \to [0,\infty] \subseteq \mathbb{R}$ such that whenever $\{X_i | i \in \omega\}$ is a family of mutually disjoint subsets of A, we have:

$$\mu(\textstyle\bigcup_{i\in\omega} X_i) = \sum_{i\in\omega} \mu(X_i)$$

The following can be derived:

— $\mu(\varnothing) = 0$

— $X \subseteq Y \Rightarrow \mu(X) \leq \mu(Y)$

— $X_i \subseteq X_{i+1} \Rightarrow \mu(\bigcup_{i\in\omega} X_i) = \sup_{i\in\omega} \mu(X_i)$

On $\mathbb{R}^n$ we have *Lebesgue measure* that has the following additional properties:

— $\mu_L([0,1]) = 1$

— $\mu_L(\{r\}) = 0$ (we say that μ_L is non-atomic).

— For $r \in \mathbb{R}$ and any $X \subseteq \mathbb{R}$ let

$$r + X = \{r + x | x \in X\}$$

The operation $X \mapsto r + X$ is called translation by r. Then μ_L is *translation invariant* i.e.

$$\mu_L(r + X) = \mu_L(X) \text{ for all } X \in \text{dom}(\mu_L)$$

The Axiom of Choice implies that $\text{dom}(\mu_L) \subsetneq \mathcal{P}(\mathbb{R})$, i.e. μ_L cannot be defined for all $X \in \mathcal{P}(\mathbb{R})$. In other words, there is a μ_L- non-measurable set.

In fact: We show there is no translation-invariant non-trivial measure that measures all subsets of $\mathbb{R}$. (Non-trivial means that $0 < \mu((a,b)) < \infty$ if $a < b$).

Let μ be a translation-invariant measure on $\mathbb{R}$. We assume that $\mathbb{R}$ is well-orderable and construct a non-measurable set X.

Define a binary relation $\sim$ on $\mathbb{R}$ by

$$x \sim y \text{ iff } x - y \in \mathbb{Q}$$

Easy to check that $\sim$ is an equivalence relation on $\mathbb{R}$ and that for every $x \in \mathbb{R} : [x] = x + \mathbb{Q}$.

Using the fact that $\mathbb{R}$ is well-orderable we can find a set $X \subseteq \mathbb{R}$ such that $X \cap [x]$ has exactly one element for each $x \in \mathbb{R}$, i.e. X picks exactly one element from each equivalence class. Moreover, we can construct X in such a way that $X \subseteq (0, 1)$. This is because each equivalence class is dense in $\mathbb{R}$ by the above definition of $[x]$.

Claim: X is not μ-measurable.
If X were μ-measurable then $\mu(X) < \infty$ because $X \subseteq (0, 1)$.

Case 1: If $\mu(X) = 0$.
Notice: $\mathbb{R} = \bigcup_{q \in \mathbb{Q}} q + X$ (if $y \in \mathbb{R}$ then $y \in [x]$ for some $x \in X$. But then if we let $q = y - x$ then q is a rational number and $y = q + x$).
But if $q \neq q'$ then $(q + X) \cap (q' + X) = \varnothing$. Why: if $q + x = q' + x'$ then $x - x' = q - q' \in \mathbb{Q}$. But $x, x' \in X$ so if $x \neq x'$ then x, x' come from different equivalence classes, so $x \not\sim x'$ i.e. $x - x' \notin \mathbb{Q}$.
So $\mu(\mathbb{R}) = \mu(\bigcup_{q \in \mathbb{Q}} q + X) = \sum_{q \in \mathbb{Q}} \mu(q + X) = \sum_{q \in \mathbb{Q}} \mu(X) = \sum_{q \in \mathbb{Q}} 0$. Contradiction.

Case 2: $\mu(X) > 0$. Say $\mu(X) = \delta > 0$.
Then $\bigcup_{q \in \mathbb{Q} \cap (0,1)} q + X \subseteq (0, 2)$ since $X \subseteq (0, 1)$. So as in Case 1:
$\infty > \mu((0, 2)) \geq \mu(\bigcup_{q \in \mathbb{Q} \cap (0,1)} q + X) = \sum_{q \in \mathbb{Q} \cap (0,1)} \mu(q + X) = \sum_{q \in \mathbb{Q} \cap (0,1)} \mu(X) =$
$\sum_{q \in \mathbb{Q} \cap (0,1)} \delta = \infty$.

