Axiomatic Set Theory I

A. C. Walczak-Typke

April 27, 2009

Contents

Chapter 1

Learning to Speak

1.1 The Motivation Behind Set Theory

Mathematicians in general work within so-called "naive set theory". That is, in a theory which is not axiomatized, and treating its objects, "sets", as platonic absolute objects. Traditionally, this is how natural and real numbers are approached in grade school mathematics classes. Sets are sets, and that is all. Unfortunately, this approach to set theory very quickly leads to contradictions. A well known example of this is the "set of all sets", in other words, *Russell's Paradox* (also known as Russell's Antinomy). Let us make it clear:

Theorem. *There is no set containing all sets.*

More formally:

Theorem. *Let R be the set of all sets not containing themselves. Then R is neither a member of itself, nor not a member of itself.*

I.e., let $R = \{x : x \notin x\}$. Then $R \in R \leftrightarrow R \notin R$.

With a naive approach, there is nothing in particular that stops us from making assertions such as, there is a set of all sets. This example shows that the naive approach to set theory is a bit unsafe, especially when we are talking about sets.

There are two threads that run through basic set theory: foundation via axiomatization, and different sizes of infinities. We will see that these two threads are closely connected, and both are very central to set theory. The discovery of different sizes of infinities was one of the driving forces behind the development of set theory. In fact, in some languages, set theory is refered to as the "theory of pluralities"! Despite this, we will make a slight emphasis on axiomatization. Another of the driving forces behind the development of set theory was the Hilbert Program, a program proposed by David Hilbert in the 1920's to formalize all of mathematics to a finite, complete set of axioms which are provably consistent. As we will see, this did not work out so well.

The difficulties caused by a lack of formalism, as illustrated by Russell's Paradox, are why we will begin this lecture with a discussion of the formal language of set theory, and a reminder of first order formal languages.

1.2 The First-Order Language of Set Theory

First, we define our "alphabet":

Definition 1.2.1. The *basic symbols* are $\wedge, \neg, \exists, (,), \in, =$, and v_j for every natural number j.

Intuition for Definition 1.2.1: The intuition behind these symbols is the following. $\wedge$ means the conjunction "and", $\neg$ is negation "not", $\exists$ is the existential quantifier "there is, there exists", the parenthesis will help with the readability of our sentences and formulas, $\in$ denotes the relation of membership ($x \in y$ means x is a member of y), $=$ is the relation of equality, and v_j are variables.

Now we will form words from these letters.

Definition 1.2.2. An *expression* is any finite sequence of basic symbols, such as $\in \wedge \wedge v_9(=)$.

Intuition for Definition 1.2.2: Similarly to natural languages like English and German, we can put together our letters. For example we can write "adkhkfd" and "banana". But, not all expressions have meaning, just as in our example. The sequence of letters "adkhkfd" means nothing, while "banana" does.

Intuition for Definition 1.2.3: The intuitive interpretation of the symbols determine which expressions are meaningful. These meaningful expressions are called *formulas.*

More precisely:

Definition 1.2.3. We definite (inductively) a *formula* to be an expression built using the following rules:

1. $v_j \in v_i$ and $v_j = v_i$ are formulas for all i and j;
2. if ϕ and ψ are formulas, then so are $(\phi) \wedge (\psi)$, $\neg(\phi)$ and $\exists v_i(\phi)$ for all i;

Abbreviations. We will use the following abbreviations:

- $\forall v_i(\phi)$ abbreviates the formula $\neg(\exists v_i(\neg(\phi)))$;
- $(\phi) \vee (\psi)$ stands for $\neg((\neg(\phi)) \wedge (\neg(\psi)))$;
- $(\phi) \rightarrow (\psi)$ abbreviates $(\neg(\phi)) \vee (\psi)$;
- $(\phi) \leftrightarrow (\psi)$ stands for $((\phi) \rightarrow (\psi)) \wedge ((\psi) \rightarrow (\phi))$
- $v_j \neq v_i$ and $v_j \notin v_i$ stand for $\neg(v_j = v_i)$ and $\neg(v_j \in v_i)$ respectively;
- we omit parentheses if their placement is clear from context;
- other letters of the Latin, Greek, or Hebrew alphabet are used as variables.
- $\forall x \in a\, \phi$ stands for $\forall x\, (x \in a \rightarrow \phi)$
- Similarly, $\exists x \in a\, \phi$ stands for $\exists x\, (x \in a \wedge \phi)$
- $\exists! x\phi$ is an abbreviation of $\exists x\, (\phi(x) \wedge (\forall y)(\phi(y) \rightarrow y = x))$. The intended meaning here is that there exists exactly one x such that ϕ holds.

Definitions 1.2.4. A *subformula* of a formula is a segment of a formula that itself constitutes a formula.

The *scope* of an occurrence of a quantifier $\exists v_i$ is the (unique) subformula beginning with that $\exists v_i$. An occurrence of a variable is called *bound* if it lies in the scope of a quantifier acting on that variable. Otherwise, a variable is called *free*.

Example 1. Look at

$$(\exists v_0\ (v_0 \in v_1)) \wedge (\exists v_1\ (v_2 \in v_1)).$$

In this example, the *subformulas* are $v_0 \in v_1$, $\exists v_0\ (v_0 \in v_1)$, $v_2 \in v_1$, $\exists v_1\ (v_2 \in v_1)$, and the whole formula $(\exists v_0\ (v_0 \in v_1)) \wedge (\exists v_1\ (v_2 \in v_1))$.

The *scope* of $\exists v_0$ in the example, is $\exists v_0\ (v_0 \in v_1)$.

The first occurrence of v_1 in the example is *free*, as is the occurrence of v_2. The second occurrence of v_1 is *bound*, as are the occurrences of v_0.

Intuition for Definitions 1.2.4: Intuitively, a formula expresses a property of its free variables. The bound variables are just used to make existential statements and are in a sense dummy variables.

We will sometimes present a formula as $\phi(x_1, \ldots, x_n)$ to emphasize its dependence (whatever that means) on $x_1, \ldots, x_n$. If $y_1, \ldots, y_n$ are other variables, $\phi(y_1, \ldots, y_n)$ denotes the formula that comes from substituting a y_i for each free occurrence of x_i. Such a substitution is called *free* or *legitimate* if no free occurrence of an x_i is in the scope of a quantifier $\exists y_i$. Here, the intuition is that $\phi(y_1, \ldots, y_n)$ says about $y_1, \ldots, y_n$ what $\phi(x_1, \ldots, x_n)$ said about $x_1, \ldots, x_n$. This may not be the case if the substitution is not free and some y_i winds up bound by a quantifier of ϕ. We will always assume that our substitutions are legitimate.

Definition 1.2.5. A *sentence* is a formula that has no free variables.

Intuition for Definition 1.2.5: Intuitively, a sentence states an assertion which is either true or false.

The axioms of set theory we will examine in this lecture, ZFC, are a certain set of sentences.

Now, we address how things can be proved.

Intuition: If S is a set of sentences and ϕ is a sentence, then intuitively, $S \vdash \phi$ means that one can prove from S by a purely logical argument in which the sentences of S may be quoted as axioms, but may not refer to the intended "interpretation" or "meaning" of the symbol $\in$.

Formally, we define $S \vdash \phi$ iff (= "if and only if" $=\Leftrightarrow$) there is a *formal deduction* of ϕ from S. That is, iff there is a finite sequence $\phi_1, \ldots, \phi_n$ of formulas such that ϕ_n is ϕ, and for each i, either ϕ_i is in S, or ϕ_i is a logical axiom, or ϕ_i follows from $\phi_1, \ldots, \phi_{i-1}$ by certain *rules of inference*.

If S is the empty set, and $S \vdash \phi$, then we write $\vdash \phi$ and say that ϕ is *logically valid*. If $\vdash (\phi \leftrightarrow \psi)$ then ϕ and ψ are *logically equivalent*.

If ϕ is a formula, a *universal closure* of ϕ is a sentence gotten by universally quantifying all free variables of ϕ.

Example 2. Let ϕ be the formula

$$x = y \to \forall z\, (z \in x \iff z \in y).$$

Then, $\forall x\, \forall y\, \phi$ and $\forall y\, \forall x\, \phi$ are universal closures of ϕ.

All universal closures of a formula are logically equivalent. If S is a set of sentences and ϕ is a formula, then $S \vdash \phi$ indicates that the universal closure of ϕ is provable from S.

We extend to formulas our notions of logical validity and logical equivalence, by saying that a formula is logically valid if its universal closure is. Similarly from logical equivalence. Using the notion of logical equivalence, we can make precise the idea that bound variables are dummy variables. If $\phi(x_1, \ldots, x_n)$ is a formula with only $x_1, \ldots, x_n$ free and $\phi'(x_1, \ldots, x_n)$ results from replacing the bound variables of ϕ with other variables, then ϕ and ϕ' are logically equivalent. This justifies the use of other letters to stand in for our "official" variables.

If S is a set of sentences, we say that S is *consistent* (symbolically written $\mathrm{Con}(S)$) if there does not exist a ϕ such that $S \vdash \phi$ and $S \vdash \neg\phi$. If S is inconsistent, then $S \vdash \psi$ for all ψ. Such S are thus of no interest. Notice that $S \vdash \psi$ iff $S \cup \{\neg\psi\}$ is inconsistent.

The fact that formal proofs are all finite gives us the following:

Theorem 1.2.6.

1. *If $S \vdash \phi$, then there is a finite $S_0 \subset S$ such that $S_0 \vdash \phi$;*
2. *If S is inconsistent, there is a finite $S_0 \subset S$ such that S_0 is inconsistent.*

Chapter 2

The Axioms of Set Theory

There is more than one possible axiomatization of set theory. In this semester we will concentrate on one - one that is generally accepted as the standard - so-called *ZFC set theory.* The letters stand for Zermelo, Fraenkel, and Choice, for two formulators of the axiom system and the 9th axiom. Zermelo formulated all but Axioms 8 and 5 by 1908. Further additions were made by Fraenkel and Skolem in the 1920's.

We underline that a *set* is anything whose existence is guaranteed by the following axioms.

2.1 Statement and discussion of the axioms of ZFC

There are 9 axioms and axiom schema of ZFC set theory, 10 if you count the 0th axiom. Different people number them differently!

Axiom 0 (Set Existence)**.**

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

Intuition: This axiom says that our universe, or domain, of sets is not empty - that we are actually talking about *something.*

Under most developments of classical formal logic, this axiom can be derived from the logical axioms. Alternatively, it can be derived from Axiom 6 (Infinity) below. Thus, this axiom does not need to be explicitly stated. We do so here for emphasis.

Axiom 1 (Extensionality (or Equality))**.**

$$\forall x\,\forall y\,((y = x) \leftrightarrow \forall z\,(z \in x \leftrightarrow z \in y)).$$

Intuition: The intuition behind the Axiom of Extensionality is that a set is determined by its members. Note that the implication $(y = x) \rightarrow \forall z(z \in x \leftrightarrow z \in y)$ is a theorem of logic, so really only the opposite implication is the important bit.

Axiom 2 ((Restricted) Comprehension Axiom Schema (or Separation Axiom Schema))**.**

For each formula $\phi \in \mathcal{L}(\in)$ without y free , the universal closure of the following is an axiom:

$$\exists y \, \forall x \, (x \in y \leftrightarrow x \in z \wedge \phi).$$

Note that in the above definition, y need not actually be used in ϕ, just if it is there it has to be bound.

Axiom 2 is not just one axiom, but rather a schema, a recipe or model, for making infinitely many axioms, one for each ϕ in which y is not a free variable.

Intuition: The idea behind this axiom is the formalization of the construction of sets of the form $\{x : P(x)\}$, where $P(x)$ is some property of x. Since we have formalized the notion of a property via formulas, one may simple-mindedly expect an axiom of the form

$$\exists y \, \forall x \, (x \in y \leftrightarrow \phi).$$

This would be the axiom scheme of (full) Comprehension. But, if we take ϕ to be the formula $x \notin x$, then we get Russell's Paradox! So, it would be a mistake to take full comprehension as an axiom!

So, instead, we use the property given by ϕ to "separate" from a set (z as written above) a subset having this property. We assert that y exists, and denote it by $\{x : x \in z \wedge \phi\}$. This y is then unique by Axiom 1, Extensionality. While the variable y is presumed not to be free, ϕ may have any number of other variables free. The free variables are considered to be parameters in this definition of a subset of z.

The requirement that y is not free eliminates the possibility of self-referential definitions of sets. For example: $\exists y \, \forall x \, (x \in y \leftrightarrow x \in z \wedge x \notin y)$, which would be inconsistent with the existence of a non-empty z.

If z is a set, then thanks to the restricted Comprehension axiom, we can form a set $\{x \in z : x \neq x\}$, which is a set with no member elements. By the Set Existence axiom, some set z exists, so there is a set with no elements. By Extensionality, the set with no elements is unique. So we can make the following:

Definition 2.1.1. $\emptyset$ is the unique set y such that $\forall x \, (x \notin y)$.

We can also prove using the restricted Comprehension axiom that there is no universal set, no set containing all sets.

Theorem 2.1.2.

$$\neg \exists z \, \forall x \, (x \in z).$$

Proof. Assume we do have such a universal set z. If there is such a set z that $\forall x \, (x \in z)$, then by the restricted Comprehension axiom schema, we can form the set $\{x \in z : x \notin x\}$. Because the set z is universal, this new set can be written $\{x : x \notin x\}$. This is a contradiction with Russell's Antinomy. $\square_{2.1.2}$

Abbreviations. At this point, we can also define some further abbreviations.

- Let $A \subseteq B$ abbreviates the formula $\forall x (x \in A \rightarrow x \in B)$.

From the axioms of logic, we have that $A \subseteq A$ and $\emptyset \subseteq A$.

The empty set $\emptyset$ is the only set that can be proven to exist from the axioms 0, 1, 2 so far. If we assume that the empty set is the only set in our domain, with $\in$ interpreted as the (vacuous) membership relation, then it is easy to see that the axioms so far hold in this interpretation. But, so do other (unwanted!) statements, such as $\forall x (x = \emptyset)$. Thus axioms cannot refute $\forall x\, (x = \emptyset)$. So, we need more axioms!

We give three further axioms for building sets, then will discuss them.

Axiom 3 (Pairing)**.**

$$\forall x\, \forall y\, \exists z\, (x \in z \wedge y \in z).$$

Intuition: The pairing axiom is meant to allow us to combine two sets.

By axioms 3,1,2 (Pairing, Extensionality, and restricted Comprehension), for all sets x and y there exists exactly one set whose elements are only x and y. We call this set $\{x, y\}$. The set $\{x\} = \{x, x\}$ is the set whose unique element is x. exercise?

We can now define:

Definition 2.1.3. A *(Kuratowski) ordered pair* is defined to be

$$\langle x, y \rangle = \{\{x\}, \{x, y\}\}.$$

Clearly, $\langle x, y \rangle = \langle x', y' \rangle \to x = x' \wedge y = y'$. exercise?

Axiom 4 (Union)**.**

$$\forall \mathcal{F}\, \exists A\, \forall Y\, \forall x\, (x \in Y \wedge Y \in \mathcal{F} \to x \in A).$$

Intuition: In the Union Axiom, we think of $\mathcal{F}$ as a family of sets, and postulate that every member of $\mathcal{F}$ is a subset of some set A, which will be called the union.

Together with Separation and Extensionality, the union axiom gives the smallest and unique set with the property mentioned above in the intuition. Thus we define:

Definition 2.1.4. The *union* of a family of sets $\mathcal{F}$, written $\bigcup \mathcal{F}$ is defined to be

$$\bigcup \mathcal{F} = \{x \in A : \exists Y \in \mathcal{F}\, (x \in Y)\}.$$

Definition 2.1.5. If $\mathcal{F}$ is a non-empty set, then we can also define the *intersection* of $\mathcal{F}$, $\bigcap \mathcal{F}$ to be

$$\bigcap \mathcal{F} = \{x : \forall Y \in \mathcal{F}\, (x \in Y)\}.$$

This intersection set exists since for each $b \in \mathcal{F}$ we have $\bigcap \mathcal{F} = \{x \in b : \forall y \in \mathcal{F}\, (x \in y)$, thus we can use restricted Comprehension. Uniqueness, as usual, follows from Extensionality.

If $\mathcal{F} = \emptyset$, then $\bigcup \mathcal{F} = \emptyset$. In this case, $\bigcap \mathcal{F}$ would have to be the set of all sets, which we have shown does not exist. So, the assumption that $\mathcal{F}$ is non-empty is a vital one.

Abbreviations. We have the following abbreviations:

- $A \cup B = \bigcup\{A, B\}$;
- $A \cap B = \bigcap\{A, B\}$;
- $A \setminus B = \{x \in A : x \notin B\}$.

Axiom 5 (Replacement Axiom Schema)**.**
For each $\phi \in \mathcal{L}(\in)$ without Y free, the universal closure of the following is an axiom:

$$\forall x \in A\, \exists! y\, \phi(x, y) \rightarrow \exists Y\, \forall x \in A\, \exists y \in Y\, \phi(x, y).$$

Intuition: This, like axiom 2 (restricted Comprehension), is an axiom schema, and so gives us infinitely many axioms - one for each ϕ. The intuition behind this axiom is that ϕ defines a function on A. Then, there should exist a set that is the image of the function, i.e., $Y = \{y : \exists x \in A\, \phi(x, y)\}$. This Y should be a set, and of size not greater than A.

Definition 2.1.6. The Replacement Schema allows us to define the *cartesian product* $A \times B$ of finitely many factors. We do this in a couple of steps. First, for every $y \in B$ we have $\forall x \in A\, \exists! z\, (z = \langle x, y\rangle)$. This allows us to define, using replacement, the set

$$\mathrm{prod}(A, y) = \{z : \exists x \in A\; z = \langle x, y\rangle\}.$$

Now, $\forall y \in B\, \exists! z\, (z = \mathrm{prod}(A, y))$. Again, thanks to the axiom of replacement, we can define

$$\mathrm{Prod}(A, B) = \{\mathrm{prod}(A, y) : y \in B\}.$$

Finally, we define

$$A \times B = \bigcup \mathrm{Prod}(A, B).$$

Other important notions can be defined already at this point in the development of the theory.

——————— HERE ENDED WINTER 2006 LECTURE 1 ———————

Definitions 2.1.7. A *relation* is a set R all of whose elements are ordered pairs.
For a given relation R we define the *domain* and *range* of R:

$$\mathrm{dom}(R) = \{x : \exists y\, (\langle x, y\rangle \in R)\},$$

$$\mathrm{rng}(R) = \{y : \exists x\, (\langle x, y\rangle \in R)\}.$$

For a relation R we define its *inverse*

$$R^{-1} = \{\langle x, y\rangle : \langle y, x\rangle \in R\}.$$

Remark 2.1.8. *The construction of the domain and range does not require the axiom of replacement. Notice that both are subsets of* $\bigcup\bigcup R$.

The definitions of range, domain, and inverse make sense for any set R. However, if R is a relation, then we have some nice properties. For example, $R \subseteq \mathrm{dom}(R) \times \mathrm{rng}(R)$. *Also,* $R = (R^{-1})^{-1}$.