4.12 Example: Filter
Let A be a set. A *filter* on A is a family $\mathcal{F} \subseteq \mathcal{P}(A)$ such that:
— $A \in \mathcal{F}$ and $\varnothing \notin \mathcal{F}$
— $\mathcal{F}$ is upwards closed, i.e. if $A \in \mathcal{F}$ and $A \subseteq B$ then $B \in \mathcal{F}$.
— If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.
Intuitively, a filter determines which subsets of A are "large".

For instance:
(a) Consider $A = (0, 1)$ and

$$\mathcal{F} = \{X \subseteq A | \mu_L(X) = 1\}$$

Then $\mathcal{F}$ is a filter on A.
(b) Consider given arbitrary infinite set A, let

$$\mathcal{F} = \{X \in \mathcal{P}(A) | A - X \text{ is finite}\}$$

This is called the *Frechér* filter on A. In analysis we use this to determine convergence of sequences.

A filter U on a set A is an *ultrafilter* iff for every $X \in \mathcal{P}(A)$:

$$X \in U \text{ or } A - X \in U$$

Notice: if $\mathcal{F}$ is a filter then we cannot have both X and $A - X$ inside $\mathcal{F}$.

Prime Ideal Theorem (PID): If A is a set and $\mathcal{F}$ is a filter on A then $\mathcal{F}$ can be extended to an ultrafilter. That is, there is an ultrafilter U on A such that $\mathcal{F} \subseteq U$.

We defined a filter $\mathcal{F} \subseteq \mathcal{P}(A)$ on a set A as a family such that:
—$A \in \mathcal{F}$, $\varnothing \notin \mathcal{F}$
—$\mathcal{F}$ is closed upwards.
—$\mathcal{F}$ is closed under intersections.

A filter U on a set A is an *ultrafilter* iff $X \in U$ or $A - X \in U$ for all $X \in \mathcal{P}(A)$.
Notice: If $\mathcal{F}$ is a filter then $\mathcal{F}$ is closed under finite intersections:

$$X_1, ..., X_n \in \mathcal{F} \Rightarrow X_1 \cap X_2 \cap ... \cap X_n \in \mathcal{F}.$$

The proof is trivial.

If $X \in \mathcal{P}(A)$ then $\mathcal{F}_X = \{Z \in \mathcal{P}(A) | X \subseteq Z\}$ is a filter on a set A. (Easy)
Filters of the form $\mathcal{F}_X$ are called *principal* filters.

If U is an ultrafilter on a set A and $a \in A$ then for every $X \in \mathcal{P}(A)$ we have

$$a \in X \text{ or } a \in A - X$$

This tells us that

$$U_a = \{Z \in \mathcal{P}(A) | a \in Z\}$$

is an ultrafilter on A. Following the terminology from above: $U_a = \mathcal{F}_{\{a\}}$ is a *principal* ultrafilter on A.
If we drop AC it may happen that ultrafilters on A are principal.
Notice also: If A is finite, then all ultrafilters on A are principal. (Easy).

Let A be a set. A family $\mathcal{B} \subseteq \mathcal{P}(A)$ is called a *filter base* iff for every finite sequence of sets $Z_1, ..., Z_n \in \mathcal{B} :\ Z_1 \cap ... \cap Z_n \neq \varnothing$.
So given a filter base $\mathcal{B}$, we let

$$\mathcal{F}(\mathcal{B}) = \{X \in \mathcal{P}(A) | \text{ there is a finite sequence } Z_1, ..., Z_n \in \mathcal{B} \text{ such that } Z_1 \cap ... \cap Z_n \subseteq X\}$$

In other words: $\mathcal{F}(\mathcal{B})$ is obtained by forming all intersections of finite subfamilies of $\mathcal{B}$ and then closing upwards.

Easy to verify: $\mathcal{F}(\mathcal{B})$ is a filter. This filter is called the *filter generated by* $\mathcal{B}$. Exercise: $\mathcal{F}(\mathcal{B})$ is the smallest filter $\mathcal{F}$ such that $\mathcal{B} \subseteq \mathcal{F}$, i.e.

$$\mathcal{F}(\mathcal{B}) = \bigcap\{\mathcal{F} \in \mathcal{P}(\mathcal{P}(A)) | \mathcal{F} \text{ is a filter and } \mathcal{B} \subseteq \mathcal{F}\}.$$

$(*)$ Let $\mathcal{F}$ be a filter on A and $X \subseteq A$. Then either $\mathcal{F} \cup \{X\}$ or $\mathcal{F} \cup \{A - X\}$ is a filter base.

Proof. It is enough to show that one of the following (1) is true:

Either $Z \cap X \neq \varnothing$ for all $Z \in \mathcal{F}$

Or else $Z \cap (A - X) \neq \varnothing$ for all $Z \in \mathcal{F}$.

Why: If $Z \cap X$ is nonempty for all $Z \in \mathcal{F}$ then $\mathcal{F} \cup \{X\}$ is a filter base. If $Z_1, ..., Z_n \in \mathcal{F} \cup \{Z\}$ then either $Z_i \in \mathcal{F}$ for all $i < n+1$ and then obviously $Z_1 \cap ... \cap Z_n \neq \varnothing$, or else $Z_i = Z$ for some $i < n+1$ and then $Z_1 \cap ... \cap Z_n = Z \cap (\bigcap_{j \neq i} Z_j) \neq \varnothing$ where $\mathcal{F} = \bigcap_{j \neq i} Z_j$.

Similarly for $A - X$.

Now prove $(*)$: assume there is some $X \in \mathcal{P}(A)$ such that (1) fails. Then we can find $Z_1, Z_2 \in \mathcal{F}$ such that

$$Z_1 \cap X = \varnothing \text{ and } Z_2 \cap (A - X) = \varnothing$$

Let $Z = Z_1 \cap Z_2 \in \mathcal{F}$. Then still $Z \cap X = \varnothing$ and $Z \cap (A - X) = \varnothing$. Then:

$$\varnothing = (Z \cap X) \cup (Z \cap (A - X)) = Z \cap (X \cup (A - X)) = Z \cap A = Z. \text{ Contradiction.}$$

□

Prime Ideal Theorem (PID): If A is a set and $\mathcal{F}$ is a filter on A, then there is an ultrafilter U on A such that $\mathcal{F} \subseteq U$.

We show: AC$\Rightarrow$ PID

Proof. We use Zorn's Lemma. We let

$$P = \text{the set of all filters on } A. \text{ The ordering: } \subseteq.$$

This poset $(P, \subseteq)$ easily satisfies Zorn's condition: If B is a chain of filters on A then $\bigcup B$ is a filter on A. (Exercise).

So there is a maximal filter U in $(P, \subseteq)$. We show that U is an ultrafilter. That is, if $X \in \mathcal{P}(A)$ then $X \in U$ or $(A - X) \in U$. But if there is some X with both $X \notin U$ and $(A - X) \notin U$ then by $(*)$

$$U \cup \{X\} \text{ or } U \cup \{A - X\}$$

is a filter base so

$$\mathcal{F}(U \cup \{X\}) \text{ or } \mathcal{F}(U \cup \{A - X\})$$

is a filter properly extending U, which contradicts the maximality.

That ends the proof □

Remark PID $\not\Rightarrow$ AC.

Corollary: A filter U on A is maximal in $(P, \subseteq)$ iff U is an ultrafilter on A.

Conclusion (AC): If A is an infinite set then there exists a nonprincipal ultrafilter on A. Why: Let $\mathcal{F}$ be the Frechér filter on A, i.e. $\mathcal{F} = \{X \in \mathcal{P}(A) | A - X \text{ is finite}\}$. Then $A - \{a\} \in \mathcal{F}$ for all $a \in A$. Let $U \supseteq \mathcal{F}$ be an ultrafilter on A.

Then $A - \{a\} \in U$ for all $a \in A$, so U is non-principal.

5. Cardinals (ZF)

5.1 Definition: Equinumerous

Two sets A, B are *equinumerous* iff there is a bijection $f : A \to B$. We write $A \sim B$.

5.2 Proposition:

$\sim$ is an equivalence relation. Moreover, $\sim$ is a class and for each set $a \in V$ the equivalence class $[a]_\sim$ is a proper class. $(a \sim a \times \{x\} \forall x \in V)$.

Our goal: Find a representative for each equivalence class.

Notice: we could do this locally, using AC. If we have some big set S we can use AC to pick representatives on $S/\sim$. However, we would like to do everything globally.