Note that traditionally we often write xRy instead of $\langle x, y\rangle \in R$.

Definitions 2.1.9. f is called a *function* iff f is a relation and

$$\forall x \in \operatorname{dom}(f)\, \exists! y \in \operatorname{rng}(f)\, (\langle x, y\rangle \in f).$$

We write $f : A \longrightarrow B$ to mean that f is a function such that $\operatorname{dom}(f) = A$ and $\operatorname{rng}(f) = B$.

If $f : A \longrightarrow B$ and $x \in A$, then $f(x)$ denotes the unique y such that $\langle x, y\rangle \in f$.

If $C \subseteq A$, then $f \restriction C = f \cap C \times B$ is the *restriction* of f to C.

Further, $f"C = \operatorname{rng}(f \restriction C) = \{f(x) : x \in C\}$. Sometimes this is also noted as $f[C]$ (also $f * C$ or $f \to (C)$).

A function $f : A \longrightarrow B$ is called *1-1* ("one-to-one") or an *injection* if f^{-1} is a function. The function f is called *onto* or a *surjection* if $\operatorname{rng}(f) = B$. A function that is both a surjection and an injection is called a *bijection*.

We can use functions to compare relations.

Definition 2.1.10. If R and S are relations and A and B are sets, then $\langle A, R\rangle$ and$\langle B, S\rangle$ are *isomorphic* ("similar") if there exists a bijection (remember: 1-1 and onto function) $f : A \longrightarrow B$ such that

$$\forall x, y \in A\, xRy \iff f(x)Sf(y).$$

This function is called an *isomorphism*. We denote the existence of such an isomorphism as $\langle A, R\rangle \cong \langle B, S\rangle$

So far, the axioms we have presented only allow us to build finite sets (whatever finite formally means). This means we cannot define, say, the set of all natural numbers. The next axiom, the axiom of infinity rectifies this problem.

Axiom 6 (Infinity)**.**

$$\exists x\, (\emptyset \in x \land \forall y \in x\, (y \cup \{y\} \in x)).$$

Abbreviations.
- Let $S(x) = x \cup \{x\}$. We call S the *successor* function (for reasons that will become clear later.)

So, we can restate the axiom of Infinity as

$$\exists x\, (\emptyset \in x \land \forall y \in x\, S(y) \in x)).$$

We call a set x that satisfies the axiom of infinity an *inductive* set. Later, we will define rigorously what "infinite" means, and that an inductive set is necessarily infinite.

——————— HERE ENDED WINTER 2007 LECTURE 1 ———————

Axiom 7 (Powerset)**.**

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

Set theory, unlike other most other branches of mathematics, has at its roots the work of one man: Georg Cantor. Cantor made the observation in 1873 that there are "more" transcendental numbers, and so more real numbers, than there are natural numbers. Zermelo later developed the axioms we are studying to take care of the paradoxes that appeared because of Cantor's less formal approach.

The infinity axiom only allows us to get sets that are the same size as the natural numbers. We need the powerset to get bigger infinities, such as the infinity that is the size of the real numbers.

Axiom 8 (Foundation (also: Axiom of Regularity)**.**

$$\forall x\,(\exists y \in x \;\rightarrow \exists y \in x\,(\neg\exists z\,(z \in y \wedge z \in x)))$$

The Axiom of Foundation is an axiom that people tend to forget about. Nevertheless, it is very important in certain inductive constructions. We will concentrate more on this axiom later in the semester.

Axiom 9 (Axiom of Choice)**.**

$$\forall \mathcal{F}\,((\forall S \in \mathcal{F}\,(S \neq \emptyset)) \rightarrow (\exists f\,\forall S \in \mathcal{F}\, f(S) \in S))$$

Intuition: The idea behind the axiom of choice is that for any family of sets that are non-empty, there is a function that picks out one element out of each member of the family.

There are many equivalent formulations of the Axiom of Choice. We'll show some of these later. This was at one time a bit of a controversial axiom (though most mathematicians nowadays accept the axiom as useful and "correct"). A lot of modern mathematics doesn't work quite so well if the axiom of choice is not assumed. For example, a lot of analysis and topology gets very ugly and messy very quickly without this axiom. We'll point out where it is used in the development of set theory as we go along.

2.2 Partial Axiom Systems

Certain theorems can be proven using only part of the full ZFC system of axioms. Here we list certain standard partial systems.

ZFC All the axioms presented here. 0-9

ZF Axioms 0-8. Here the Axiom of Choice is omitted.

ZF^- Axioms 0-7. So, in particular, the Axiom of Foundation and Choice are omitted.

$ZF^- - P$ Axioms 0-6. So, Choice, Foundation, and the Powerset Axiom are omitted.

$ZF - P$ Axioms 0-6 and 8. So, no Choice or Powerset.

The systems ZFC^-, $ZFC^- - P$, and $ZFC - P$ are defined in the obvious way.

We will sometimes note when a theorem can be proved within one of these partial systems.

——————— HERE ENDED SPRING 2009 WEEK 1 (2hrs 15 min) ———————

Chapter 3

Orders and Ordinals

3.1 Orders

We now concentrate on a particular kind of relation: that of the ordering.

Definition 3.1.1. A *linear ordering* (or *total ordering*) is a pair $\langle A, R\rangle$ where A is a set and R is a relation that linearly orders A. That is, R is

- *transitive,* i.e. $\forall x, y, z \in A\, xRy \;\wedge yRz \rightarrow xRz$;
- *irreflexive,* i.e. $\forall x \in A\, \neg(xRx)$;
- *linear,* i.e. $\forall x, y \in A\, xRy \vee x = y \vee yRx$.

Notice that we are not assuming that $R \subseteq A \times A$. Thus, if $\langle A, R\rangle$ is a linear ordering and $B \subseteq A$, then $\langle B, R\rangle$ is also a linear ordering.

We will be particularly concerned with a particular type of linear ordering:

Definition 3.1.2. A relation R is a *well ordering* on A if $\langle A, R\rangle$ is a linear ordering and every non-empty subset of A has a R-least element.

Examples of well-orderings include: $(\mathbb{N}, <)$ and $(\{0, 1, 2\}, <)$. The following are NOT well-orderings:$(\mathbb{Z}, <)$,$(\mathbb{Q}, <)$, and $(\mathbb{R}, <)$.

A basic tool for studying well orderings is the set of predecessors of an element:

Definition 3.1.3. Let $\langle A, R\rangle$ be an ordering. If $x \in A$, then the *initial segment determined by* x is defined as

$$\mathrm{pred}(A, x, R) = \{y \in A : yRx\}.$$

A basic property of well orderings is as follows:

Lemma 3.1.4. *If $\langle A, R\rangle$ is a well ordering, then for all $x \in A$, $\langle A, R\rangle \not\cong \langle \mathrm{pred}(A, x, R), R\rangle$*

Proof. Assume, to the contrary, that $f : A \longrightarrow \mathrm{pred}(A, x, R)$ is an isomorphism. Then $f(x)Rx$, by definition of an isomorphism. Let z be the R-least element of the set $X = \{y \in A : f(y)Ry\}$, which exists because we have assumed that R is a well ordering. But then $f(z)Rz$. Thus immediately we have $ff(z)Rf(z)$. Thus, $f(z) \in X$, which means that z wasn't the R-least element in X after all. A contradiction. $\square_{3.1.4}$

A further very important property of well orders is given by:

Lemma 3.1.5. *If $\langle A, R\rangle$ and $\langle B, S\rangle$ are isomorphic well-orderings, then the isomorphism between them is unique.*

Proof. For a contradiction, let f and g be two different isomorphisms between the isomorphic well orderings $\langle A, R\rangle$ and $\langle B, S\rangle$. Let $X = \{y \in A : f(y) \neq g(y)\}$. Since we have assume that $f \neq g$, it must be that $X \neq \emptyset$. Let z be the R-least element of the set X. Since $f(z) \neq g(z)$, then either $f(z)Sg(z)$ or $g(z)Sf(z)$. Let us assume that $f(z)Sg(z)$. Let $t \in A$ be such that $g(t) = f(z)$. Then, $g(t) \neq g(z)$, and therefore $t \neq z$, so further, we have $f(t) \neq g(t) = f(z)$. So, $g(t)Sg(z)$, which gives tRz because g is an isomorphism. This means that t is R-smaller than z and $t \in X$. Contradiction. $\square_{3.1.5}$

This leads us to the fact that any two well orderings are comparable.

Theorem 3.1.6. *Let $\langle A, R\rangle$ and $\langle B, S\rangle$ be two well orderings. Then, exactly one of the following holds:*

1. $\langle A, R\rangle \cong \langle B, S\rangle$;
2. $\exists y \in B\,(\langle A, R\rangle \cong \langle \mathrm{pred}(B, y, S), S\rangle)$;
3. $\exists x \in A\,(\langle \mathrm{pred}(A, x, R), R\rangle \cong \langle B, S\rangle)$.

Proof. Let

$$f = \{\langle v, w\rangle : v \in A \wedge w \in B \wedge \langle \mathrm{pred}(A, v, R), R\rangle \cong \langle \mathrm{pred}(B, w, S), S\rangle\};$$

here f is an isomorphism from some initial segment of A onto some initial segment of B. Use the previous lemmas to show that these initial segments cannot both be proper. details of this as exercise? $\square_{3.1.6}$

At this point, we can mention a statement that is equivalent to the Axiom of Choice, Axiom 9. This statement is often given as THE statement of the Axiom of Choice.

Axiom (9', Well-ordering Principle (Zermelo's Theorem))**.**

$$\forall A\,\exists R\,(R \text{ well orders } A).$$

Theorem 3.1.7. *The following statements are equivalent:*

AC Axiom of choice

WOP Well-ordering Principle

We postpone the proof of Theorem 3.1.7 until the next section.

3.2 Ordinals

We begin with some definitions:

Definition 3.2.1. A set z is *transitive* if every element of z is also a subset of z.

Examples of transitive sets are: $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\},\emptyset\}$, and $\{\{\{\emptyset\}\},\{\emptyset\},\emptyset\}$. On the other hand, $\{\{\emptyset\}\}$ is not transitive.

Definition 3.2.2. A set α is called an *ordinal* if it is transitive and well-ordered by $\in$.

There is a formal subtlety here: formally, the statement "α is well-ordered by $\in$" means that $\langle \alpha, \in_\alpha \rangle$ is a well-order, where $\in_\alpha = \langle \{\langle x, y\rangle \in \alpha \times \alpha : x \in y\}\rangle$. We make this distinction because one must differentiate between the relation $\in$, which is a relation in the sense of our formal language of set theory, and the relation $\in_\alpha$ that well-orders α. We need the latter to be a set, and hence part of the domain of things our formal language talks about, that is $\in_\alpha$ is a relation in the sense that it is a set composed of ordered pairs.

When we talk about ordinals, we do not explicitly mention $\in_\alpha$. So, we will write $\alpha \cong \langle A, R\rangle$ instead of $\langle \alpha, \in_\alpha\rangle \cong \langle A, R\rangle$, and when $\beta \in \alpha$, we write $\operatorname{pred}(\alpha, \beta)$ instead of $\operatorname{pred}(\alpha, \beta, \in_\alpha)$.

Theorem 3.2.3.

1. *If α is an ordinal and $y \in \alpha$, then y is also an ordinal and $y = \operatorname{pred}(\alpha, y)$;*
2. *If α and β are ordinals and $\alpha \cong \beta$, then $\alpha = \beta$;*
3. *If α and β are ordinals, then exactly one of the following holds: $\alpha \in \beta$, $\beta \in \alpha$, or $\alpha = \beta$;*
4. *If α, β, and γ are ordinals, $\alpha \in \beta$ and $\beta \in \gamma$, then $\alpha \in \gamma$;*
5. *If C is a non-empty set of ordinals, then $\exists \alpha \in C\, \forall \beta \in C\, (\alpha \in \beta \vee \alpha = \beta)$.*

Proof. **(1):** Let $y \in \alpha$. Then $y \subseteq \alpha$ because α is transitive. If y itself is not transitive, then there is some $x \in y$ such that $x \not\subseteq y$. Then, let $z \in x$ be such that $z \notin y$. But, since both z and y are elements of α, then either $z = y$ or $y \in z$, because α is ordered by $\in$. Both of these possibilities contradict the fact that $\in$ well-orders α (for example $x \in y \in z$ but $x \notin y$!). Therefore, y must be transitive. Because $y \subseteq \alpha$, $\in$ well-orders y.

(2): Notice first that because α is a well-ordering, either $\alpha = \emptyset$ or $\emptyset \in \alpha$. Now, if $\alpha \cong \beta$, then by Lemma 3.1.5, the isomorphism $f : \alpha \longrightarrow \beta$ is unique. Of course, $f(\emptyset) = \emptyset$. If f is not the identity mapping, then let γ be the first element of α such that $f(\gamma) \neq \gamma$. It is easy to check that such a thing does not exists (there will be a loop). Details as exercise?

(3): To prove this, use (1), (2), and Theorem 3.1.6. If more than one of the possibilities were to occur, then this would imply the existence of an x such that $x \in x$, which would in turn imply that $\in$ is not irreflexive.

(4): This is an obvious result of the other things we have shown.

(5): Thanks to (3), it suffices to show that $\exists x \in C\, (x \cap C = \emptyset)$. Let $x \in C$ be arbitrary. If $x \cap C \neq \emptyset$, then, since x is well-ordered by $\in$ (because it is an ordinal, and C is a set of ordinals), there is a $\in$-least element y of $x \cap C$. Then $y \cap C = \emptyset$. $\square_{3.2.3}$

——————— HERE ENDED SPRING 2009 WEEK 2 (1hr 15 min) ———————

Theorem 3.2.3 implies that the *set* of all ordinals, if it existed, would itself be an ordinal. This is the so-called Burali-Forti paradox. Precisely:

Theorem 3.2.4 (Burali-Forti paradox)**.**

$$\neg\exists z\,\forall x\,(x \textit{ is an ordinal } \rightarrow x \in z).$$

Proof. If there were such a z, then we would have a set ON such that

$$ON = \{x : x \text{ is an ordinal }\}.$$

Then ON is transitive by (1) of Theorem 3.2.3 and well-ordered by $\in$ (by (3), (4), and (5) of the same Theorem). Thus ON would be an ordinal. But, as pointed out in the proof of Theorem 3.2.3, no ordinal is a member of itself. $\square_{3.2.4}$

Lemma 3.2.5. *If A is a transitive set of ordinals, then A itself is an ordinal.*

proof of the above lemma is clear from the definitions exercise?

The following gives us a main point of ordinals.

Theorem 3.2.6. *If $\langle A, R\rangle$ is a well-ordering then there exists a unique ordinal C such that $\langle A, R\rangle \cong C$.*

Proof. **Uniqueness** is a result of Theorem 3.2.3 (2).

Existence: Let $B = \{a \in A : \exists x\,(x \text{ is an ordinal } \wedge \text{pred}(A, a, R) \cong x\}$. Then, we can define on B a function f such that for every $a \in B$,

$$f(a) = \text{ the unique ordinal } x \text{ such that } \text{pred}(A, a, R) \cong x.$$

Let $C = \text{rng}(f)$. By the Replacement Axiom, C is a set. Using Lemma 3.2.5, one can see that C is an ordinal (just need to check transitivity!). One can also easily see that f is an isomorphism between $\langle B, R\rangle$ and C. Now, either $A = B$, in which case we are done, or there is some $b \in A$ such that $B = \text{pred}(A, b, R)$. In the latter case, this would mean that $b \in B$, which is not possible. $\square_{3.2.6}$

Remark 3.2.7.

1. *The proof of Theorem 3.2.6 used the axiom of Replacement in an essential way to justify the existence of the set f. Formally: let $\phi(a, x)$ be the formula asserting that $\langle \text{pred}(A, a, R), R\rangle \cong x$. Then, $\forall a \in B\,\exists! x\,\phi(a, x)$. So, by Replacement (and restricted Comprehension) one can form the set $C = \{x : \exists a \in B\,\phi(a, x)\}$, then we use restricted Comprehension to define $f \subset B \times C$.*

2. *If one drops the axiom of Replacement from ZFC, then one can develop much of usual mathematics, but one cannot then prove Theorem 3.2.6.*

3. *Theorem 3.2.6 allows us to use ordinals as representatives of well-order types.*

Definition 3.2.8. If $\langle A, R\rangle$ is a well-ordering, then $\text{type}(\langle A, R\rangle)$ is the unique ordinal α such that $\langle A, R\rangle \cong \alpha$.

Definition 3.2.9. If X is a set of ordinals, then $\sup(X) = \bigcup X$ and, if $X \neq \emptyset$, $\inf(X) = \bigcap X$.

Notation: From now on, we will use small Greek letters to stand for ordinals. So, for example, we will write $\exists \alpha\, \phi$ to mean $\exists x\, (x \text{ is an ordinal } \wedge \phi)$. Also, since $\in$ orders the ordinals, we will write $\alpha < \beta$ to mean $\alpha \in \beta$, and $\alpha \leqslant \beta$ to mean $\alpha \in \beta \vee \alpha = \beta$.

Lemma 3.2.10.

1. $\forall \alpha, \beta\, (\alpha \leqslant \beta \iff \alpha \subseteq \beta)$.

2. *If X is a set of ordinals, then* $\sup(X)$ *is the smallest ordinal that is $\geqslant$ than all the ordinals in X. Similarly, if $X \neq \emptyset$, then* $\inf(X)$ *is the smallest ordinal in X.* *proof of this little fact as exercise?*

———————— HERE ENDED WINTER 2006 LECTURE 2 ————————

3.3 The Axiom of Infinity and the fundamentals of Peano Arithmetic

The first few ordinals are the natural numbers, which are used to count finite sets. If we assume the Axiom of Choice, Theorem 3.1.7 (AC $\Leftrightarrow$ WOP) (which we have not yet proved) means that we can well-order every set. Theorem 3.2.6 promises that we can count each well-ordered set with an ordinal. So, assuming AC, we can count each set with an ordinal.

We can extend the definition of many of the standard arithmetic operations that are familiar from the natural numbers to the ordinals.

Definition 3.3.1. We define the *successor* of an ordinal:

$$S(\alpha) = \alpha \cup \{\alpha\}.$$

A simple lemma: exercise?

Lemma 3.3.2. *For any ordinal α,*

- *$S(\alpha)$ is an ordinal;*
- $\alpha < S(\alpha)$*;*
- $\forall \beta\, (\beta < S(\alpha) \iff \beta \leqslant \alpha)$.

Definition 3.3.3. An ordinal α is called a *successor ordinal* if $\exists \beta\, (\alpha = S(\beta)$. An ordinal α is a *limit ordinal* iff $\alpha \neq \emptyset$ and α is not a successor ordinal.

Now we can formally define the natural numbers:

Definition 3.3.4. $0 = \emptyset$, $1 = S(0)$, $2 = S(1)$, $3 = S(2)$, $4 = S(3)$,... etc.

So, $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, ..., etc.

Definition 3.3.5. An ordinal α is a *natural number* iff $\forall \beta \leqslant \alpha\, (\beta = 0 \vee \beta \text{ is a successor ordinal})$

Intuition: The natural numbers are obtained by applying the successor function S to $\emptyset$ finitely many times. If β is the smallest ordinal which cannot be obtained in this manner, then β cannot be a successor. So, neither β, nor any ordinal greater than β, can be a natural number.

Many mathematical arguments use the concept of the *set* of natural numbers. It is the Axiom of Infinity that allows us to define this set. Recall that it is:

$$\exists x\,(\emptyset \in x \wedge \forall y \in x\,(y \cup \{y\} \in x)).$$

Intuition: If a set x satisfies the Axiom of Infinity, then "by induction", x contains all of the natural numbers.

More formally: Suppose x satisfies Infinity, and suppose n is a natural number and $n \notin x$. By assumption, $0 \in x$, so $n \neq 0$. This means that $n = S(m)$ for some m. Then, $m < n$, m is a natural number, and $m \notin x$. From this we get that $m \setminus x \neq \emptyset$. Let k be the smallest element of $m \setminus x$. If we apply this same argument to k, we get an $l < k$ such that $l \in m \setminus x$, which leads to a contradiction.

By the axiom of Comprehension, there exists a set of natural numbers. (formal version of below definition: $\omega = \{z \in x : z \text{ is a natural number.}\}$

Definition 3.3.6. ω is the set of natural numbers.

The set ω is an ordinal by Lemma 3.2.5. All ordinals smaller than ω (i.e. the elements of ω, are either 0 or successors. So, ω is a limit ordinal (since otherwise it would itself be a natural number), and hence is the smallest limit ordinal. So, in essence, the Axiom of Infinity is equivalent to the existence of a limit ordinal.

The set of natural numbers ω satisfies the Peano Postulates (Peano Axioms):

Theorem 3.3.7. *ω satisfies the Peano Postulates:*

1. $0 \in \omega$;
2. $\forall n \in \omega\,(S(n) \in \omega)$;
3. $\forall n, m \in \omega\,(n \neq m \rightarrow S(n) \neq S(m))$;
4. *(Induction)* $\forall X \subset \omega\,((0 \in X \wedge \forall n \in X\,(S(n) \in X)) \rightarrow X = \omega)$.

Proof.

1. 0 is a natural number.
2. For every natural number n, $S(n)$ is also a natural number.
3. If $S(n) = S(m)$, then we have $n \cup \{n\} = m \cup \{m\}$. Then we have $n = \sup(n \cup \{n\}) = \sup(m \cup \{m\}) = m$.
4. Assume $X \neq \omega$ satisfies the induction requirements. This means that $\omega \setminus X \neq \emptyset$. Then, let $n = \min(\omega \setminus X)$. Then it must be that $n \neq 0$, since this would mean that $X = \omega$. So, this means that $n = S(m)$ for some m. Then, $m \in X$ because we assumed n to be minimal not in X. But, by assumption, $n = S(m) \in X$, a contradiction.

$\square_{3.3.7}$

Now that we have the natural numbers and the Peano postulates, we could for the moment forget about ordinals, and develop elementary mathematics from here: construct the integers, the rationals, then use the Power-Set axiom to develop the real numbers. The first step to doing this would be to define $+$ and $\cdot$. We will not do that, but instead we will define $+$ and $\cdot$ on all the ordinals.

——————— HERE ENDED WINTER 2007 LECTURE 2 ———————

3.4 Ordinal Addition and Multiplication

Now we define some basic arithmetic operations on the ordinals.

Definition 3.4.1. $\alpha + \beta = \text{type}(\langle \alpha \times \{0\} \cup \beta \times \{1\}, R\rangle)$, where the relation R is defined as follows:

$$R = \{\langle\langle\zeta, 0\rangle, \langle\eta, 0\rangle\rangle : \zeta < \eta < \alpha\} \cup \{\langle\langle\zeta, 1\rangle, \langle\eta, 1\rangle\rangle : \zeta < \eta < \beta\} \cup (\alpha \times \{0\}) \times (\beta \times \{1\}).$$

Intuition: When learning addition in first grade, the analogy is that $2 + 5$ means that if I lay down 2 pieces of chocolate followed by 5 carrots, I will have a row of 7 sweet things. The idea here is the same. Less formally, the mess above just means that the elements $\alpha \times \{0\}$ ordered like α precede the elements of $\beta \times \{1\}$ ordered like β.

Lemma 3.4.2. *For arbitrary ordinals α, β, and γ, we have:*

1. *(Associativity of addition) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$;*
2. *$\alpha + 0 = \alpha$;*
3. *$\alpha + 1 = S(\alpha)$;*
4. *$\alpha + S(\beta) = S(\alpha + \beta)$;*
5. *if β is a limit ordinal, then $\alpha + \beta = \sup\{\alpha + \zeta : \zeta < \beta\}$.*

Note that $+$ **is not always commutative!!** For example $\omega + 1 \neq 1 + \omega = \omega$. However, on the natural numbers, the operation is commutative.

Proof. The proof is comes straight from the definition. For example, for (1), notice that both $\alpha + (\beta + \gamma)$ and $(\alpha + \beta) + \gamma$ are isomorphic to the set $\alpha \times \{0\} \cup \beta \times \{1\} \cup \gamma \times \{2\}$. $\square_{3.4.2}$

Now, we define ordinal multiplication $(\cdot)$.

Definition 3.4.3. For ordinals α and β, we define $\alpha \cdot \beta = \text{type}(\langle \beta \times \alpha, R\rangle)$, where R is the lexicographic relation on $\beta \times \alpha$. I.e.

$$\langle\zeta, \eta\rangle R \langle\zeta', \eta'\rangle \Leftrightarrow (\zeta < \zeta' \vee (\zeta = \zeta' \wedge \eta < \eta')).$$

Intuition: Again, the intuition is the same as in elementary school: $4 \cdot 5$ is counting 4 chairs 5 times.

From the definition, we can easily get the following lemma: no proof. maybe as an exercise.

Lemma 3.4.4. *For arbitrary ordinals α, β, and γ, we have the following:*

1. $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$;
2. $\alpha \cdot 0 = 0$;
3. $\alpha \cdot 1 = \alpha$;
4. $\alpha S(\beta) = \alpha\beta + \alpha$;
5. *if β is a limit ordinal, then* $\alpha \cdot \beta = \sup\{\alpha\zeta : \zeta < \beta\}$;
6. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Note that ordinal multiplication is NOT COMMUTATIVE! For example: $2\omega = \omega \neq \omega 2$. Similarly, multiplication is not distributive from the right: $(1+1)\omega = \omega \neq \omega + \omega$. However, on the natural numbers, the operation is both commutative and distributive.

Natural numbers let us deal with finite sequences:

Definition 3.4.5. (a) A^n is the set of all functions from n into A.

(b) $A^{<\omega} = \bigcup\{A^n : n \in \omega\}$.

With this definition, $A \times A$ is not the same thing as A^2. However, there is a 1-1 correspondence between them.

Note that it is not obvious that the above definition 3.4.5 makes sense without the Power-set axiom. This is done thus: Let $\phi(n, y)$ be a formula that says that $\forall s\, (s \in y \iff s$ is a function from n into $A)$. Then, using induction on n (via the Peano Axioms, for example), one shows that, using Extensionality, $\forall n\, \exists! y\, \phi(n, y)$. At the inductive step, we use the Replacement Axiom as well as identifying A^{n+1} with $A^n \times A$. Again, by Replacement, we can form the set $\{y : (\exists n \in \omega)\, \phi(n, y)\} = \{A^n : n < \omega\}$. Finally, using the Union Axiom, we have $A^{<\omega}$.

One generally thinks of the elements of A^n as sequences of elements of A of length n.

Definition 3.4.6. For every n, $\langle x_0, x_1, \ldots, x_{n-1}\rangle$ is a function s with domain n, such that $s(0) = x_0, \ldots, s(n-1) = x_{n-1}$.

Note that in the case of $n = 2$, the above definition does not agree with our earlier definition of the Kuratowski ordered pair. The Kuratowski definition is useful for introducing basic properties of relations and functions. On the other hand, the definition above is more convenient when dealing with sequences of varying lengths. In cases where it matters, we will explicitly indicate which definition we are using.

Generally if s is a function such that $\mathrm{dom}(s) = I$, then we can think of I as an index set, and of s as a sequence that is indexed by I. Thus, we will often write s_i instead of $s(i)$.

Definition 3.4.7. If s and t are sequences such that $\mathrm{dom}(s) = \alpha$ and $\mathrm{dom}(t) = \beta$, then the function $s^\frown t$ with domain $\alpha + \beta$ is defined by $(s^\frown t \restriction \alpha) = s$ and $(s^\frown t)(\alpha + \zeta) = t(\zeta)$ for all $\zeta < \beta$.

3.5 Classes, Transfinite Induction, and Transfinite Recursion

As we have established, sets of the form $\{x : \phi(x)\}$ do not have to exist. It is however, quite convenient to think about such collections. Since they lie outside of the domain that is describable with our axioms, one should never use them in formal proofs.

Informally, we call collections of the form $\{x : \phi(x)\}$ *classes*. Here, we allow ϕ to have other variables than x, and think about them as parameters on which our class depends. A *proper class* is a class that is not a set (because it is "too big"). The Axiom of Restricted Comprehension says that a subclass of a set is a set. Boldface letters are often used to denote classes. Two classes, which we have shown to be proper classes are given in the following:

Definition 3.5.1.

$$\mathbf{V} = \{x : x = x\}$$

$$\mathbf{ON} = \{x : x \text{ is an ordinal.}\}.$$

Formally, proper classes do not exist, and expressions containing them must be thought of as abbreviations for expressions not involving them. For example, $x \in \mathbf{ON}$ is an abbreviation of the formula "x is an ordinal". The expression $\mathbf{ON} = \mathbf{V}$ abbreviates the (false!) sentence $\forall x\,(x \text{ is an ordinal} \iff x = x)$.

Formally, there is no difference between a formula and a class; the difference is only in the informal presentation. So, we could, instead of the above definition, consider the class $\mathbf{ON}$ an abbreviation of the formula $\mathbf{ON}(x)$ which says that "x is an ordinal". The usefulness of thinking about $\mathbf{ON}$ as a collection of sets is, for example, such that we can write $\mathbf{ON} \cap y$ instead of the formal $\{x \in y : x \text{ is an ordinal.}\}$. Any of our defined predicates and functions can be thought of as a class. For example, we can think of the union operation as defining a class $\mathbf{UN} = \{\langle\langle x, y\rangle, z\rangle : z = x \cup y\}$. Intuitively, $\mathbf{UN} : \mathbf{V} \times \mathbf{V} \longrightarrow \mathbf{V}$. This motivates using an abbreviation like $\mathbf{UN} \restriction (a \times b)$ for

$$\{\langle\langle x, y\rangle, z\rangle : z = x \cup y \wedge x \in a \wedge y \in b\}.$$

This kind of abbreviation obtained with a class is very useful when discussing general properties of classes. Asserting that a statement is true for all classes is equivalent to asserting that a statement is a theorem schema. An example of this are the principles of induction and recursion on $\mathbf{ON}$.

Theorem 3.5.2 (Transfinite Induction on $\mathbf{ON}$). *If $\mathbf{C} \subset \mathbf{ON}$ and $\mathbf{C} \neq \emptyset$ then $\mathbf{C}$ has a least element.*

Proof. The proof is exactly like the proof of Theorem 3.2.3(5), which stated the same thing for C being a set. Fix $\alpha \in \mathbf{C}$. If α is not the least element of $\mathbf{C}$, then $\alpha \cap \mathbf{C}$ is a nonempty set by Replacement. By Theorem 3.2.3(5), let β be the smallest element of $\alpha \cap \mathbf{C}$. Clearly, β is then the smallest element of $\mathbf{C}$. $\square_{3.5.2}$

Mathematically, Theorems 3.2.3(5) and 3.5.2. are very similar. Formally, there is an enormous difference between them. Theorem 3.2.3(5) is the abbreviation of one provable sentence. On the other hand, Theorem 3.5.2 is a theorem schema which represents infinitely many theorems.

It is possible, of course, to state Theorem 3.5.2 without classes. To do this, we would have to say: for each formula $\mathbf{C}(x, z_1, \ldots, z_n)$, the following is a theorem:

$$\forall z_1, \ldots, z_n \, ((\forall x \, (\mathbf{C} \to x \text{ is an ordinal}) \wedge \exists x \, \mathbf{C}) \to \\ \to (\exists x \, (\mathbf{C} \wedge \forall y \, (\mathbf{C}(y, z_1, \ldots, z_n) \to y \geqslant x)))).$$

Note that here we think of $\mathbf{C}$ as defining $\{x : \mathbf{C}(x, z_1, \ldots, z_n)\}$, with $z_1, \ldots, z_n$ as parameters.

The fact that we can use parameters in the definition of classes implies that theorems about all classes (like our theorem schema (Theorem 3.5.2) has as one special case, the universal statement about all sets. To see this, let $\mathbf{C}(x, z)$ be the formula $x \in z$. Then, our schema takes the form:

$$\forall z \, ((z \text{ is a non-zero set of ordinals}) \to (\exists x \in z \, \forall y \in z \, (y \geqslant x))) \, ,$$

which is exactly Theorem 3.2.3(5).

What is our point here? Well, a proof "*by transfinite induction on* α" establishes $\forall \alpha \, \psi(\alpha)$ by showing, for each α, that $((\forall \beta < \alpha) \, \psi(\beta)) \to \psi(\alpha)$. Then, the fact that $\forall \alpha \, \psi(\alpha)$ must hold, for otherwise $\exists \alpha \, \neg \psi(\alpha)$, and the least α such that $\neg \psi(\alpha)$ will lead to a contradiction.

A similar result says that one can define a function of α recursively from information about the function below α.

——————— HERE ENDED WINTER 2006 LECTURE 3 ———————

Theorem 3.5.3 (Transfinite recursion for **ON**)**.** *If* $\mathbf{F} : \mathbf{V} \longrightarrow \mathbf{V}$*, then there is a unique* $\mathbf{G} : \mathbf{ON} \longrightarrow \mathbf{V}$ *such that*

$$\forall \alpha \, (\mathbf{G}(\alpha) = \mathbf{F}(\mathbf{G} \restriction \alpha)). \tag{3.1}$$

Proof. To show **uniqueness**, assume that there are functions $\mathbf{G_1}$ and $\mathbf{G_2}$ that both satisfy 3.1. Then, it is possible to prove that $\forall \alpha \, (\mathbf{G_1}(\alpha) = \mathbf{G_2}(\alpha))$ by transfinite induction on α. pedantic details as an exercise?

To show **existence:** Call g a δ*-approximation* of the class $\mathbf{G}$ iff g is a function with domain δ and $\forall \alpha < \delta \, (g(\alpha) = \mathbf{F}(g \restriction \alpha))$. Similarly to the proof of uniqueness, if g is a δ-approximation and g' is a δ'-approximation, then $g \restriction (\delta \cap \delta') = g' \restriction (\delta \cap \delta')$. Next, by transfinite induction on δ, we can show that for each δ, there exists exactly one δ-approximation. Finally, we define $\mathbf{G}(\alpha)$ as $g(\alpha)$, where g is the δ-approximation for some (any) $\delta > \alpha$. $\square_{3.5.3}$

To state Theorem 3.5.3 one has to work a lot harder: For a given formula $\mathbf{F}(x, y)$ (which could also have other free variables), we can explicitly define a formula $\mathbf{G}(x, y)$ (and the explicit manner in which to do this is the content of the proof of Theorem 3.5.3), so that the expression

$$\forall x \, \exists! y \, \mathbf{F}(x, y) \to (\forall \alpha \, \exists! y \, \mathbf{G}(\alpha, y) \wedge \forall \alpha \, \exists x \, \exists y \, (\mathbf{G}(\alpha, y) \wedge \mathbf{F}(x, y) \wedge x = \mathbf{G} \restriction \alpha))$$

is a theorem. **Note:** Here $x = \mathbf{G} \restriction \alpha$ is an abbreviation of the expression "x is a function $\wedge \operatorname{dom}(x) = \alpha \wedge \forall \beta \in \operatorname{dom}(x) \, \mathbf{G}(\beta, x(\beta))$".

Fortunately, it is rare that we need to translate mathematical language with classes to mathematical language without classes! The point is, it is possible, and this is how you do it.

3.6 More ordinal arithmetic

In this section we will take advantage of transfinite recursion to define some further ordinal arithmetic operations. It is possible to define $+$ and $\cdot$ inductively too. To see this, and for details on why these definitions are equivalent, look at Heike's notes. I don't wanna dwell on this point.

Where transfinite recursion is really useful is in the definition of ordinal exponentiation. This is because the purely combinatorial definition is very messy.

Definition 3.6.1. α^β is defined by recursion on β by

1. $\alpha^0 = 1$;
2. $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$;
3. If β is a limit, $\alpha^\beta = \sup\{\alpha^\zeta : \zeta < \beta\}$.

Lemma 3.6.2. *If $\alpha > 0$ and γ is arbitrary, then there exist a unique β and a unique $\rho < \alpha$ such that $\gamma = \alpha \cdot \beta + \rho$.*

Proof. Let β be the greatest ordinal such that $\alpha \cdot \beta \leqslant \gamma$. Details as exercise? $\square_{3.6.2}$

——————— HERE ENDED SPRING 2009 WEEK 3 (3hrs 45 min) ———————

Sometimes the following can be useful:

Theorem 3.6.3 (Cantor's Normal Form Theorem)**.** *Every ordinal $\alpha > 0$ can be represented uniquely in the form*

$$\alpha = \omega^{\beta_1} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n,$$

where $n \geqslant 1$, $\alpha \geqslant \beta_1 > \ldots > \beta_n$, and $k_1, \ldots, k_n$ are nonzero natural numbers.

Proof. We proceed by transfinite induction on α to prove **existence:**.

For $\alpha = 1$, we have $1 = \omega^0 \cdot 1$.

For arbitrary $\alpha > 0$, let β be the largest ordinal such that $\omega^\beta \leqslant \alpha$ (if $\alpha < \omega$, then $\beta = 0$). By Lemma 3.6.2 there is a unique δ and a unique $\rho < \omega^\beta$ such that $\alpha = \omega^\beta \cdot \delta + \rho$. Since $\omega^\beta \leqslant \alpha$, we have $\delta > 0$ and $\rho < \alpha$. Now, δ must necessarily be finite: if δ were infinite, then $\alpha \geqslant \omega^\beta \cdot \delta \geqslant \omega^\beta \cdot \omega = \omega^{\beta+1}$, contradicting the maximality of β. So, let $\beta_1 = \beta$ and $k_1 = \delta$.

If $\rho = 0$, then $\alpha = \omega^{\beta_1} \cdot k_1$ is the normal form, and we can stop.

If $\rho > 0$, then by the induction hypothesis,

$$\rho = \omega^{\beta_2} \cdot k_2 + \ldots + \omega^{\beta_n} \cdot k_n,$$

for some $\beta_2 > \cdots > \beta_n$ and finite $k_2, \ldots, k_n > 0$. Since $\rho < \omega^{\beta_1}$, we have $\omega^{\beta_2} \leqslant \rho\omega^{\beta_1}$, and so $\beta_1 > \beta_2$. So, $\alpha = \omega^{\beta_1} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n$, is expressed in normal form.