5.3 Definition: Cardinal Assignment

A *cardinal assignment* is a class function $c : V \to V$ such that

$$c(x) = c(y) \text{ iff } x \sim y$$

for all $x, y \in V$.

In other words, a cardinal assignment is a class selector on $V/\sim$.

5.4 Proposition (ZF):

There is such a cardinal assignment.

Proof. First define a function $x \mapsto \alpha_x \in O_n$ by

$$\alpha_x = \text{the least ordinal } \alpha \text{ such that } [x]_\sim \cap V_\alpha \neq \varnothing$$

It is easy to check that $x \mapsto \alpha_x$ is a class function defined on the entire V. Then let

$$c(x) = [x]_\sim \cap V_{\alpha_x}$$

Again easy to check: $c : V \to V$ is a class function and $c(x) \neq \varnothing$ for all $x \in V$. Notice $c(x) \subseteq [x]_\sim$ so every element of $c(x)$ is equinumerous to x. This tells us: $c(x) = c(y) \Rightarrow x \sim y$ since every $z \in c(x) = c(y)$ is equinumerous to both x and y.

On the other hand, if $x \sim y$ then $[x]_\sim = [y]_\sim \Rightarrow \alpha_x = \alpha_y$, hence

$$c(x) = V_{\alpha_x} \cap [x]_\sim = V_{\alpha_y} \cap [y]_\sim = c(y).$$

□

5.5 Remark:

P.5.4 Uses the Axiom of Foundation, as it assumes that $V = \bigcup_{\alpha \in O_n} V_\alpha$. The method in P.5.4 is important, as it gives a recipe how to replace proper classes by sets in a uniform way (i.e. we can express this process of replacement in the language of set theory).

The drawback: Typically, $c(x)$ is much larger than x in the sense that there will be no surjection $g : x \longrightarrow c(x)$ in most cases.

However, one can show that without AC we cannot do better.

Our goal is to find a cardinal assignment $c : V \rightarrow V$ such that $c(x) \sim x$. This assignment will be much more useful. However, we will use AC here.

5.6 Definition:

Given sets A, B, we write

(i) $A \leq B$ iff there is an injection $f : A \rightarrow B$

(ii) $A \leq^s B$ iff there is a surjection $g : B \rightarrow A$

Notice: (ZF) $A \leq B \Rightarrow A \leq^s B$

(ZFC) $A \leq^s B \Rightarrow A \leq B$

So these are equivalent under (ZFC).

5.7 Theorem (ZF): (Schröder-Bernstein) ($A \leq B$ and $B \leq A$) $\Rightarrow A \sim B$

Proof. $f : A \longrightarrow B$, $g : B \longrightarrow A$ are injections.

Let $A_1 = g[B]$. Then $A_1 \sim B$.

Let $A_2 = g \circ f[A]$ then $A_2 \sim A$ and $A_2 \subseteq A_1 \subseteq A$. It suffices to show: $A_1 \sim A$.

We prove: assume $A' \subseteq A$ and $h : A \longrightarrow A'$ is an injection. Then $A \sim A'$. Comparing with the above: A' plays the role of A_1 and h plays the role of $g \circ f$. Also:

$$\begin{gathered} A'' = h[A] \\ A''' = h[A'] \\ \vdots \end{gathered}$$

Idea: use h to map $A - A'$ onto $A'' - A'''$, $A'' - A'''$ onto $A'''' - A'''''$ etc. And don't move the remaining elements if there are any.

An economic way of writing the things:

For $n \in \omega$ let $h^n(x) = h \circ h \circ ... \circ h(x)$ n times, $h^0(x) = x$. Let

$$X = \text{the set of all } x \in A \text{ such that } (\exists \bar{x} \in A - A')(\exists n)(x = h^n(\bar{x})$$

Now define $H : A \longrightarrow A'$ by

$$H(a) = \begin{cases} h(a) & \text{if } \ a \in X \\ a & \text{if } \ a \notin X \end{cases}$$

Then $H : A \longrightarrow A'$ is a bijection. First notice that the values of $H(a)$ are indeed elements of A'. This is because either $H(a) = h(a)$ in which case h is already mapping into A'. Or else, $H(a) = a$ but this is only if $a \in A - X$ and $A - X \subseteq A - (A - A') = A'$ since $X \supseteq A'$.

H is injective: Assume $H(a) = H(b)$

Proof. Case 1: $a, b \in X$. This means: we have $\bar{a}, \bar{b} \in A - A'$ and $m, n \in \omega$ such that $a = h^m(\bar{a})$ and $b = h^n(\bar{b})$. We know that $H(a) = h(a)$ and $H(b) = h(b)$, so

$$h^{m+1}(\bar{a}) = h^{n+1}(\bar{b})$$

What if $m \neq n$. Assume WLOG that $m < n$. Using the injectivity of h, we can cancel $m+1$ times and we get that

$$A - A' \ni \bar{a} = h^{n-m}(\bar{b}) \in A'$$

Because $n - m > 0$. Contradiction. So $m = n$ and therefore $a = h^m(\bar{a}) = h^m(\bar{b}) = b$ by injectivity of h.

Case 2: WLOG $a \in X$ and $b \notin X$. In this case, we cannot have $H(a) = H(b)$ because by definition of H,

$$a \in X \Rightarrow H(a) = h(a) \in X$$
$$a \in A - X \Rightarrow H(a) = a \in X.$$

Case 3: $a, b \in A - X$. But in this case, $b = H(b) = H(a) = a$. □

H is surjective:

Proof. Let $a \in A'$. Then either $a \in X$ in which case $a = h^n(\bar{a})$ for some $\bar{a} \in A - A'$ and $n \in \omega$. But $n > 0$ otherwise $A \ni a = \bar{a} \in A - A'$. So $n = \bar{n} + 1$ for some $\bar{n} \in \omega$ and

$$a = h^{\bar{n}+1}(\bar{a}) = h(h^{\bar{n}}(\bar{a})) = H(h^{\bar{n}}(\bar{a}))\ (h^{\bar{n}}(\bar{a}) \in X)$$

or else: $a \in A - X$ in which case $a \in A'$ and $H(a) = a$. In either case, $a \in \mathrm{rng}(H)$.
That ends the proof of surjectivity. □

That ends the proof. □

5.8 Corollary:
The relation $\leq$ is a partial ordering modulo $\sim$. That is:
$A \leq A$ for all A
$(A \leq B \wedge B \leq A) \Rightarrow A \sim B$ (Schröder-Bernstein)
$(A \leq B \wedge B \leq C) \Rightarrow A \leq C$

In particular: if c is the cardinal assignment from P.5.4: the second clause above can be replaced by

$$(A \leq B \wedge B \leq C) \Rightarrow c(A) = c(B)$$

So $\leq$ induces a partial ordering on $\mathrm{rng}(c)$.
We let $c(A) \preceq c(B)$ iff $A \leq B$. Then $\preceq$ is a partial ordering.

Notice: if $\alpha < \beta$ are ordinals such that $\alpha \sim \beta$ and $\alpha < \gamma < \beta$ then $\alpha \sim \gamma \sim \beta$. Why:
$id_\gamma : \gamma \to \beta$ is an injection
$f : \beta \to \gamma$ is an injection when $f : \beta \to \alpha$ is a bijection.
So by Schröder-Bernstein: $\gamma \sim \beta$.

This means: if $\alpha \in O_n$ then $I_\alpha = \{\xi \in O_n | \xi \sim \alpha\}$ is an interval. I_α is open from above whenever α is *infinite* ($\alpha \geq \omega$), i.e. I_α does not have a largest element: if $\xi \in I_\alpha$ then also $\xi + 1 \in I_\alpha$ because if $f : \alpha \to \xi$ is a bijection then $f' : \alpha \to \xi + 1$ is a bijection when

$f'(0) = \xi$
$f'(n+1) = f(n)$ for $n < \omega$
$f'(n) = f(n)$ for $n \geq \omega$.

Also: each I_α has a least element, because $I_\alpha \subseteq O_n$. So I_α looks as follows: $[\kappa_\alpha, \lambda_\alpha)$ when possibly $\lambda_\alpha = \infty$.

Important point: The least element κ_α of I_α is not equinumerous to any $\beta < \kappa_\alpha$.

5.9 Definition: Cardinals
An ordinal α is a *cardinal* iff α is not equinumerous to any $\beta < \alpha$.

5.10 Definition:
We say that a set A has a cardinality iff A is equinumerous to some cardinal. In this case we denote the cardinality of A by card(A) or $|A|$. Some literature also uses $\bar{\bar{A}}$.