Now we show **uniqueness:** First, observe that if $\beta < \gamma$, then $\omega^\beta \cdot k < \omega^\gamma$, for every finite k. This is because $\omega^\beta \cdot k < \omega^\beta \cdot \omega = \omega^{\beta+1} \leqslant \omega^\gamma$. So, if $\alpha = \omega^{\beta_1} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n$ is in normal form, and $\gamma > \beta_1$, then $\alpha < \omega^\gamma$.

We show uniqueness by induction on α.

For $\alpha = 1$, the expansion $1 = \omega^0 \cdot 1$ is clearly unique.

Now, assume for a contradiction that for $\alpha > 0$, $\alpha = \omega^{\beta_1} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n = \omega^{\gamma_1} \cdot l_1 + \ldots + \omega^{\gamma_m} \cdot l_m$. The observation implies that $\beta_1 = \gamma_1$. If we let

$\delta = \omega^{\beta_1} = \omega^{\gamma_1}$, $\rho = \omega^{\beta_2} \cdot k_1 + \ldots + \omega^{\beta_n} \cdot k_n$, and $\sigma = \omega^{\gamma_2} \cdot l_1 + \ldots + \omega^{\gamma_m} \cdot l_m$, we get $\alpha = \delta \cdot k_1 + \rho = \delta \cdot l_1 + \sigma$. Since $\rho < \delta$, and $\sigma < \delta$, Lemma 3.6.2 implies that $k_1 = l_1$ and $\rho = \sigma$. By the induction hypothesis, the normal form for ρ is unique, so the two normal forms we have written for α must also be the same. $\square_{3.6.3}$

maybe a fun application of this would be (weak) Goodstein sequences - sequences that look like they explode hugely, but actually terminate with 0. This stuff is on p126-127 of Hrbacek and Jech. Maybe exercise, maybe a time waster.

Note that it is possible to have $\alpha = \omega^\alpha$. The least ordinal with this property is called ϵ_0. This ordinal is countable (a term that will be explained later!) and rather important for formal arithmetic and recursion theory.

3.7 Proof of AC ⇔ WOP

Ok, now we can return to the proof that the well-ordering principle is equivalent to the axiom of choice.

Proof. **AC⇒WOP:** Assume that the axiom of Choice holds. Let S be any set. We will show that S can be well-ordered. To do this, we find an ordinal α and a one-to-one α-sequence

$$a_0, a_1, \ldots, a_\gamma, \ldots \qquad (\gamma < \alpha)$$

which enumerates S.

Let F be a choice function on the family of all non-empty subsets of S. We use this to construct the desired sequence by transfinite recursion: Let $a_0 = F(S)$. Let $a_\gamma = F(S - \{a_\beta : \beta < \gamma\})$. The construction stops when the elements of S are all used up.

WOP⇒AC: Let $\mathcal{F}$ be any family of sets that are non-empty. By assumptions, each member $S \in \mathcal{F}$ of the family can be well-ordered. For each $S \in \mathcal{F}$, Define $f(S)$ to be the smallest element of S. This satisfies the requirements of a choice function. $\square_{3.1.7}$

———————— HERE ENDED WINTER 2007 LECTURE 3 ————————

Chapter 4

Cardinal Numbers

4.1 Definition and Very Basic Properties of Cardinals

A fundamental property of a set is its size: how big is it? We use cardinal numbers to describe this aspect of a set.

We compare the sizes of sets using injective functions.

Definition 4.1.1.

1. $A \preccurlyeq B$ iff there is a 1-1 function from A into B.
2. $A \approx B$ if there is a 1-1 function from A onto B.
3. $A \prec B$ if $A \preccurlyeq B$ and $B \not\preccurlyeq A$.

It is easy to see that the use here of $\preccurlyeq$ is transitive, and that $\approx$ as used here is an equivalence relation on sets.

One of the most important theorems of the theory of cardinal numbers is the following:

Theorem 4.1.2 (Cantor, Bernstein, (Schröder) Theorem)**.** *If $A \preccurlyeq B$ and $B \preccurlyeq A$ then $A \approx B$.*

Theorem 4.1.2 is a theorem in the partial system $ZF^- - P$.

Proof. Let $f : A \longrightarrow B$ and $g : B \longrightarrow A$ be injective functions. We use these to build a bijection between A and B.

First, let $C_0 = A \setminus \mathrm{rng}(g)$. Inductively define $C_{n+1} = g\text{"}f\text{"}C_n$. (The C_i are progressively smaller sets.)

We define a function $h : A \longrightarrow B$ by

$$h(x) = \begin{cases} f(x) & x \in \bigcup_{n<\omega} C_n, \\ g^{-1}(x) & x \in A \setminus \bigcup_{n<\omega} C_n. \end{cases}$$

This is a well defined function, since if $x \notin C_0$, then $x \in \mathrm{rng}(g)$.

We show that h is **injective:** Let $x \neq x'$ be given. When x and x' are in the same case of the function (i.e. both $x, x' \in \bigcup_{n<\omega} C_n$ or $x, x' \in A \setminus \bigcup_{n<\omega} C_n$), then there is nothing to prove – f is an injective function, on the one hand, and because g is a function, g^{-1} is also always injective.

Assume therefore that $x \in C_m$ for some m and $x' \notin \bigcup_{n<\omega} C_n$. Then in this case, $h(x) = f(x) \in f"C_m$, by definition. On the other hand, $h(x') = g^{-1}(x') \notin f"C_m$, because otherwise, we would have $x' \in g"f"C_m = C_{m+1}$.

We now show that h is **surjective:** Let $y \in B$. Assume that $y \in \bigcup_{n<\omega} f"C_n$. Then $y \in \operatorname{rng}(h)$. Now, assume that $y \notin \bigcup_{n<\omega} f"C_n$. Then $g(y) \notin \bigcup_{n<\omega} C_{n+1}$ and $g(y) \notin C_0$. This means that $h(g(y)) = g^{-1}(g(y)) = y$. $\square_{4.1.2}$

Intuition for the definition of cardinality: One finds the size of a finite set by counting its elements. If a set A can be well ordered, then $A \approx \alpha$ for some ordinal α. The smallest such ordinal α is called the *cardinality* of the set A.

Definition 4.1.3. If A is a set that can be well ordered, then $|A|$ is the smallest ordinal α such that $A \approx \alpha$.

If we write down a statement using $|A|$ (such as $|A| > \alpha$), then we are assuming that A can be well-ordered. If we assume the Axiom of Choice, then every set A can be well-ordered, and hence $|A|$ is defined for every set. Since $A \approx B$ implies $|A| = |B|$ and $|A| \approx A$, assuming the Axiom of Choice, $|A|$ picks a unique representative of each $\approx$-equivalence class.

Regardless of the assumption of the Axiom of Choice, $|\alpha|$ is defined for every ordinal α, and $|\alpha| \leq \alpha$.

Definition 4.1.4. α is a *cardinal* if $\alpha = |\alpha|$.

Lemma 4.1.5. *If* $|\alpha| \leq \beta \leq \alpha$, *then* $|\beta| = |\alpha|$.

Proof. $\beta \subseteq \alpha$, so $\beta \preccurlyeq \alpha$. And, $\alpha \approx |\alpha| \subseteq \beta$, so $\alpha \preccurlyeq \beta$. By Theorem 4.1.2, we get the result. $\square_{4.1.5}$

Lemma 4.1.6. *If* $n \in \omega$, *then*

1. $n \not\approx n + 1$;

2. $\forall \alpha\, (\alpha \approx n \rightarrow \alpha = n)$.

Proof. **(1):** This is proved by induction on n.

(2): This is a corollary of Lemma 4.1.5. $\square_{4.1.6}$

Corollary 4.1.7. ω *is a cardinal, and each* $n \in \omega$ *is a cardinal.*

Definition 4.1.8. We say that a set A is *finite* if $|A| < \omega$. We say that A is *countable* if $|A| \leqslant \omega$. *Infinite* means not finite. *Uncountable* means not countable.

Later, we will show that you really really need the Powerset Axiom for an uncountable set to exist.

4.2 Basic Cardinal Arithmetic

Let us make a notational convention that κ and λ denote cardinals.

We can define arithmetic on cardinals. We'll use circled symbols to distinguish cardinal addition and multiplication from ordinal addition and multiplication.

Definition 4.2.1.

$$\kappa \oplus \lambda = |\kappa \times \{0\} \cup \lambda \times \{1\}|;$$

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

Unlike the addition and multiplication of ordinals, cardinal addition and multiplication are commutative. In addition $|\kappa + \lambda| = |\lambda + \kappa| = \kappa \oplus \lambda$ and $|\kappa \cdot \lambda| = |\lambda \cdot \kappa| = \kappa \otimes \lambda$. So, for example, we have $\omega \oplus 1 = |1 + \omega| = \omega < \omega + 1$. Similarly $\omega \otimes 2 = |2 \cdot \omega| = \omega < \omega \cdot 2$.

Lemma 4.2.2. *For every $n, m \in \omega$, we have $n \oplus m = n + m < \omega$. Similarly, we have $n \otimes m = n \cdot m < \omega$.*

Proof. First, using induction on m, prove that $n + m < \omega$. Then, show $n \cdot m < \omega$ by induction on m. The rest follows from Lemma 4.1.6 (2). $\square_{4.2.2}$

From this point on, we will concentrate on $\oplus$ and $\otimes$ in the context of infinite cardinals.

Lemma 4.2.3. *Every infinite cardinal is a limit ordinal.*

Proof. If $\kappa = \alpha + 1$, then since we have $1 + \alpha = \alpha$, we thus have $\kappa = |\kappa| = |\alpha + 1| = |1 + \alpha| = |\alpha| \leqslant \alpha < \kappa$. A contradiction. $\square_{4.2.3}$

Note that the principle of transfinite induction can be applied to prove results about cardinals, since every class of cardinals is a class of ordinals. The following theorem is an example of this.

Theorem 4.2.4. *If κ is an infinite cardinal, then $\kappa \otimes \kappa = \kappa$.*

Proof. We proceed by transfinite induction on κ. Assume the hypothesis holds for all infinite cardinals smaller than κ, where κ is an infinite cardinal. Then, for $\alpha < \kappa$ we have

$$|\alpha \times \alpha| = |\alpha| \otimes |\alpha| < \kappa.$$

Note that for finite α we apply Lemma 4.2.2.

Now, we define a well-ordering $\ll$ on $\kappa \times \kappa$ in the following manner: $\langle \alpha, \beta \rangle \ll \langle \gamma, \delta \rangle$ iff

$$\max(\alpha, \beta) < \max(\gamma, \delta) \vee$$
$$(\max(\alpha, \beta) = \max(\gamma, \delta) \wedge (\langle \alpha, \beta \rangle \text{ precedes } \langle \gamma, \delta \rangle \text{ lexicographically.})$$

Then, every $\langle \alpha, \beta \rangle$ has no more than $|(\max(\alpha, \beta) + 1) \times (\max(\alpha, \beta) + 1)| < \kappa$ $\ll$-predecessors. (For intuition, see Figure 4.1.) So, type$(\kappa \times \kappa, \ll) \leqslant \kappa$, so $|\kappa \times \kappa| \leqslant \kappa$. Since clearly $|\kappa \times \kappa| \geqslant \kappa$, we have equality. $\square_{4.2.4}$

——————— HERE ENDED SPRING 2009 WEEK 4 (2hrs 15 min) THANKS PETER! ——————

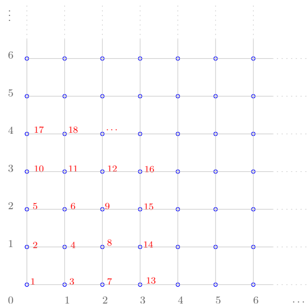

Figure 4.1: A tiny initial portion of the well-ordering $\ll$. This shows that, in the worst case, the predecessors of a given pair are contained in the square defined by that pair.

Corollary 4.2.5. *Let κ and λ be infinite cardinals. Then,*

1. $\kappa \oplus \lambda = \kappa \otimes \lambda = \max(\kappa, \lambda)$;
2. $|\kappa^{<\omega}| \leqslant \omega \otimes \kappa = \kappa$. *(This was defined in Definition 3.4.5.)*

Proof. We prove only **(2)**: We use the proof of Theorem 4.2.4 to define, by induction on n, a 1-1 map $f_n : \kappa^n \longrightarrow \kappa$. This yields a 1-1 map $f : \bigcup_n \kappa^n \longrightarrow \omega \times \kappa$. This gives us $|\kappa^{<\omega}| \leqslant \omega \otimes \kappa = \kappa$. $\square_{4.2.5}$

4.3 The influence of the Powerset Axiom

We begin the discussion of Axiom 7

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

with the following definition:

Definition 4.3.1. The set

$$\mathscr{P}(x) = \{z : z \subseteq x\}$$

is called the *power set* of the set x.

The existence of a power set is guaranteed by the Power Set Axiom and the Restricted Comprehension Schema. The operation $\mathscr{P}()$ allows us to build sets of greater cardinalities.

Theorem 4.3.2 ((ZF$^-$) Cantor)**.** $x \prec \mathscr{P}(x)$.

Proof. This is a proof in ZF^-. Let $f : x \longrightarrow \mathscr{P}(x)$. We will show that f cannot be surjective. Let

$$u = \{y \in x : y \notin f(y)\} \in \mathscr{P}(x).$$

Then, there is no $y \in x$ such that $f(y) = u$ – otherwise, if $f(y) = u$, then we would have $y \in u \iff y \notin f(y) = u$, which would be a contradiction. $\square_{4.3.2}$

With the help of the Axiom of Choice, one can deduce from Theorem 4.3.2 that there exists a cardinal $> \omega$, in particular, $|\mathscr{P}(\omega)|$.

——————— HERE ENDED WINTER 2006 LECTURE 4 ———————

One does not actually need the Axiom of Choice to reach this conclusion:

Theorem 4.3.3 ((ZF$^-$) Hartogs, 1906)**.**

$$\forall\alpha\,\exists\kappa\,(\kappa > \alpha \textit{ and } \kappa \textit{ is a cardinal}).$$

Proof. This is a proof in ZF^-. Let $\alpha \geqslant \omega$. Let $W = \{R \in \mathscr{P}(\alpha \times \alpha) : R \text{ well orders } \alpha\}$. Let $S = \{\text{type}(\langle\alpha, R\rangle) : R \in W\}$. The set S exists by the Replacement Axiom, and is a set of ordinals, so has a supremum. Then, $\sup(S) \notin S$, since $\forall\beta \in S\,(\beta + 1 \in S)$. Thus it is clear that $\sup(S)$ is an ordinal $> \alpha$.

Now we show that $\sup(S)$ is a cardinal: If $\sup(S)$ were not a cardinal, then there would be a $\beta < \sup(S)$ such that $\beta \approx \sup(S)$. Let such a β be minimally chosen. Then, β is a cardinal. Since $\beta < \sup(S)$, there is a well-ordering R of α, such that $\beta \leqslant \text{type}(\alpha, R)$. Thus, we have $|\beta| \leqslant |\alpha|$.

Let $f : \beta \longrightarrow \sup(S)$ be a bijection, and define $R_\beta \subseteq \beta \times \beta$ by $\gamma R_\beta \gamma' \iff f(\gamma) <_{\sup(S)} f(\gamma')$. Then β can be well ordered using type(β, R_β). (And naturally, α can also be well-ordered, using a similar argument) This contradicts the fact that $\sup(S) \notin S$ and the definition of $\sup(S)$. $\square_{4.3.3}$

Definition 4.3.4 ((ZF$^-$))**.** Define α^+ to be the smallest cardinal $> \alpha$.

κ is a *successor cardinal* iff $\kappa = \lambda^+$ for some cardinal λ.

κ is a *limit cardinal* iff κ is not a successor cardinal and $\kappa > \omega$.

Definition 4.3.5. $\aleph_\alpha = \omega_\alpha$ is defined by transfinite recursion on α by:

1. $\aleph_0 = \omega_0 = \omega$;
2. $\aleph_{\alpha+1} = \omega_{\alpha+1} = (\aleph_\alpha)^+$;
3. For γ a limit, $\aleph_\gamma = \bigcup\{\aleph_\alpha : \alpha < \gamma\}$.

That funny letter in the previous definition is aleph, the first letter of the Hebrew alphabet.

Lemma 4.3.6.

1. *Every $\aleph_\alpha$ is a cardinal.*
2. *Every infinite cardinal is equal to $\aleph_\alpha$ for some α.*
3. *$\alpha < \beta \rightarrow \aleph_\alpha < \aleph_\beta$.*
4. *$\aleph_\alpha$ is a limit cardinal iff α is a limit ordinal. $\aleph_\alpha$ is a successor cardinal iff α is a successor ordinal.*

Proof. **(1) and (3)** are both proved by induction on α. The successor steps should be clear. For the limit step, note that every limit of cardinals is itself a cardinal. We prove this in general: Let $\kappa = \sup\{\kappa_i : i \in I\}$ and let κ_i be pairwise different cardinals. With perhaps some reordering, let the κ_i be in a strictly $\prec$-increasing sequence. Then, I is an ordinal number, say $I = \beta$. Thus, $\{\kappa_i : i \in I\} \subset ON$ and so is well-ordered and so we do not have to use the

Axiom of Choice here. So, with these assumptions, $\kappa_i \prec \kappa_j$ for $i < j < \beta$. So, by a previous Lemma, κ is an ordinal because it is a supremum of a set of ordinals. By the definition of supremum, κ is the smallest ordinal larger than all the κ_i. Thus, every $\kappa' < \kappa$ (in the ordering of ordinals) is $\leqslant \kappa_i$ for some $i \in \beta$. Thus $\kappa' \leqslant \kappa_i \prec \kappa_{i+1} \leqslant \kappa$. Therefore, $\kappa' \not\approx \kappa$. Therefore, κ is a cardinal number.

(2) is proved by transfinite induction along the ordinals, and is a direct consequence of the definition of the $\aleph$s.

(4): The statement also holds for the third case in our trichotomy (limit, successor, 0): $\aleph_0$ and 0 are both the only members of the third case. Inductively, we get the truth of the statement for the successor case from $(\aleph_\alpha)^+ = \aleph_{\alpha+1}$. Similar reasoning to that in part (3) of this proof yields the limit case. $\square_{4.3.6}$

Many important facts about cardinals do, however, heavily rely on the Axiom of Choice.

Lemma 4.3.7 ((ZFC$^-$))**.** *If there exists a function f from X onto Y, then $|Y| \leqslant |X|$.*

Proof. Let R be a well-ordering of X (as guaranteed by the Axiom of Choice). Define $g : Y \longrightarrow X$ so that $g(y)$ is the R-least element of $f^{-1}(\{y\})$. Then, g is a 1-1 function, so $Y \preccurlyeq X$. $\square_{4.3.7}$

Note: As in the Cantor's Theorem 4.3.2, one can prove, even without the Axiom of Choice, that there exists a mapping from $\mathscr{P}(\omega)$ onto $\aleph_1$, but one cannot prove the existence of a 1-1 function from $\aleph_1$ into $\mathscr{P}(\omega)$.

Lemma 4.3.8 ((ZFC$^-$))**.** *If $\kappa \geqslant \omega$ and $|X_\alpha| \leqslant \kappa$ for all $\alpha < \kappa$, then we have $|\bigcup\{X_\alpha : \alpha < \kappa\}| \leqslant \kappa$.*

Proof. Let $\mathcal{F} = \{\{f : \text{the function } f : X_\alpha \longrightarrow \kappa \text{ is injective.}\} : \alpha < \kappa\}$. By the assumption of the Axiom of Choice, we can well-order $\mathcal{F}$: Let $h = \{\langle \alpha, \{f : \text{the function } f : X_\alpha \longrightarrow \kappa \text{ is injective.}\}\rangle : \alpha < \kappa\}$, i.e., h well-orders $\mathcal{F}$ with ordertype κ.

In addition, from the assumption of the Axiom of Choice, we have a choice function for $\mathcal{F}$. Taking into account h, these choices can be well-ordered. Thus, we have an injective function $g : \kappa \longrightarrow \bigcup \mathcal{F}$. The function g is defined so that for $\alpha < \kappa$, we have $g(\alpha) : X_\alpha \longrightarrow \kappa$.

Then, we have the following injection: $g' : \bigcup_{\alpha<\kappa} X_\alpha \longrightarrow \kappa \times \kappa$, defined by $g'(x) = \langle \alpha, g(\alpha)(x)\rangle$, where $\alpha = \min\{\beta < \kappa : x \in X_\beta\}$. .

The fact that $\kappa \otimes \kappa = \kappa$ gives us the final result. $\square_{4.3.8}$

The use of the Axiom of Choice in the preceding Lemma is vital. It is possible to show (Azriel Levy did this) that without the Axiom of Choice, it is consistent with ZF that both $\mathscr{P}(\omega)$ and ω_1 are countable unions of countable sets.

A very important modification of Lemma 4.3.8 is the so-called Downward Löwenheim-Skolem Theorem, an important theorem from model theory that is used quite often in set theory. To state this, we need a definition:

Definition 4.3.9. An *n-ary function* on A is an $f : A^n \longrightarrow A$ if $n > 0$, or an element of A if $n = 0$. If $B \subset A$, B is closed under f iff $f\text{''}B^n \subset B$ (or $f \in B$ when $n = 0$). A *finitary function* is an n-ary function for some n. If $\mathscr{S}$ is a set of finitary functions and $B \subset A$, the *closure* of B under $\mathscr{S}$ is the least $C \subset A$ such that $B \subset C$ and C is closed under all the functions in $\mathscr{S}$. Note that there is a least C, namely $\bigcap\{D : B \subset D \subset A \wedge D \text{ is closed under } \mathscr{S}\}$.

Theorem 4.3.10 (Downward Löwenheim-Skolem Theorem)**.** *Let κ be an infinite cardinal. Suppose $B \subset A$, $|B| \leqslant \kappa$, and $\mathscr{S}$ is a set of $\leqslant \kappa$ finitary functions on A. Then the closure of B under $\mathscr{S}$ has cardinality $\leqslant \kappa$.*

Note that this is a purely combinatorial version of this theorem. GIVE THE MODEL THEORETIC INTERPRETATION OF THIS THEOREM!!

Proof. For $f \in \mathscr{S}$ and $D \subseteq A$ we define

$$f * D = \begin{cases} f''D, & \text{if } f \ n\text{-ary}, n > 0, \\ \{f\}, & \text{if } f \ 0\text{-ary}. \end{cases}$$

Notice that if $|D| \leqslant \kappa$, then we have $|f * D| \leqslant \kappa$, since $|\kappa^n| = \kappa$.
We define inductively on $n < \omega$ the sets C_n by

$$\begin{aligned} C_0 &= B, \\ C_{n+1} &= C_n \cup \bigcup\{f * C_n : f \in \mathscr{S}\}. \end{aligned}$$

Inductively on n, and by Lemma 4.3.8 we can show that $|C_n| \leqslant \kappa$.
Clearly, $C_\omega := \bigcup_{n\in\omega} C_n$ is closed under $\mathscr{S}$.
Again, from Lemma 4.3.8, we have $|C_\omega| \leqslant \kappa$. $\square_{4.3.10}$

———————— HERE ENDED WINTER 2007 LECTURE 4 ————————

4.4 Cardinal Exponentiation

Definition 4.4.1 (ZF$^-$)**.** $A^B = {}^BA = \{f : f \text{ is a function } \wedge \operatorname{dom}(f) = B \wedge \operatorname{rng}(f) \subseteq A\}$.

This set exists, because, for example, ${}^BA \subseteq \mathscr{P}(A \times B)$. Thus, ${}^BA$ exists by the Powerset and Comprehension Axioms.

Definition 4.4.2 (ZFC$^-$)**.** $\kappa^\lambda = |{}^\lambda\kappa|$.

Both notations A^B and ${}^BA$ appear in the literature. In this lecture, to avoid misunderstandings, κ^λ when we are talking about cardinals, and ${}^\lambda\kappa$ when we are talking about functions.

Lemma 4.4.3. *If $\lambda \geqslant \aleph_0$ and $2 \leqslant \kappa \leqslant \lambda$, then ${}^\lambda\kappa \approx {}^\lambda 2 \approx \mathscr{P}(\lambda)$.*

Proof. The fact that ${}^\lambda 2 \approx \mathscr{P}(\lambda)$ follows from the identification of sets with their characteristic functions. Further, we have ${}^\lambda 2 \preccurlyeq {}^\lambda\kappa \preccurlyeq {}^\lambda\lambda \preccurlyeq \mathscr{P}(\lambda \times \lambda) \approx \mathscr{P}(\lambda) \approx 2^\lambda$. $\square_{4.4.3}$

So, cardinal exponentiation is not the same as ordinal exponentiation. For example with ordinals, 2^ω is ω, but $2^{\aleph_0} = |\mathscr{P}(\omega)| > \aleph_0$. In future, if ordinal exponentiation is meant, I will explicitly say so. So exponent notation will mean cardinal exponentiation by default.

The same rules from normal arithmetic apply here too:

Lemma 4.4.4 (ZFC$^-$). *If κ, λ, and μ are cardinals, then $\kappa^{\lambda\oplus\mu} = \kappa^\lambda \otimes \kappa^\mu$ and $(\kappa^\lambda)^\mu = \kappa^{\lambda\otimes\mu}$.*

Proof. Without assuming the Axiom of Choice, it is possible to show that if the sets B and C *are disjoint* then we have ${}^{(B\cup C)}A \approx {}^BA \times {}^CA$ and ${}^C({}^BA) \approx {}^{C\times B}A$. details left as an exercise. $\square_{4.4.4}$

Definition 4.4.5 (AC).

1. *CH* (the Continuum Hypothesis) is the statement $2^{\aleph_0} \approx \aleph_1$.
2. *GCH* (the Generalized Continuum Hypothesis) is the statement $\forall\alpha\,(2^{\aleph_\alpha} = \aleph_{\alpha+1})$.

Cantor showed that $2^{\aleph_\alpha} \geqslant \aleph_{\alpha+1}$ (Theorem 4.3.2), but couldn't do any more than that. This problem drove set theory for a good portion of the first half of the 20th century. Gödel showed in 1938 that if ZFC is consistent, then so is ZFC + CH. (Maybe we will get to this this semester.....). But! Cohen showed in 1963 that if ZFC is consistent, then so is ZFC + ¬CH. So, the continuum hypothesis is independent of ZFC. The latter fact will be proved next semester.

4.5 Cofinalities and different kinds of Cardinals

Now, what exactly is GCH good for? Well, for one, κ^λ becomes easy to compute. We show this, but first we need some definitions.

Definition 4.5.1. If $f : \alpha \longrightarrow \beta$, f maps α *cofinally* iff $\mathrm{rng}(f)$ is unbounded in β. The *cofinality* of β, written $\mathrm{cf}(\beta)$, is the least α such that there is a map from α cofinally into β.

So, $\mathrm{cf}(\beta) \leqslant \beta$ and, if β is a successor ordinal, then $\mathrm{cf}(\beta) = 1$.

——————— HERE ENDED SPRING 2009 WEEK 5 (2 hrs 45 min) ———————

Lemma 4.5.2. *There is a cofinal map $f : \mathrm{cf}(\beta) \longrightarrow \beta$ which is strictly increasing (i.e. $\zeta < \nu \to f(\zeta) < f(\nu)$).*

Proof. Let $g : \mathrm{cf}(\beta) \longrightarrow \beta$ be any cofinal map. We define f recursively by

$$f(\mu) = \max(g(\mu), \sup\{f(\zeta) + 1 : \zeta < \mu\}) < \beta.$$

$\square_{4.5.2}$

Lemma 4.5.3. *If α is a limit ordinal, and $f : \alpha \longrightarrow \beta$ is a strictly increasing cofinal function, then $\mathrm{cf}(\alpha) = \mathrm{cf}(\beta)$.*

Proof. The fact that $\mathrm{cf}(\alpha) \leqslant \mathrm{cf}(\beta)$ follows by composing a cofinal map from $\mathrm{cf}(\alpha)$ into α with f.

We show $\mathrm{cf}(\alpha) \geqslant \mathrm{cf}(\beta)$: Let $g : \mathrm{cf}(\beta) \longrightarrow \beta$ be a cofinal mapping. Put $h(\zeta) = \min\{\eta : f(\eta) > g(\zeta)\}$. Then, h is a cofinal function because f is strictly increasing and cofinal. Thus, $h \circ g : \mathrm{cf}(\beta) \longrightarrow \alpha$ gives the desired inequality. $\square_{4.5.3}$

Corollary 4.5.4. $\mathrm{cf}(\mathrm{cf}(\beta)) = \mathrm{cf}(\beta)$.

Proof. We use Lemma 4.5.3 on a strictly increasing function $f : \mathrm{cf}(\beta) \longrightarrow \beta$, whose existence is guaranteed by Lemma 4.5.2. $\square_{4.5.4}$

Definition 4.5.5. An ordinal β is *regular* iff β is a limit ordinal and $\mathrm{cf}(\beta) = \beta$.

So, by Corollary 4.5.4, $\mathrm{cf}(\beta)$ is regular for every limit β.

Lemma 4.5.6. *If an ordinal β is regular, then it is a cardinal.*

Proof. We prove this by contradiction. Assume that there is $\alpha < \beta$ such that there exists an onto function $f : \alpha \longrightarrow \beta$. Then, we would have $\mathrm{cf}(\beta) \leqslant \alpha < \beta$. This would imply that β is not regular, a contradiction. $\square_{4.5.6}$

——————— HERE ENDED WINTER 2006 LECTURE 5 ———————

Definition 4.5.7. An infinite cardinal κ is *regular* if $\mathrm{cf}(\kappa) = \kappa$. It is *singular* if $\mathrm{cf}(\kappa) < \kappa$.

Lemma 4.5.8. *ω and all infinite* $\mathrm{cf}(\beta)$ *are regular.*

Lemma 4.5.9 (ZFC^-)**.** *For every cardinal κ, κ^+ is regular.*

Proof. We prove this by contradiction. Assume that there is a cofinal mapping $f : \alpha \longrightarrow \kappa^+$, where $\alpha < \kappa$. Then we have $\kappa^+ = \bigcup\{f(\zeta) : \zeta < \alpha\}$. But then, the union of $\leqslant \kappa$ sets of cardinality $\leqslant \kappa$ is, by Lemma 4.3.8 also of cardinality $\leqslant \kappa$ (and in particular $\neq \kappa^+$). Contradiction. $\square_{4.5.9}$

Without the assumption of the Axiom of Choice, it is consistent that $\mathrm{cf}(\aleph_1) = \omega$. For a long time, it was not known if one can prove in ZF that there exists a cardinal of cofinality $> \omega$. This was finally done by M. Gitik in 1980. He built a model of set theory without Choice containing a singular cardinal of uncountable cofinality.

Limit cardinals are often not regular. For example $\mathrm{cf}(\aleph_\omega) = \omega$. More generally, we have the following:

Lemma 4.5.10. *If α is a limit ordinal, then* $\mathrm{cf}(\aleph_\alpha) = \mathrm{cf}(\alpha)$.

Proof. This results from Lemma 4.5.3. $\square_{4.5.10}$

So, the question is, are there regular limit cardinals $\aleph_\alpha$? If $\aleph_\alpha$ is a regular limit cardinal, then $\aleph_\alpha = \alpha$. But, the condition $\aleph_\alpha = \alpha$ is not enough to guarantee that $\aleph_\alpha$ is a regular limit cardinal. To see this, define $\sigma_0 = \aleph_0$, $\sigma_{n+1} = \aleph_{\sigma_n}$. Let $\alpha = \{\sigma_n : n < \omega\}$. Then, α is the first cardinal satisfying $\aleph_\alpha = \alpha$, but $\mathrm{cf}(\alpha) = \omega$.

Regular limit cardinals, despite the problem stated above, play a very vital role. They are among the so-called "large cardinals". We define:

Definition 4.5.11.

1. κ is *weakly inaccessible* iff κ is a regular limit cardinal.

2. (AC) κ is *strongly inaccessible* iff $\kappa > \omega$, κ is regular, and

$$\forall \lambda < \kappa\, (2^\lambda < \kappa).$$

So, a strongly inaccessible cardinal is a weakly inaccessible cardinal. Under the assumption of GCH, the two notions coincide. It is consistent that 2^ω is weakly inaccessible. It is also consistent that it is larger than the first weakly inaccessible cardinal. One cannot prove in ZFC that weakly inaccessible cardinals exist.

By modifying an argument of Cantor, we have $(\omega_\omega)^\omega > \omega_\omega$. More generally:

Lemma 4.5.12 (ZFC$^-$ König's Lemma 1905)**.** *If κ is an infinite cardinal, and $\mathrm{cf}(\kappa) \leqslant \lambda$, then $\kappa^\lambda > \kappa$.*

Proof. Let $f : \lambda \longrightarrow \kappa$ be a cofinal mapping. Let $G : \kappa \longrightarrow {}^\lambda\kappa$. We show that G cannot be onto: Define $h : \lambda \longrightarrow \kappa$ so that $h(\alpha)$ is the smallest element of the set $\kappa \setminus \{G(\mu)(\alpha) : \mu < f(\alpha)\}$. Then, $h \notin \mathrm{rng}(G)$. For if otherwise, $h = G(\mu)$ for some μ. Take α such that $f(\alpha) > \mu$ (this is possible because f is a cofinal mapping). Then $G(\mu)(\alpha) \neq h(\alpha)$. Thus $G(\mu) \neq h$, a contradiction. $\square_{4.5.12}$

Corollary 4.5.13 (AC)**.** *If $\lambda \geqslant \omega$, then $\mathrm{cf}(2^\lambda) > \lambda$.*

Proof. By the properties of cardinal arithmetic, we have $(2^\lambda)^\lambda = 2^{\lambda\otimes\lambda} = 2^\lambda$. Now, compare this to Lemma 4.5.12 with $\kappa = 2^\lambda$. $\square_{4.5.13}$

Lemma 4.5.14 (ZFC$^-$ +GCH)**.** *Assume that $\kappa, \lambda \geqslant 2$ and at least one of them is infinite. Then,*

1. $\kappa \leqslant \lambda \rightarrow \kappa^\lambda = \lambda^+$;
2. $\kappa > \lambda \geqslant \mathrm{cf}(\kappa) \rightarrow \kappa^\lambda = \kappa^+$;
3. $\lambda < \mathrm{cf}(\kappa) \rightarrow \kappa^\lambda = \kappa$.

Proof.

1. This part results from Lemma 4.4.3.
2. By Lemma 4.5.12 we have $\kappa^\lambda > \kappa$. On the other hand, we have $\kappa^\lambda \leqslant \kappa^\kappa = 2^\kappa = \kappa^+$.
3. IF $\lambda < \mathrm{cf}(\kappa)$, then ${}^\lambda\kappa = \bigcup\{{}^\lambda\alpha : \alpha < \kappa\}$, but $|{}^\lambda\alpha| \leqslant \max(\alpha, \lambda)^+ \leqslant \kappa$.

$\square_{4.5.14}$

To finish this section off, we give a couple of useful definitions.

Definition 4.5.15 (AC)**.** $A^{<\beta} = \{{}^\alpha A : \alpha < \beta\}$.

Note: If $\kappa \geqslant \omega$, then $|\kappa^{<\omega}| = \kappa$ and

$$|\kappa^{<\lambda}| = \sup\{\kappa^\theta : \theta < \lambda \wedge \theta \text{ is a cardinal.}\}.$$

Definition 4.5.16 (AC)**.** By transfinite recursion over the ordinal numbers, we define the $\beth$-operation (beth operation - beth is the second Hebrew letter.):

1. $\beth_0 = \omega$;
2. $\beth_{\alpha+1} = 2^{\beth_\alpha}$;
3. for limit γ we have $\beth_\gamma = \sup\{\beth_\alpha : \alpha < \gamma\}$.

We will see the $\beth$-operation again, but not too soon - only once we get to Lemma 5.1.13. maybe some exercises covering the last two definitions would be good. They flap in the wind aimlessly for now.

Chapter 5

The Axiom of Regularity

In this chapter, we will work in ZF$^-$.

We will define the class **WF** of well-founded sets. Intuitively, **WF** is the class of all sets that can be defined from $\emptyset$ with the help of various set-theoretic operations. Then, we will prove some theorems that show that all of mathematics takes place within **WF**. This will lead us to the Axiom of Regularity, which in effect says that $\mathbf{WF} = \mathbf{V}$, i.e. that our domain of discourse of sets is restricted to the well-founded ones.

5.1 Properties of well-founded sets

Definition 5.1.1. By transfinite recursion, we define the sets R_α, for $\alpha \in \mathbf{ON}$ by:

1. $R_0 = \emptyset$;
2. $R_{\alpha+1} = \mathscr{P}(R_\alpha)$;
3. $R_\alpha = \bigcup\{R_\zeta : \zeta < \alpha\}$, for limit α.

The R_αs are sometimes called the *von Neumann Hierarchy*.

Definition 5.1.2. $\mathbf{WF} = \bigcup\{R_\alpha : \alpha \in \mathbf{ON}\}$.

So, in other words, well-founded sets are sets which occur in some R_α.

Lemma 5.1.3. *For every α, we have*

1. *R_α is transitive, and*
2. *$\forall\zeta < \alpha\,(R_\zeta \subseteq R_\alpha)$.*

Proof. We proceed by transfinite induction on α.

The case $\alpha = 0$ is trivial.

Let us assume that the lemma is true for all $\beta < \alpha$.

Assume α is a limit ordinal. Then, point 2 follows straight from the definition. Point 1 follows from the fact that the union of transitive sets is transitive.

Now, assume $\alpha = \beta + 1$. Let $x \in R_\alpha = \mathscr{P}(R_\beta)$. Let $y \in x$. Then, we have that $y \in R_\beta$, because $y \in x \subseteq R_\beta$. Because R_β is transitive, $y \subseteq R_\beta$. Therefore, $y \in \mathscr{P}(R_\beta) = R_\alpha$. Since this is true of all elements $y \in x$, x is a subset of R_α. We get point 2 from the transitivity of R_β. $\square_{5.1.3}$

From this, we see that the R_α increase in size along with α. We also note that if $x \in \mathbf{WF}$, then the smallest α such that $x \in R_\alpha$ must be a successor. This is clear by part 3 of Definition 5.1.1.