5.11 Proposition:
AC implies that every set has a cardinality. So under AC, we have a cardinality assignment $A \mapsto$card(A) such that $A \sim$card(A) for all sets A.

Proof. Let A be a set. By AC, we can find a well-ordering $<$ on A. We proved that every well-ordering is isomorphic to an ordinal, so we can find $\alpha \in O_n$ such that $(A, \leq)$ is isomorphic to $(\alpha, \in)$ via some isomorphism f. Let $\kappa \in I_\alpha$ be a cardinal. Then we have a bijection $g : \alpha \to \kappa$. Then $g \circ f : A \to \kappa$ so $A \sim \kappa$. □

5.12 Remark:
Notice that under AC: all cardinalities are linearly ordered. As a consequence if A, B are sets then $A \leq B$ or $B \leq A$. This is not true without AC. Given a partial ordering $(I, \leq_I)$ one can construct a model of ZF that contains $(I, \leq_I)$ and a system of sets $(A_i | i \in I)$ such that for each $i, j \in I$:

$$i \leq_I j \text{ iff } A_i \leq A_j$$

(one can also prove the converse of this).

5.13 Proposition:
(a) Every $n \in \omega$ is a cardinal.
(b) ω is a cardinal.

Proof. (a) By induction on n show that there is no injection of n into $m < n$. (Exercise).
(b) Follows from (a): If $f : \omega \to n$ is an injection, then $f : n+1 \to n$ is an injection □

5.14 Theorem: Existence of Cardinals
Let $C_n = \{\kappa \in O_n | \kappa \text{ is a cardinal}\}$. Then C_n is a proper class.

Proof. Easy: C_n is a class.

We show that C_n is proper. Suppose not. This means that $C_n \subseteq \alpha$ for some $\alpha \in O_n$. This means that I_α is of the form $[\kappa, \infty)$, i.e. every ordinal $\beta > \alpha$ is an element of I_α. In other words: I_α is a proper class. From this we derive a contradiction. Define a map

$$F : \mathcal{P}(\alpha \times \alpha) \to I_\alpha$$

by

$$F(a) = \begin{cases} \text{otp}(a) & \text{if} \quad a \text{ is a well-ordering on } \alpha \\ \alpha & \text{otherwise} \end{cases}$$

Then F is a surjection. Why: if $\gamma \in I_\alpha$ then we have a bijection $g : \alpha \to \gamma$. Define a set $a \subseteq \alpha \times \alpha$ by $\langle \xi, \xi' \rangle \in a \Leftrightarrow g(\xi) < g(\xi')$.

A routine verification yields that (α, a) is a well-ordering isomorphic to $(\gamma, \in)$ via g. (exercise).

This means that $\text{otp}(a) = \gamma$, so $F(a) = \gamma$. Hence, $\gamma \in \text{rng}(F)$. This shows $\text{rng}(F) = I_\alpha$. Now $\mathcal{P}(\alpha \times \alpha)$ is a set by the Power Set axiom. Also, α is a class function. So by Replacement, $\text{rng}(F)$ must be a set. But $\text{rng}(F) = I_\alpha$. Contradiction. □

5.15 Remark:

This proof substantially depends on the use of Power set and Replacement axioms. One can show that if we drop even one of them, we are unable to prove that there are cardinals other than those in $\omega \cup \{\omega\}$.

5.16 Definition:

A set A is
(i) *finite* iff $A \sim n$ for some $n \in \omega$.
(ii) *countable* iff $A \sim \omega$.
(iii) *uncountable* otherwise.

5.17 Proposition:

C_n is a closed class.

Proof. Assume $A \subseteq C_n$ and $\kappa \in O_n$ is a limit point of A. This means that $\kappa = \bigcup(A \cap \kappa)$. Then κ is a cardinal: if not, there would be some $\alpha < \kappa$ and some bijection $f : \kappa \to \alpha$. Since κ is a limit point of A, we can pick $\lambda \in A$ such that $\alpha < \lambda < \kappa$. Then $f \restriction \lambda : \lambda \to \alpha$ is an injection and $\alpha < \lambda$, λ a cardinal, this is a contradiction. (A cardinal may not be equinumerous to any other cardinal). □

5.18 Definition:

Since $C_n - \omega$, the class of all infinite cardinals is a proper class, we can use recursion on ordinals to define an enumeration of $C_n - \omega$. This enumeration is denoted by $\aleph$ (Hebrew "Aleph"). So

$$\aleph : \; O_n \to C_n - \omega$$

and is defined by recursion, i.e. $\aleph(\alpha) = \min(C_n - \omega - \aleph[\alpha])$. In human language: $\aleph(\alpha)$ is the α^{th} infinite cardinal. We write $\aleph_\alpha$ instead of $\aleph(\alpha)$. Also, we often write ω_α instead of $\aleph_\alpha$ if we want to stress our interest in the well-ordering of $\aleph_\alpha$.