——————— HERE ENDED WINTER 2007 LECTURE 5 ———————

Definition 5.1.4. If $x \in \mathbf{WF}$, then $\text{rank}(x)$ is the smallest β such that $x \in R_{\beta+1}$.

So, in other terms, if $\text{rank}(x) = \beta$, then $x \subseteq R_\beta$, $x \notin R_\beta$, and $x \in R_\alpha$ for all $\alpha > \beta$.

Lemma 5.1.5. *For any α, $R_\alpha = \{x \in \boldsymbol{WF} : \text{rank}(x) < \alpha\}$.*

Proof. For $x \in \mathbf{WF}$, $\text{rank}(x) < \alpha$ iff $\exists \beta < \alpha\,(x \in R_{\beta+1})$ iff $x \in R_\alpha$. $\square_{5.1.5}$

The following lemma is quite useful when calculating the the rank of a set:

Lemma 5.1.6. *If $y \in \boldsymbol{WF}$, then*

1. $\forall x \in y\,(x \in \boldsymbol{WF} \wedge \text{rank}(x) < \text{rank}(y))$, *and*

2. $\text{rank}(y) = \sup\{\text{rank}(x) + 1 : x \in y\}$.

Proof. We prove 1.:
Let $\text{rank}(y) = \alpha$. Then, $y \in R_{\alpha+1} = \mathscr{P}(R_\alpha)$. If $x \in y$, then $x \in R_\alpha$. Thus, $\text{rank}(x) < \text{rank}(y)$.

For 2.: Let $\alpha = \sup\{\text{rank}(x) + 1 : x \in y\}$. By point 1, $\alpha \leqslant \text{rank}(y)$. Furthermore, every $x \in y$ has $\text{rank}(x) < \alpha$. Therefore, $y \subseteq R_\alpha$, and so $y \in R_{\alpha+1}$. Thus, $\text{rank}(y) \leqslant \alpha$. $\square_{5.1.6}$

Lemma 5.1.6, point 1, says that **WF** is transitive, and that we can think of the elements $y \in \mathbf{WF}$ as having been "constructed" by transfinite recursion from well-founded sets of smaller rank. Thus, **WF** excludes sets that are built up from themselves. Formally, there is no $x \in \mathbf{WF}$ such that $x \in x$, since we would then have that $\text{rank}(x) < \text{rank}(x)$. Similarly **WF** rules out cycles of the type $x \in y \wedge y \in x$, since we would in such a case have $\text{rank}(x) < \text{rank}(y) < \text{rank}(x)$.

Further, each ordinal is in **WF**, and its rank is itself:

Lemma 5.1.7.

1. $\forall \alpha \in \boldsymbol{ON}(\alpha \in \boldsymbol{WF} \wedge \text{rank}(\alpha) = \alpha)$;

2. $\forall \alpha \in \boldsymbol{ON}(R_\alpha \cap \boldsymbol{ON} = \alpha)$.

Proof.

1. We prove this via transfinite induction on α. Assume that the hypothesis holds for all $\beta < \alpha$. Then, for $\beta < \alpha$, we have $\beta \in R_{\beta+1} \subseteq R_\alpha$, and so $\alpha \subseteq R_\alpha$. Thus, $\text{rank}(\alpha) \leqslant \alpha$. By Lemma 5.1.6, 2, $\text{rank}(\alpha) = \sup\{\beta + 1 : \beta < \alpha\} = \alpha$. And so, the hypothesis holds for α as well.

2. This is a direct consequence of Lemma 5.1.5 and part 1 of this lemma.

$\square_{5.1.7}$

The class **WF** contains not only the ordinals, but also other sets that arise through standard mathematical constructions, since **WF** is closed under such constructions.

Lemma 5.1.8.

1. *If $x \in \boldsymbol{WF}$, then $\bigcup x$, $\mathscr{P}(x)$ and $\{x\} \in \boldsymbol{WF}$. The ranks of these sets are smaller than* $\operatorname{rank}(x) + \omega$.

2. *If $x, y \in \boldsymbol{WF}$, the $x \cap y$, $x \cup y$, $x \times y$, $\{x, y\}$, $\langle x, y\rangle$, and ${}^y x$ are also in $\boldsymbol{WF}$. The ranks of these sets are smaller than* $\max\{\operatorname{rank}(x), \operatorname{rank}(y)\} + \omega$.

Proof.

1. Let $\operatorname{rank}(x) = \alpha$. Then, $x \subseteq R_\alpha$. Thus, $\mathscr{P}(x) \subseteq \mathscr{P}(R_\alpha) = R_{\alpha+1}$. Similarly, $\{x\} \subseteq R_\alpha$ and $\bigcup x \subseteq R_\alpha$. Hence, $\bigcup x \in R_{\alpha+1}$.

2. Let $\alpha = \max\{\operatorname{rank}(x), \operatorname{rank}(y)\}$. Just as in Case 1, we can calculate: $\{x, y\} \in R_{\alpha+2}$, $\langle x, y\rangle = \{x, \{x, y\}\} \in R_{\alpha+3}$. Any ordered pair of elements of $x \cup y$ is in $R_{\alpha+2}$, so ${}^y x \subset R_{\alpha+3}$, so ${}^y x \in R_{\alpha+4}$.

$\square_{5.1.8}$

Familiar mathematical objects are also in the well-founded hierarchy.

Lemma 5.1.9. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, *and* $\mathbb{C}$ *are elements of* $R_{\omega+\omega}$.

Any of the standard definitions for the above sets is good. For example, take $\mathbb{Z} = (\omega \times \omega)/\equiv$, where the relation $\equiv$ is defined so that $\langle n, m\rangle$ represents $m - n$. Similarly, we can define $\mathbb{Q} = (\omega \times (\omega \setminus \{0\})/\equiv$, where $\langle x, y\rangle/\equiv$ represents the fraction x/y. Finally, let

$$\mathbb{R} = \{X \in \mathscr{P}(\mathbb{Q}) : X \neq \emptyset \wedge X \neq \mathbb{Q} \wedge \forall x \in X\, \forall y \in \mathbb{Q}\, (y < x \rightarrow y \in X)\}.$$

In other words, let $\mathbb{R}$ be the set of the "left" parts of Dedekind cuts. Further, let $\mathbb{C} = \mathbb{R} \times \mathbb{R}$.

Proof. The proof follows from Lemma 5.1.8 and the definitions of these sets. $\square_{5.1.9}$

Lemma 5.1.10. $\forall x\, (x \in \boldsymbol{WF} \iff x \subseteq \boldsymbol{WF})$.

Proof. The implication $(x \in \mathbf{WF} \rightarrow x \subseteq \mathbf{WF})$ is just restates the transitivity of **WF** (Lemma 5.1.6). For the opposite implication: if $x \subseteq \mathbf{WF}$, then let $\alpha = \sup\{\operatorname{rank}(y) + 1 : y \in x\}$. Then, by the definition, $x \subseteq R_\alpha$, and hence, $x \in R_{\alpha+1}$. $\square_{5.1.10}$

Note that it is possible to get the closure properties as given by Lemma 5.1.8 directly from Lemma 5.1.10. However, Lemma 5.1.10 is much stronger than just that. For any R_γ, γ a limit, satisfies the same closure properties, but any class satisfying Lemma 5.1.10 must contain **WF**. maybe a little exercise? see Kunen ex 3 p 107.

We now concentrate on the cardinalities of the sets R_α:

Lemma 5.1.11. $\forall n \in \omega\, (|R_n| < \omega)$.

Proof. The proof proceeds by induction on n. $\square_{5.1.11}$

Lemma 5.1.12. $|R_\omega| = \omega$.

Proof. Since $\omega \subset R_\omega$, it suffices to show that R_ω is countable. If we assume the Axiom of Choice, then we get the result from Lemma 5.1.11. If we wish to avoid the Axiom of Choice, then notice that it is possible to define a well-ordering on R_n by induction on $n < \omega$. For example, if we already have a well-ordering on R_n, we can identify R_{n+1} with ${}^{R_n}2$, and then order it lexicographically. $\square_{5.1.12}$

The powers of the R_α increase exponentially: $|R_\omega| = \aleph_0$; $|R_{\omega+1}| = 2^{\aleph_0}$; $|R_{\omega+2}| = 2^{2^{\aleph_0}}$;...etc. More generally:

Lemma 5.1.13 (AC)**.** $|R_{\omega+\alpha}| = \beth_\alpha$.

Proof. The proof proceeds by induction. $\square_{5.1.13}$

All reasonable mathematics can take place in **WF**. That is, one can find the usual mathematical structures in **WF**, or at least, isomorphic copies of them. To illustrate, we have the next lemma:

Lemma 5.1.14 (AC)**.**

1. *Every group is isomorphic with a group in* ***WF***.
2. *Every topological space is homeomorphic with a topological space in* ***WF***.

Proof. Formally, a group is an ordered pair $\langle G, \cdot \rangle$, where $G \times G \longrightarrow G$. By Lemmas 5.1.8 and 5.1.10 we have that $\langle G, \cdot \rangle \in \mathbf{WF}$ iff $G \in \mathbf{WF}$ iff $G \subseteq \mathbf{WF}$. For our group $\langle G, \cdot \rangle$, by the assumption of the Axiom of Choice, there is an ordinal α such that $|G| = \alpha$. Let f, thus, be a bijective mapping from α onto G. Then we can define an operation $\circ$ on α in the following manner: $\zeta \circ \eta = f^{-1}(f(\zeta) \cdot f(\eta))$. This means that f is an isomorphism from $\langle \alpha, \circ \rangle$ onto $\langle G, \cdot \rangle$.

The proof of 2. is similar. $\square_{5.1.14}$

So, **WF** contains concrete mathematical objects like $\mathbb{Z}$ and $\mathbb{R}$, and identical copies of the various abstract objects like groups, topological spaces, and so on.

——————— HERE ENDED WINTER 2006 LECTURE 6 ———————

5.2 Well-founded relations

The idea of a well-founded relation is a generalization of a well-order. This will be very important in constructions of models of set theory.

Despite the fact that the definition of **WF** uses the Powerset Axiom in an important way, many results in this section about well-founded relations will be done in the theory $\mathrm{ZF}^- - \mathrm{P}$. This will also be an important assumption in later constructions of models of set theory.

Definition 5.2.1 ($\mathrm{ZF}^- - \mathrm{P}$)**.** A relation R is *wf* (well-founded) on a set A iff

$$(\forall X \subseteq A)(X \neq \emptyset \rightarrow (\exists y \in X)(\neg(\exists z \in X)(zRy))).$$

An element y as in the formula above is called the *R-minimal element* of X.

In other words, R is wf on A if and only if every non-empty subset of A has a R-minimal element.

Lemma 5.2.2 (ZF$^-$)**.** *If* $A \in \boldsymbol{WF}$*, then* $\in$ *is a well founded relation on* A*.*

Proof. Let $X \neq \emptyset$ and $X \subseteq A$. Let $\alpha = \min\{\text{rank}(y) : y \in X\}$. Fix $y \in X$ so that $\text{rank}(y) = \alpha$. Then, y is $\in$-minimal by Lemma 5.1.6. $\square_{5.2.2}$

NOTE: The converse of Lemma 5.2.2 does not necessarily hold. For example, if $x = \{y\}$, $y = \{x\}$, and $x \neq y$, then $y \notin \mathbf{WF}$, but $\in$ is well-founded (in fact, empty) on y. The converse of Lemma 5.2.2 is true, however, if we make further assumptions:

Lemma 5.2.3 (ZF$^-$)**.** *If* A *is a transitive set and* $\in$ *is well-founded on* A*, then* $A \in \boldsymbol{WF}$*.*

Proof. By Lemma 5.1.10, it suffices to show that $A \subseteq \mathbf{WF}$. If $A \not\subseteq \mathbf{WF}$, then let $X = A \setminus \mathbf{WF}$, and let y be the $\in$-minimal element in X. Such an element exists because $X \neq \emptyset$ and $\in$ is well-founded on A. If $z \in y$, then $z \notin X$. But, $z \in y \in X \subseteq A$, hence by the transitivity of the set A, we have $z \in A \setminus X \subseteq \mathbf{WF}$. So, every element z is an element of $\mathbf{WF}$, which implies that $y \subseteq \mathbf{WF}$. Hence from this and Lemma 5.1.10, $y \in \mathbf{WF}$, contrary to the definition of y as an element of A outside of $\mathbf{WF}$. $\square_{5.2.3}$

We now show that a set $A \in \mathbf{WF}$ if and only if $\in$ is a well-founded relation on the *transitive closure* of A, that is, on the least transitive set containing A as a subset. To do this, we need some definitions.

Definition 5.2.4 (ZF$^-$ – P)**.**

1. By induction on $n < \omega$, we define $\bigcup^0 A = A$, and $\bigcup^{n+1} A = \bigcup(\bigcup^n A)$);
2. $\text{trcl}(A) = \bigcup\{\bigcup^n A : n \in \omega\}$.

Thus $\text{trcl}(A) = A \cup \bigcup A \cup \bigcup^2 A \cup \cdots$, and has as elements the elements of A, and the elements of the elements of A, and so on.

Lemma 5.2.5 (ZF$^-$ – P)**.**

1. $A \subseteq \text{trcl}(A)$*;*
2. $\text{trcl}(A)$ *is a transitive set*
3. *If* $A \subseteq T$ *and* T *is a transitive set, then* $\text{trcl}(A) \subseteq T$*;*
4. *If* A *is transitive, then* $\text{trcl}(A) = A$*;*
5. *If* $x \in A$*, then* $\text{trcl}(x) \subseteq \text{trcl}(A)$*;*
6. $\text{trcl}(A) = A \cup \bigcup\{\text{trcl}(x) : x \in A\}$*.*

Proof.

1. The statement is obvious from the definition.
2. Notice that if $y \in \bigcup^n A$, then $y \subseteq \bigcup^{n+1} A$.

3. To show this, show by induction that $\bigcup^n A \subseteq T$.
4. This is a consequence of 1 and 3 and taking $A = T$.
5. If $x \in A$, then $x \in \text{trcl}(A)$, and consequently, $x \subseteq \text{trcl}(A)$. Now, apply 3 to x.
6. Let $T = A \cup \bigcup\{\text{trcl}(x) : x \in A\}$. Then T is transitive. Whence, by 3, we have $\text{trcl}(A) \subseteq T$. In the other direction, by 1 and 3, we have $T \subseteq \text{trcl}(A)$.

$\square_{5.2.5}$

Theorem 5.2.6 (ZF$^-$)**.** *For any set A, the following are equivalent.*

1. $A \in \mathbf{WF}$;
2. $\text{trcl}(A) \in \mathbf{WF}$;
3. $\in$ *is well-founded on* $\text{trcl}(A)$.

Proof. **1⇒2:** If $A \in \mathbf{WF}$, then by Lemma 5.1.8 and induction on n, we have that $\bigcup^n A \in \mathbf{WF}$. Thus, $\bigcup^n A \subseteq \mathbf{WF}$ and further $\text{trcl}(A) \subseteq \mathbf{WF}$. And so $\text{trcl}(A) \in \mathbf{WF}$ by Lemma 5.1.10.

2⇒3: This is the content of Lemma 5.2.2.

3⇒1: With this assumption 3 and Lemma 5.2.3, we have that $A \subseteq \text{trcl}(A) \subseteq \mathbf{WF}$. Hence, $A \subseteq \mathbf{WF}$. Consequently, by Lemma 5.1.10, $A \in \mathbf{WF}$. $\square_{5.2.6}$

Our definition of **WF** used the Powerset Axiom in a vital way. The equivalent statement 3 in Theorem 5.2.6 is useful particularly if one wants to define the class **WF** in some weaker theory that does not assume the Powerset axiom, for example ZF$^-$ − P.

5.3 The Axiom of Foundation

Since all of mathematics can take place in the class **WF**, one can make the case that it is reasonable to take as an axiom the statement $\mathbf{V} = \mathbf{WF}$. That is, it is reasonable to restrict our domain of discourse of sets to only the well-founded ones. Clearly, the axioms of ZF$^-$ are still true under such an interpretation since **WF** is closed under those set theoretic operations like $\bigcup$ and $\mathscr{P}()$ whose existence is given by the axioms of ZF$^-$. (more about this perhaps later (relativization)) In this section, we will talk about some of the consequences of taking as an axiom the statement $\mathbf{V} = \mathbf{WF}$.

The statement $\mathbf{V} = \mathbf{WF}$ is very non-elementary since it requires such a huge lot of definitions. So, instead, we assume an equivalent statement that is easily stated in the first-order language of set-theory. This is simply the Axiom of Foundation, or Regularity, which was mentioned in the first lecture:

$$\forall x\,(\exists y \in x \rightarrow \exists y \in x\,(\neg\exists z\,(z \in y \wedge z \in x))).$$

Equivalently: if $x \neq \emptyset$, then $\exists y \in x\,(x \cap y = \emptyset)$. Or: every non-empty set has a $\in$-minimal element. Or, if we extend the definition of well-foundedness to proper classes: $\in$ is well-founded on $\mathbf{V}$.

Theorem 5.3.1 (ZF$^-$)**.** *The following are equivalent:*

1. *The Axiom of Foundation*

2. $\forall A\,(\in$ *is well-founded on* $A)$*;*

3. $\mathbf{V} = \boldsymbol{WF}$.

Proof. That **1⇔2** is obvious from the definition of well-foundedness.

2⇒3: Statement 2 implies that for any set A, $\in$ is well-founded on $\operatorname{trcl}(A)$, and hence $A \in \mathbf{WF}$.

Lemma 5.2.2 gives us the implication **3⇒2**. $\square_{5.3.1}$

Unlike the "normal" axioms of ZFC, the Axiom of Foundation does not have applications in ordinary mathematics, since assuming this axiom limits our attentions to **WF**, where all of normal mathematics takes place anyway. Assuming this axiom just lets us be rid of certain pathologies, such as sets x where $x \in x$, or sets x and y where $x \in y \wedge y \in x$.

Since the Axiom of Foundation is equivalent with $\mathbf{WF} = \mathbf{V} = \bigcup\{R_\alpha : \alpha \in \mathbf{ON}\}$, it gives us a picture of all sets as being created by an iterative process, starting from nothing.

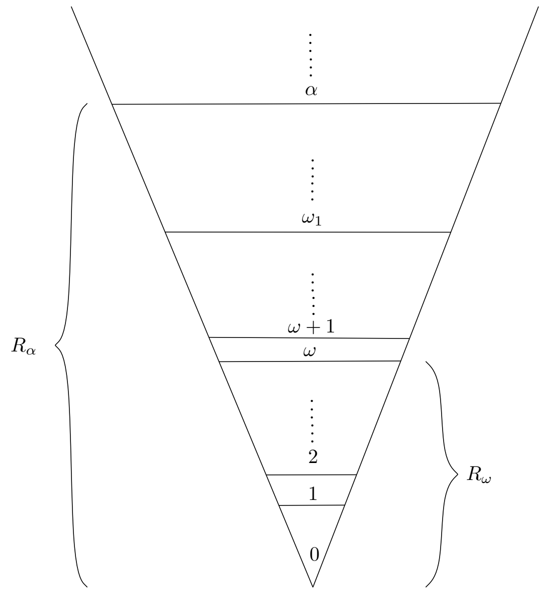

Figure 5.1: *The well-founded universe.*

Assuming that $\in$ is well-founded on every set simplifies certain definitions. The simplest example is probably the following theorem:

Theorem 5.3.2 (ZF − P)**.** *A set A is an ordinal iff A is transitive and linearly ordered by $\in$.*

The above theorem will be important later on.