By P.5.17 and P+D.3.33 and HW6: $\aleph$ is a normal function.

5.19 Definition: Operations on Cardinals

We define:

(a) *Cardinal Addition:*

$$\kappa + \lambda = |(\{0\} \times \kappa) \cup (\{1\} \times \lambda)|$$

(b) *Cardinal Multiplication:*

$$\kappa \cdot \lambda = |\kappa \times \lambda|.$$

Remark: the same notation was is used to ordinal operations, it will be always clear from the context was is meant.

5.20 Proposition:

Let κ be an infinite cardinal. Then $\kappa \cdot \kappa = \kappa$.

Proof. By induction on κ.

— If $\kappa = \omega$: Look at the otp$(\omega \times \omega, <_{mlex})$. Notice: if $(m, n) \in \omega \times \omega$ then the initial segment below (m, n) under $<_{mlex}$ is contained in the set $\max(m+1, n+1) \times \max(m+1, n+1)$. That is, the initial segment is *finite*. That is, true of any such initial segment. So: $(\omega \times \omega, <_{mlex})$ is an infinite well-ordering all of whose proper initial segments are finite. So otp$(\omega \times \omega, <_{mlex}) = \omega$.

Hence $|\omega \times \omega| = \omega$.

—If $\kappa > \omega$ look at $(\kappa \times \kappa, <_{mlex})$.

Similarly as above: If $(\alpha, \beta) \in \kappa \times \kappa$, the initial segment below (α, β) is contained in $\gamma \times \gamma$ where $\gamma = \max\{\alpha, \beta\} + 1$.

Now $\gamma < \kappa$. Hence if $\lambda = |\gamma|$ then $\lambda < \kappa$. If γ is infinite then so is λ and $\gamma \times \gamma \sim \lambda \times \lambda \sim \lambda$ where the last $\sim$ is from the induction hypothesis.

(Picture)

Conclusion: $(\kappa \times \kappa, <_{mlex})$ is a well-ordering all of whose initial segments are of cardinality smaller than κ.

But if $\alpha \in O_n$ and $\alpha > \kappa$ then κ is a proper initial segment of α. And $|\kappa| = \kappa$. It follows that otp$(\kappa \times \kappa, <_{mlex}) \leq \kappa$. On the other hand: the map $\xi \mapsto \langle o, \xi \rangle$ is an order-reserving map from $(\kappa, <)$ to $(\kappa \times \kappa, <_{mlex}) \geq \kappa$. Conclusion: otp$(\kappa \times \kappa, <_{mlex}) = \kappa$. So $|\kappa \times \kappa| = \kappa$ □

5.21 Proposition:

Let $\kappa, \lambda \in C_n$.

(a) If both $\kappa, \lambda \in \omega$ then $\kappa + \lambda$ (the ordinal)= $\kappa + \lambda$ (the cardinal). And similarly with multiplication.

(b) If at least one of κ, λ is infinite, then $\kappa + \lambda = \kappa \cdot \lambda = \max\{\kappa, \lambda\}$. (Holds if both κ, λ nonzero).

Proof. (a) Exercise.

(b) Let $\mu = \max\{\kappa, \lambda\}$. Note: μ is infinite.

$\mu \to^1 (\{0\} \times \kappa) \cup (\{1\} \times \lambda) \to^2 \kappa \times \lambda \to^3 \mu \times \mu \to^4 \mu$

1. Obvious injection.

2. If $\min\{\kappa, \lambda\} \geq 2$. Identity if $\kappa \leq \lambda$.
3. Obvious injection by identity.
4. Injection by P.5.20.
And if $\min\{\kappa, \lambda\} \in \{0, 1\}$ then we can just do these cases separately. □

Exponentiation:
If A, B are sets, then ${}^{A}B$ denotes the set of all functions from A to B. That is,

$$ {}^{A}B = \{f \in \mathcal{P}(A \times B) | f : A \to B\} $$

5.22 Proposition:
For every set A we have that $\mathcal{P}(A) \sim^{A} \{0,1\}$.

Proof. The function $F : \mathcal{P}(A) \to^{A} \{0,1\}$ defined by

$$ F(x) = \chi_x = \text{the characteristic function of } x. $$

Is the desired bijection. Here,

$$ \chi_x : A \to \{0,1\} \text{ defined by} $$
$$ \chi_x = \begin{cases} 1 & \text{if} \quad a \in x \\ 0 & \text{if} \quad a \notin x \end{cases} $$

□

5.23 Theorem: Cantor
There is no surjection of a set A onto $\mathcal{P}(A)$.

Proof. Assume $f : A \to \mathcal{P}(A)$ is a surjection. Define $B \subseteq A$ by

$$ b \in B \iff b \notin f(b), \text{ i.e. } B = \{a \in A | a \notin f(a)\} $$

Claim: $B \notin \text{rng}(f)$. Otherwise, we would have some b such that $B = f(b)$. But then

$$ b \notin f(b) \Leftrightarrow b \in B \Leftrightarrow b \in f(b). $$

Contradiction. □

5.24 Corollary:
(a) $A \lneq \mathcal{P}(A)$
(b) If B has at least two elements, then $A \lneq ({}^{A}\{0,1\}) \leq ({}^{A}B)$.

Proof: Schröder-Bernstein

5.25 Proposition:
Let A, B, C be sets. Then
(a)${}^{(\{0\} \times B) \cup (\{1\} \times C)}A \sim ({}^{B}A) \times ({}^{C}A)$.
(b)${}^{(B \times C)}A \sim ({}^{C}({}^{B}A))$.
(c) ${}^{C}(A \times B) \sim ({}^{C}A) \times ({}^{C}B)$.
Proof: Exercise.

From now on we work with AC. Under this assumption, $\mathcal{P}(A)$ and the set ${}^{A}B$ are equinumerous to some cardinals.

5.26 Definition (AC):
Let κ, λ be cardinals. We define

$$\kappa^{\lambda} = \text{card}({}^{\lambda}\kappa).$$

5.27 Problem:
We know that $|\mathbb{R}| = 2^{\aleph_0}$.
What is the α such that $2^{\aleph_0} = \aleph_\alpha$? (What is the α such that $\aleph_\alpha = |\mathbb{R}|$?).

The statement: "There is no set $X \subseteq \mathbb{R}$ such that $\aleph_0 < |X| < |\mathbb{R}|$ " is called the **Continuum Hypothesis**, or **(CH)** . The above makes sense without AC. In the presence of AC:

$$\text{CH} \Leftrightarrow 2^{\aleph_0} = \aleph_1$$

Recall: If ZF / ZFC is consistent, the proof of consistency cannot be formalized in ZF. This requires some care in formulating the results. Recall: the statement "T is consistent" is abbreviated by "Con(T)".

1938 Gödel: Con(ZF)$\Rightarrow$ Con(ZF+AC+CH)

So ZFC cannot deny CH.

1963 Cohen: Con(ZF) $\Rightarrow$ Con(ZF+AC+$\neg$CH)

So ZFC cannot deny $\neg$CH.

The statement: "For every infinite set A there is no $X \subseteq \mathcal{P}(A)$ such that $|A| < |X| < |\mathcal{P}(A)|$." Is called the *Generalized Continuum Hypothesis, (GCH)*. This makes sense without AC.

In the presence of AC:

$$\text{GCH} \Longleftrightarrow (\forall\alpha)(2^{\aleph_\alpha} = \aleph_{\alpha+1})$$

Sierpiński: GCH $\Rightarrow$ AC

Gödel 1938: Con(ZF) $\Rightarrow$ Con(ZF+AC+GCH)

Solavay (sp?) 1967 (?): Let $\alpha \in O_n$. Then

$$\text{Con(ZF)} \Rightarrow \text{Con(ZF+AC+}2^{\aleph_0} > \aleph_\alpha)$$
$$\text{Con(ZF)} \Rightarrow \text{Con(ZF+AC+}2^{\aleph_0} > \aleph_{\alpha+1})$$

König early 1900s : $2^{\aleph_0} \neq \aleph_\omega$.