——————— HERE ENDED WINTER 2007 LECTURE 6 ———————

——————— HERE ENDED SPRING 2009 WEEK 6 (4 hrs) ———————

5.4 Induction and Recursion on Well-founded Relations

If R is a well founded relation on A, then a proof by transfinite induction on R is one in which one proves that $\forall X \in A\, \phi(x)$ by first showing that for all $x \in A$,

$$\forall y \in A\, (yRx \to \phi(y)) \to \phi(x).$$

The conclusion that $\forall x\, \phi(x)$ is justified, because an R-smallest element of $\{x \in A : \neg\phi(x)\}$ would lead to a contradition.

For example, we can look at the proof of Lemma 5.2.3 (which says that if A is transitive and $\in$ is well-founded on A, then $A \subset \mathbf{WF}$) as a transfinite induction. Here $\phi(x)$ is the statement "x is well-founded" and the formula above that is to be proved reduces to $x \subset \mathbf{WF} \to x \in \mathbf{WF}$.

It is often useful to consider the notions of well-foundedness on proper classes as well.

Definition 5.4.1 (ZF$^-$ − P)**.** A class $\mathbf{R}$ is well-founded on a class $\mathbf{A}$ iff

$$\forall X \subset \mathbf{A}\, (X \neq \emptyset \to \exists y \in X (\neg \exists z \in X\, (z\mathbf{R}y))). \tag{5.1}$$

This is exactly the translation of the definition of well-foundedness for sets. There is on formal difference: here we are working with *classes*. The definition of well-foundedness for sets defines a formula with two variables, R and A. The definition for classes instead is a schema of definitions. Given formulas definition $\mathbf{R}$ and $\mathbf{A}$, 5.1 becomes an abbreviation for another formula. For example, "$\in$ is well-founded on $\mathbf{V}$" is a sentence in the language of set theory which is equivalent to the Axiom of Foundation.

Let us note also that the variable X in the above definition must range over *subsets* of the class $\mathbf{A}$, since there is no formal way to quantify over classes. This can cause problems if we try to justify a proof by transfinite induction, since we would need the existence of an $\mathbf{R}$-minimal element in the class $\{x \in \mathbf{A} : \neg\phi(x)\}$. This last class might be a proper class! In practice, we will only be concerned with relations that satisfy an addition condition which removes this problem:

Definition 5.4.2 (ZF$^-$ − P)**.** A class $\mathbf{R}$ is *set-like* on $\mathbf{A}$ iff for all $x \in \mathbf{A}$, the class $\{y \in \mathbf{A} : y\mathbf{R}x\}$ is a set.

For example, the relation $\in$ is set-like on every class $\mathbf{A}$, and every relation on a set is set-like.

Definition 5.4.3 (ZF$^-$ − P)**.** If $\mathbf{R}$ is set-like on $\mathbf{A}$ and $x \in \mathbf{A}$, then

1. $\mathrm{pred}(\mathbf{A}, x, \mathbf{R}) = \{y \in \mathbf{A} : y\mathbf{R}x\}$.
2. $\mathrm{pred}^0(\mathbf{A}, x, \mathbf{R}) = \mathrm{pred}(\mathbf{A}, x, \mathbf{R})$;
   $\mathrm{pred}^{n+1}(\mathbf{A}, x, \mathbf{R}) = \bigcup\{\mathrm{pred}(\mathbf{A}, y, \mathbf{R}) : y \in \mathrm{pred}^n(\mathbf{A}, x, \mathbf{R})\}$.
3. $\mathrm{cl}(\mathbf{A}, x, \mathbf{R}) = \bigcup\{\mathrm{pred}^n(\mathbf{A}, x, \mathbf{R}) : n \in \omega\}$.

Note that all the objects defined above are sets. If $\mathbf{R}$ is $\in$, and $\mathbf{A}$ is transitive, then $\mathrm{pred}(\mathbf{A}, x, \mathbf{R}) = x$, $\mathrm{pred}^n(\mathbf{A}, x, \mathbf{R} = \bigcup^n x$, and $\mathrm{cl}(\mathbf{A}, x, \mathbf{R}) = \mathrm{trcl}(x)$. The fact that $\mathrm{trcl}(x)$ is a transitive set is generalizes to the following lemma, the proof of which is clear from the definitions:

Lemma 5.4.4 ($\mathrm{ZF}^- - \mathrm{P}$)**.** *Let* $\mathbf{R}$ *be well-founded and set-like on* $\mathbf{A}$*. Then for all* $y \in \mathrm{cl}(\mathbf{A}, x, \mathbf{R})$ *we have that* $\mathrm{pred}(\mathbf{A}, y, \mathbf{R}) \subseteq \mathrm{cl}(\mathbf{A}, x, \mathbf{R})$*.*

Theorem 5.4.5 (($\mathrm{ZF}^- - \mathrm{P}$) Transfinite Induction on well-founded relations)**.** *If* $\mathbf{R}$ *is well-founded and set-like on* $\mathbf{A}$*, then for every non-empty class* $\mathbf{X} \subseteq \mathbf{A}$*, the class* $\mathbf{X}$ *has an* $\mathbf{R}$*-minimal element.*

Proof. Fix $x \in \mathbf{X}$. If x is not $\mathbf{R}$-minimal in $\mathbf{A}$, then $\mathbf{X} \cap \mathrm{cl}(\mathbf{A}, x, \mathbf{R})$ is a non-empty *subset* of $\mathbf{A}$, and hence has an $\mathbf{R}$-minimal element y. By Lemma 5.4.4, y is clearly the $\mathbf{R}$-minimal element of $\mathbf{X}$. $\square_{5.4.5}$

The special case of this theorem where $\mathbf{A} = \mathbf{ON}$ and $\mathbf{R} = \in$ was already proved (Theorem 3.5.2). There we also mentioned how one can give this result without using classes.

By Theorem 5.4.5, proofs using transfinite recursion on well-founded set-like relations are justified.

We can also define functions by transfinite recursion on well-founded set-like relations. Again, the special case of function on $\mathbf{ON}$ has already been discussed (Theorem 3.5.3). Again, we have already discussed how to give similar results without using classes.

Theorem 5.4.6 ($\mathrm{ZF}^- - \mathrm{P}$ Transfinite Recursion on well-founded relations)**.** *Assume that* $\mathbf{R}$ *is well-founded and set-like on* $\mathbf{A}$*. If* $\mathbf{F} : \mathbf{V} \times \mathbf{V} \longrightarrow \mathbf{V}$*, then there exists a unique* $\mathbf{G} : \mathbf{A} \longrightarrow \mathbf{V}$ *such that*

$$(\forall x \in \mathbf{A})(\mathbf{G}(x) = \mathbf{F}(x, \mathbf{G} \restriction \mathrm{pred}(\mathbf{A}, x, \mathbf{R}))).$$

Proof. We directly generalize the proof of Theorem 3.5.3 which was the special case for $\mathbf{A} = \mathbf{ON}$ and $\mathbf{R} = \in$.

The uniqueness of $\mathbf{G}$ is easily shown using transfinite induction on $\mathbf{R}$, so we will now concern ourselves with the proof for existence.

We will call a *set* $d \subset \mathbf{A}$ *closed* iff $(\forall x \in d)(\mathrm{pred}(\mathbf{A}, x, \mathbf{R}) \subseteq d)$. Closed sets will play the role that ordinals played in the proof of Theorem 3.5.3. Let us first notice that every element $x \in \mathbf{X}$ is contained in some closed set, that is, in $\{x\} \cup \mathrm{cl}(\mathbf{A}, x, \mathbf{R})$. If d is closed, then we will call a function g with domain d a d-approximation if

$$(\forall x \in d)(g(x) = \mathbf{F}(x, g \restriction \mathrm{pred}(\mathbf{A}, x, \mathbf{R}))).$$

As in the proof of uniqueness, we show that if g is a d-approximation and g' is a d'-approximation, then $g \restriction (d \cap d') = g' \restriction (d \cap d')$.

We now show by induction on $\mathbf{R}$, that for all x, there exists an $(\{x\} \cup \mathrm{cl}(\mathbf{A}, x, \mathbf{R}))$-approximation: Let us assume that this holds for all $y\mathbf{R}x$. Let g_y be a $(\{y\} \cup \mathrm{cl}(\mathbf{A}, y, \mathbf{R}))$-approximation. Then $h = \bigcup\{g_y : y\mathbf{R}x\}$ is a $\mathrm{cl}(\mathbf{A}, x, \mathbf{R}))$-approximation, and $h \cup \{\langle x, \mathbf{F}(x, h)\rangle\}$ is an $(\{x\} \cup \mathrm{cl}(\mathbf{A}, x, \mathbf{R}))$-approximation.

Now, we define $\mathbf{G}(x)$ as the value $g(x)$, where g is a d-approximation for some (any) closed set d containing x. $\square_{5.4.6}$

As an application of Theorem 5.4.6, let us look at the rank function. Consider the equation

$$\mathrm{rank}(y) = \sup\{\mathrm{rank}(x) + 1 : x \in y\},$$

which was defined earlier for $y \in \mathbf{WF}$. We can now look at this as a definition of $\mathrm{rank}(x)$ defined by transfinite recursion on $\in$, which is well-founded on $\mathbf{WF}$. More generally, we can define rank in the following manner:

Definition 5.4.7 ($\mathrm{ZF}^- - \mathrm{P}$)**.** If $\mathbf{R}$ is a well-founded set-like relation on the class $\mathbf{A}$, then

$$\mathrm{rank}(x, \mathbf{A}, \mathbf{R}) = \sup\{\mathrm{rank}(y, \mathbf{A}, \mathbf{R}) + 1 : y\mathbf{R}x \wedge y \in \mathbf{A}\}.$$

Note that formally the $\mathbf{F}$ from Theorem 5.4.6 is here given by $\mathbf{F}(x, h) \sup\{\alpha + 1 : \alpha \in \mathrm{rng}(h)\}$.

Lemma 5.4.8 (ZF^-)**.** *If $\mathbf{A}$ is transitive and $\in$ is well-founded on $\mathbf{A}$, then $\mathbf{A} \subseteq \boldsymbol{WF}$, and $\mathrm{rank}(x, \mathbf{A}, \in) = \mathrm{rank}(x)$.*

Proof. If $\mathbf{A} \not\subset \mathbf{WF}$, the let x be the $\in$-minimal element of $\mathbf{A} \setminus \mathbf{WF}$. Then, $x \subseteq \mathbf{A}$, since $\mathbf{A}$ is transitive. Hence, $x \subseteq \mathbf{WF}$, thus $x \in \mathbf{WF}$ by Lemma 5.1.10. Similarly, the $\in$-minimal element of the class $\{x \in \mathbf{A} : \mathrm{rank}(x, \mathbf{A}, \in) \neq \mathrm{rank}(x)\}$ gives a contradiction by Lemma 5.1.6. $\square_{5.4.8}$

Definition 5.4.7 allows us to define rank on $\mathbf{WF}$ without the use of the Powerset Axiom.

Another application of recursion on well-founded relations generalizes the fact that every well-order R on a set A is isomorphic to an ordinal. One could look at the isomorphism G as defined by $G(a) = \{G(b) : bRa\}$. We can generalize this:

Definition 5.4.9 ($\mathrm{ZF}^- - \mathrm{P}$)**.**

1. Let $\mathbf{R}$ be a well-founded set-like relation on a class $\mathbf{A}$. We define the *Mostowski collapsing function* $\mathbf{G}$ for $\mathbf{A}$ and $\mathbf{R}$ by

$$\mathbf{G}(x) = \{\mathbf{G}(y) : y \in \mathbf{A} \wedge y\mathbf{R}x\}.$$

2. The *Mostowski collapse* $\mathbf{M}$ of $\mathbf{A}$ and $\mathbf{R}$ is the image of $\mathbf{G}$.

The "function" $\mathbf{G} : \mathbf{A} \longrightarrow \mathbf{M}$ does not have to be 1-1. For example, if $\mathbf{R} = \emptyset$, then $\mathbf{G}(x) = \emptyset$ for every $x \in \mathbf{A}$. Then, $\mathbf{M} = \{\emptyset\}$ if $\mathbf{A} \neq \emptyset$.

Lemma 5.4.10 ($\mathrm{ZF}^- - \mathrm{P}$)**.** *With notation as in Definition 5.4.9,*

1. $\forall x, y \in \mathbf{A}\,(x\mathbf{R}y \rightarrow \mathbf{G}(x) \in \mathbf{G}(y)$.
2. $\mathbf{M}$ *is transitive.*
3. *(ZF^-)* $\mathbf{M} \subset \boldsymbol{WF}$.
4. *(ZF^-) If $x \in \mathbf{A}$, then* $\mathrm{rank}(x, \mathbf{A}, \mathbf{R}) = \mathrm{rank}(\mathbf{G}(x))$.

Proof. The proofs of 1 and 2 are immediate from the definition.

To prove 3, we show that $(\forall x \in \mathbf{A})(\mathbf{G}(x) \in \mathbf{WF})$ using induction on x.

To show 4, notice that

$$\mathrm{rank}(\mathbf{G}(x)) = \sup\{\mathrm{rank}(y) + 1 : y \in \mathbf{G}(x)\} = \sup\{\mathrm{rank}(\mathbf{G}(y)) + 1 : y\mathbf{R}x\}.$$

Then, $\mathrm{rank}(\mathbf{G}(x)) = \mathrm{rank}(x, \mathbf{A}, \mathbf{R})$ by induction on x. $\square_{5.4.10}$

In many interesting cases, the Mostowski collapsing function is in fact an isomorphism. There is a special condition for this to be the case.

Definition 5.4.11 ($\mathrm{ZF}^- - \mathrm{P}$)**.** $\mathbf{R}$ is *extensional* on $\mathbf{A}$ iff

$$\forall x, y \in \mathbf{A}\, (\forall z \in \mathbf{A}\, (z\mathbf{R}x \iff z\mathbf{R}y) \to x = y).$$

This is equivalent to saying that the Axiom of Extensionality is true in $\mathbf{A}$ if $\in$ is interpreted as $\mathbf{R}$. One can also put this in another, also convenient way: $\mathbf{R}$ is extensional on $\mathbf{A}$ iff for all $x, y \in \mathbf{A}$, if $x \neq y$, then $\mathrm{pred}(\mathbf{A}, x, \mathbf{R}) \neq \mathrm{pred}(\mathbf{A}, y, \mathbf{R})$. From this way of stating extensionality, it is clear that, for example, all linear orderings are extensional. Another class of examples is given by the following:

Lemma 5.4.12 ($\mathrm{ZF}^- - \mathrm{P}$)**.** *If* $\mathbf{N}$ *is transitive, then* $\in$ *is extensional on* $\mathbf{N}$*.*

Proof. Notice that $\mathrm{pred}(\mathbf{N}, x, \in) = x$ $\square_{5.4.12}$

By Lemmas 5.4.10(2) and 5.4.12, we see that the collapsing function cannot be an isomorphism unless $\mathbf{R}$ is extensional on $\mathbf{A}$. Conversely, the following applies:

Lemma 5.4.13 ($\mathrm{ZF}^- - \mathrm{P}$)**.** *Using the notation of Definition 5.4.9, if* $\mathbf{R}$ *is extensional on* $\mathbf{A}$*, then* $\mathbf{G}$ *is an isomorphism. In other words,* $\mathbf{G}$ *is 1-1 and* $\forall x, y \in \mathbf{A}\, (x\mathbf{R}y \iff \mathbf{G}(x) \in \mathbf{G}(y))$*.*

Proof. First we show that $\mathbf{G}$ is 1-1. Assume that it is not, and take x $\mathbf{R}$-minimal in $\{x \in \mathbf{A} : \exists y \in \mathbf{A}\, (x \neq y \wedge \mathbf{G}(x) = \mathbf{G}(y))\}$, and fix some $y \neq x$ such that $\mathbf{G}(x) = \mathbf{G}(y)$.

Since $\mathbf{R}$ is extensional, we have two possible cases:

Case 1: For some $z \in \mathbf{A}$, $z\mathbf{R}x$ and $\neg z\mathbf{R}y$. Since $\mathbf{G}(z) \in \mathbf{G}(x) = \mathbf{G}(y)$, we have that $\mathbf{G}(z) = \mathbf{G}(w)$ for some w such that $w\mathbf{R}y$. Then $w \neq z$, and z contradicts the minimality of x.

Case 2: For some $w \in \mathbf{A}$, $w\mathbf{R}y$ and $\neg w\mathbf{R}x$. Then, as in Case 1, there exists z such that $z\mathbf{R}x$ and $\mathbf{G}(z) = \mathbf{G}(w)$. Again, the existence of such a z contradicts the minimality of x.

Since $\mathbf{G}$ is 1-1, the fact that $\mathbf{G}$ is an isomorphism results directly from the definition. $\square_{5.4.13}$

We summarize in the following Theorem:

Theorem 5.4.14 (($\mathrm{ZF}^- - \mathrm{P}$) Mostowski Collapsing Theorem)**.** *Suppose* $\mathbf{R}$ *is well-founded, set-like, and extensional on* $\mathbf{A}$*. Then there exists a transitive class* $\mathbf{M}$ *and a 1-1 mapping* $\mathbf{G}$ *from* $\mathbf{A}$ *onto* $\mathbf{M}$ *such that* $\mathbf{G}$ *is an isomorphism between* ($\mathbf{A}$,$\mathbf{R}$) *and* $(\mathbf{M}, \in)$*. Furthermore,* $\mathbf{M}$ *and* $\mathbf{G}$ *are unique.*

As you read this proof, you may wish to also look back at the proof of Theorem 3.2.6, the argument of which is very very similar.

Proof. We have existence from Lemma 5.4.13.

For uniqueness, assume that $\mathbf{M}'$ and $\mathbf{G}'$ also satisfy the theorem. Then, by induction on x, $\mathbf{G}'(x) = \mathbf{G}(x)$ for all $x \in \mathbf{A}$. This implies that $\mathbf{M}' = \mathbf{M}$. $\square_{5.4.14}$

As an example for an application of Theorem 5.4.14, let us look at the situation where **R** well-orders **A**. If **A** is a set, the **A** is an ordinal. If **A** is a proper class, then $\mathbf{M} = \mathbf{ON}$. The assumption that **R** is set-like prevents **R** from having "'type' $> \mathbf{ON}$". For example, $\mathbf{ON} \times 2$ ordered lexicographically has "'type' $\mathbf{ON} + \mathbf{ON}$", but cannot be isomorphic with $\in$ on any class.

Corollary 5.4.15 (($\mathrm{ZF}^- - \mathrm{P}$))**.** *If $\in$ is extensional on **A**, then there exists a transitive **M** and a 1-1 mapping **G** from **A** onto **M** which is an isomorphism for the $\in$ relation. In other terms,*

$$\forall x, y \in \mathbf{A}\,(x \in y \iff \mathbf{G}(x) \in \mathbf{G}(y)).$$

——————— HERE ENDED WINTER 2006 LECTURE 7 ———————

Chapter 6

Relativization, Absoluteness, and Reflection in Consistency Results

AGAIN, THIS NEEDS TO BE PUT INTO LINE WITH ACCEPTED MODEL-THEORETIC TERMS. ESPECIALLY THE TARSKI-VAUGHT CRITERION IS UNCLEAR HERE.

In this section, we will introduce some techniques that will be needed for later consistency results. Along the way, we will present some easier applications, such as

$$\mathrm{Con}(\mathrm{ZF}^-) \to \mathrm{Con}(\mathrm{ZF})$$

to demonstrate how these techniques are used. We will also prove the reflection theorem, which we will use to show that ZF is not finitely axiomatizable.

Earlier, we showed that ZF^- encompasses Peano Arithmetic. Therefore, the Gödel Incompleteness theorem applies. Thus, one cannot prove the consistency of ZF^- by an argument formalizable within ZF^-. We will take the consistency of ZF^- therefore as an article of faith. Our consistency results will then actually be *relative* consistency results. That is, dependent upon the assumption that ZF^- is consistent, we will show that various other systems (such as ZF, ZF $-$ Infinity $+\neg$Infinity, etc.) are consistent.

6.1 Relativization

We make the idea of truth relative to a model precise.

Definition 6.1.1. Let $\mathbf{M}$ be any class. Then for any formula ϕ, we define $\phi^{\mathbf{M}}$, the *relativization* of ϕ to $\mathbf{M}$, by induction on the complexity of ϕ by:

1. $(x = y)^{\mathbf{M}}$ is $x = y$.
2. $(x \in y)^{\mathbf{M}}$ is $x \in y$.
3. $(\phi \wedge \psi)^{\mathbf{M}}$ is $\phi^{\mathbf{M}} \wedge \psi^{\mathbf{M}}$.

4. $(\neg\phi)^{\mathbf{M}}$ is $\neg(\phi^{\mathbf{M}})$.

5. $(\exists x\phi)^{\mathbf{M}}$ is $\exists x\,(x \in \mathbf{M} \wedge \phi^{\mathbf{M}})$.

If one wishes to be excessively formal, then note that $\mathbf{M}$ is in reality a formula $\mathbf{M}(v)$, ϕ is another formula, and we are defining in the metalanguage a third formula $\phi^{\mathbf{M}}$. Hence, (5) really should be $\exists x\,(\mathbf{M}(x) \wedge \phi^{\mathbf{M}})$.

In the definition of $\phi^{\mathbf{M}}$, the interpretation of the symbol $\in$ is unchanged. One could also consider other interpretations of this symbol. I will probably not cover this in this lecture. If the student is interested, he or she is advised to look at Kunen's Set Theory, Chapter IV, section 8 for a discussion of this.

We have defined $\phi^{\mathbf{M}}$ only for the official unabbreviated formulas as given at the beginning of the lecture. Note that the *logical* abbreviations we have defined thus far will have their intended meaning. For example, $(\phi \vee \psi)^{\mathbf{M}}$ is indeed $\neg(\neg\phi \wedge \neg\psi)^{\mathbf{M}}$ which from the above definition is $\neg(\neg(\phi^{\mathbf{M}}) \wedge \neg(\psi^{\mathbf{M}}))$, which is $\phi^{\mathbf{M}} \vee \psi^{\mathbf{M}}$. Similarly, $(\forall x\,\psi)^{\mathbf{M}}$ is the formula $\forall x\,(x \in \mathbf{m} \to \psi^{\mathbf{M}})$. Note that the situation for *set theoretical* abbreviations such as $\subset$ and $\mathscr{P}()$ can be a lot more complicated.

Definition 6.1.2. Let $\mathbf{M}$ be any class.

1. For a sentence ϕ, "*ϕ is true in* $\mathbf{M}$" means $\phi^{\mathbf{M}}$.

2. For a set of sentences S, the statement "*S is true in* $\mathbf{M}$" or "$\mathbf{M}$ *is a model for S*", means that every sentence in S is true in $\mathbf{M}$.

Intuitively, 1 and 2 are variants of the same idea, but formally, they are entirely different. The sentence "*ϕ is true in* $\mathbf{M}$" is an abbreviation, or another way of writing, $\phi^{\mathbf{M}}$. On the other hand, "*S is true in* $\mathbf{M}$" is in essence an abbreviation of a sentence in the *metatheory* that for each ϕ in S, we can prove $\phi^{\mathbf{M}}$ from the axioms we are presently using.

We need a basic result from logic to be able to get relative consistency results:

Lemma 6.1.3. *Let S and T be two sets of sentences in the language of set theory. Assume that for some class (i.e. predicate)* $\mathbf{M}$*, we can prove from T that* $\mathbf{M} \neq \emptyset$ *and that* $\mathbf{M}$ *is a model for S. Then,* $\mathrm{Con}(T) \to \mathrm{Con}(S)$.

Proof. If S were inconsistent, then we could prove $\phi \wedge \neg\phi$ from S. We have assumed that, using T, we can show that S is true in $\mathbf{M}$. Therefore, we can show using T that $\phi^{\mathbf{M}} \wedge \neg\phi^{\mathbf{M}}$, which gives a contradiction. Therefore, T is inconsistent. $\square_{6.1.3}$

In practice (in this lecture), the theory T in the lemma above will be some version of set theory, e. g. ZF^-, ZF, ZFC, or something similar.

——————— HERE ENDED WINTER 2007 LECTURE 7 ———————

We now examine our axioms in terms of the properties of the models that can satisfy them.

Note that $((\exists x)(x = x))^{\mathbf{M}}$ is equivalent to the statement that $\mathbf{M} \neq \emptyset$. We will always assume that $\mathbf{M}$ is non-empty, and hence, that the Set Existence Axiom, Axiom 0 is satisfied by $\mathbf{M}$

Let us now look at the Axiom of Extensionality, Axiom 1. When relativized to $\mathbf{M}$, it is

$$\forall x, y \in \mathbf{M}\,(\forall z \in \mathbf{M}\,(z \in x \iff z \in y) \to x = y).$$

Note that this is exactly the definition of $\in$ being extensional in a class $\mathbf{M}$. Since $\in$ is extensional on transitive classes, we get the following fact:

Lemma 6.1.4. *If* $\mathbf{M}$ *is transitive, then the Axiom of Extensionality is true in* $\mathbf{M}$.

If $\mathbf{M}$ is some given class, Axiom 2, the Separation Axiom, is typically not true in $\mathbf{M}$. That is, Separation generally holds only in very carefully constructed classes. The proof that it holds in some class is not entirely simple. One can, however, reduce the satisfaction of Separation by $\mathbf{M}$ to a closure property of $\mathbf{M}$:

Lemma 6.1.5. *Assume that for every formula* $\phi(x, z, v_1, \ldots, v_n)$ *without free variables other than the ones listed,*

$$\forall z, v_1, \ldots, v_n \in \mathbf{M}\,(\{x \in z : \phi^{\mathbf{M}}(x, z, v_1, \ldots, v_n)\} \in \mathbf{M}),$$

then each instance of the Separation Axiom Schema is true in $\mathbf{M}$.

Proof. We must verify that for every formula ϕ as in the statement of the lemma, we have

$$(\forall z, v_1, \ldots, v_n \in \mathbf{M})(\exists y \in \mathbf{M})(x \in y \iff x \in z \wedge \phi^{\mathbf{M}}(x, z, v_1, \ldots, v_n)),$$

since this is just the relativized version of Separation.

For given $z, v_1, \ldots, v_n \in \mathbf{M}$, let $y = \{x \in z : \phi^{\mathbf{M}}(x, z, v_1, \ldots, v_n)\}$. By assumption, $y \in \mathbf{M}$. Hence, for all x, and particularly for all $x \in \mathbf{M}$,

$$x \in y \iff x \in z \wedge \phi^{\mathbf{M}}(x, z, v_1, \ldots, v_n).$$

$\square_{6.1.5}$

It is also clear that if $\mathbf{M}$ also happens to be transitive, the requirement of Lemma 6.1.5 is still necessary for the Axiom of Separation to hold in $\mathbf{M}$. In practice, it is difficult to check that the requirements of Lemma 6.1.5 are satisfied because one would then have to look at the meaning of all possible formulas when relativized to $\mathbf{M}$. However, for this lecture, we will look at very simple models, and Separation will hold trivially thanks to:

Corollary 6.1.6. *If* $\forall z \in \mathbf{M}\,(\mathscr{P}(z) \subset \mathbf{M})$*, then the Separation Axiom is true in* $\mathbf{M}$.

We can now prove a very easy consistency result:

Theorem 6.1.7 (ZF$^-$)**.** *If* $\mathbf{M} = \{\emptyset\}$*, then the Axioms of Set Existence, Extensionality, and Separation, together with* $\forall y\,(y = \emptyset)$ *hold in* $\mathbf{M}$.

Proof. Here, we will consider the formula $\forall y\,(y = \emptyset)$ an abbreviation of $\forall y\, \forall x\,(x \notin y)$. This is true in $\mathbf{M}$ since $\emptyset \notin \emptyset$. Set Existence and Extensionality are true in $\mathbf{M}$ because $\mathbf{M}$ is transitive and non-empty. Separation is true by Corollary 6.1.6: every subset of every element of $\mathbf{M}$ (here, $\emptyset$) is in $\mathbf{M}$. $\square_{6.1.7}$

Hence, by Lemma 6.1.3, we have the following:

Corollary 6.1.8. $\mathrm{Con}(\mathit{ZF}^-) \to \mathrm{Con}(\mathit{Extensionality} + \mathit{Separation} + (\forall y\,(y = \emptyset)))$.

We have only defined relativization for formulas in the first-order language of set theory. This should cause no problems in theory because we have taken that the only proper formulas are those that only use $\in$ and $=$ and nothing else. Any other statement we have made has been understood to be just an abbreviation for a first-order formula. But, there are many statements that are of interest to us, such as CH and AC that are expressed using quite a few defined notions. We would like to check their validity in a model without actually writing down the unabbreviated statement.

If we had used abbreviation only in defining relations, then we would have no problems. We would just replace the relation with the formula that defines it. For example $z \subset x$ abbreviates $\forall v\,(v \in z \to v \in x)$, so $(z \subset x)^{\mathbf{M}}$ abbreviates $\forall v \in \mathbf{M}\,(v \in z \to v \in x)$. This is equivalent to $z \cap \mathbf{M} \subset x$.

If we now want to check that a statement that uses $\subset$ (for example, the Powerset Axiom) holds in $\mathbf{M}$, we do not need to write out the unabbreviated statement. The Powerset Axiom relativized to $\mathbf{M}$ is equivalent to

$$\forall x \in \mathbf{M}\,\exists y \in \mathbf{M}\,\forall z \in \mathbf{M}\,(z \cap \mathbf{M} \subset x \to z \in y).$$

In the special case that $\mathbf{M}$ is transitive (which is what will usually be the case in the examples we will look at), the relativized statement of the Powerset Axiom becomes still simpler. Then, $z \cap \mathbf{M} = z$ for all $z \in \mathbf{M}$, so for $z, y \in \mathbf{M}$, we have $(z \subset y)^{\mathbf{M}} \iff z \subset y$ (or, to use the terminology of the next section, $\subset$ is *absolute* for $\mathbf{M}$). Thus, for transitive $\mathbf{M}$, the Powerset Axiom holds in $\mathbf{M}$ iff

$$\forall x \in \mathbf{M}\,\exists y \in \mathbf{M}\,\forall z \in \mathbf{M}\,(z \subset x \to z \in y).$$

Therefore we have:

Lemma 6.1.9. *If $\mathbf{M}$ is transitive, the Power Set Axiom holds in $\mathbf{M}$ iff*

$$\forall x \in \mathbf{M}\,\exists y \in \mathbf{M}(\mathscr{P}(x) \cap \mathbf{M} \subset y).$$

When considering function and constants that we have defined via an abbreviation, we have to be a bit more careful. If S is a set of axioms and

$$S \vdash \forall x_1, \ldots, x_n\,\exists! y\,\phi(x_1, \ldots, x_n, y),$$

we can "define" $F(x_1, \ldots, x_n)$ to be the y such that $\phi(x_1, \ldots, x_n, y)$ holds. Formally, however, expressions using F are abbreviations for expressions that do not use F. If we wish to "unabbreviate" a given such F, we may have lots of possibilities. It is not clear which one we should take, since they will be all equivalent *on the basis of* S. However, they don't have to be equivalent in a class where S does not hold.

To give an example of this, let $\phi(y)$ be $\forall v\,(v \notin y)$. As long as S contains the axioms of Extensionality and Comprehension, we have that $S \vdash (\exists! y)\phi(y)$. We can then define that $\emptyset$ is exactly that y. Then, the expression $\emptyset \in z$ could be an abbreviation of either of the following formulas:

$$\psi(z) = \exists y\,(\phi(y) \wedge y \in z),$$

or

$$\chi(z) = \forall y(\phi(y) \to y \in z).$$

These two statements are equivalent when $(\exists! y)\phi(y)$.

Now, assume that $\mathbf{M}$ is $\{a, b, c\}$, where $a = \emptyset$, $b = \{\emptyset\}$, $c = \{\{\{\emptyset\}\}\}$. Then $\phi^{\mathbf{M}}(a)$ and $\phi^{\mathbf{M}}(c)$ are true, but $\psi^{\mathbf{M}}(b)$ is true, while $\chi^{\mathbf{M}}(b)$ is false.

To avoid these problems, we will only consider the relativizations to $\mathbf{M}$ of abbreviations involving F for which we have already checked that

$$\forall x_1, \ldots, x_n \, \exists! y \, \phi(x_1, \ldots, x_n, y) \tag{6.1}$$

holds in $\mathbf{M}$.

Usually, $\mathbf{M}$ will be a model for some axioms from which we can prove 6.1. If 6.1 holds in $\mathbf{M}$, then we can use $F^{\mathbf{M}}(x_1, \ldots, x_n)$ for the unique $y \in \mathbf{M}$ such that $\phi^{\mathbf{M}}(x_1, \ldots, x_n, y)$.

NOTE: If $\mathbf{M} = \{1, 2\}$, then $\exists! y \forall v \, (v \notin y)$ holds, and $\emptyset^{\mathbf{M}} = 1$. Since $1 \in 2$, we see that the sentence (abbreviated by) $\exists x \, (\emptyset \in x)$, is true in $\mathbf{M}$. On the other hand, if $\mathbf{M} = \{\emptyset\}$, then $\emptyset^{\mathbf{M}} = \emptyset$ and $\exists x \, (\emptyset \in x)$ is false in $\mathbf{M}$.

Now we can use these considerations to make more precise some of our statements about R_ω and $\mathbf{WF}$: Let $\mathbf{N}$ be one of these. Since $\mathbf{N}$ is transitive, it satisfies the Extensionality. If $x \in \mathbf{N}$, then $\mathscr{P}(x) \in \mathbf{N}$. Thus, by Corollary 6.1.6, $\mathbf{N}$ satisfies Separation, and by Lemma 6.1.9, it also satisfies Powerset. That $\mathbf{N}$ is closed under the Pair and Union Axioms follows from the fact that $\mathbf{N}$ is closed under the pairing and union operators, along with the following general fact:

Lemma 6.1.10. *If* $\forall x, y \in \mathbf{M} \, \exists z \in \mathbf{M} \, (x \in z \wedge y \in z)$ *and* $\forall x \in \mathbf{M} \, \exists z \in \mathbf{M} \, (\bigcup x \subset z)$*, then the Pairing and Union Axioms are true in* $\mathbf{M}$*.*

The Replacement Axiom, similarly to the Comprehension Axiom, can be difficult to check since it involves considering an arbitrary formula, but also like Comprehension, it is easy in R_ω and $\mathbf{WF}$. First, for convenience, we translate the relativization of this axiom:

Lemma 6.1.11. *Assume that we can show, for every formula* $\phi(x, y, A, v_1, \ldots, v_n)$ *and for every* $A, v_1, \ldots, v_n \in \mathbf{M}$*, if:*

$$(\forall x \in A)(\exists! y \in \mathbf{M}) \, \phi^{\mathbf{M}}(x, y, A, v_1, \ldots, v_n),$$

then

$$\exists Y \in \mathbf{M} \, (\{y : (\exists x \in A) \, \phi^{\mathbf{M}}(x, y, A, v_1, \ldots, v_n)\} \subset Y).$$

Then the Replacement Axiom Schema is true in $\mathbf{M}$*.*

We apply this to our class $\mathbf{N}$. Let

$$Y = \{y \in \mathbf{N} : (\exists x \in A) \, \phi^{\mathbf{N}}(x, y, A, v_1, \ldots, v_n)\}.$$

Then $Y \subset \mathbf{N}$. So, if $\mathbf{N} = \mathbf{WF}$, $Y \in \mathbf{N}$. If $\mathbf{N} = R_\omega$, then $|Y| \leqslant |A| < \omega$, so for some n, $Y \subset R_n$, and $Y \in R_{n+1} \subset \mathbf{N}$. Hence, Replacement holds in $\mathbf{N}$.

The Axiom of Foundation relativized to $\mathbf{M}$ is

$$(\forall x \in \mathbf{M}) \, (\exists y \in \mathbf{M}) \, (y \in x) \rightarrow (\exists y \in \mathbf{M}) \, (y \in x \wedge (\neg \exists z \in \mathbf{M}) \, (z \in x \wedge z \in y))).$$

If $\mathbf{M} \subseteq \mathbf{WF}$, then for a given $x \in \mathbf{M}$, we can take $y \in \mathbf{M} \cap x$ of minimal rank. In particular, we see that, working in ZF$^-$, Foundation holds in R_ω and $\mathbf{WF}$. More generally:

Lemma 6.1.12 (ZF$^-$)**.** *The Axiom of Foundation is true in any* $\mathbf{M} \subset \boldsymbol{WF}$*.*

Therefore:

Lemma 6.1.13 (ZF$^-$)**.** *The classes* $\boldsymbol{WF}$ *and* R_ω *are models of* $ZF - $ *Infinity.*

The Axiom of Infinity,

$$\exists x\,(\emptyset \in x \wedge (\forall y \in x)\,(S(y) \in x)),$$

involves both the notions $S()$ and $\emptyset$. Intuitively, this axiom is true in **WF** (just take $x = \omega$), and false in R_ω. The proof of this involves carefully checking that $S()$ and $\emptyset$ mean the same in R_ω and **WF** that they do in **V**. In other words, that these two notions are *absolute* for R_ω and **WF**. Instead of doing this specifically for these two cases, we will do a more general study of absoluteness in the next section. ———— HERE ENDED WINTER 2006 LECTURE 8 ————

6.2 Absoluteness

6.2.1 General facts about absoluteness

IF YOU LECTURE THIS AGAIN, YOU SHOULD PERHAPS BRING THIS SECTION ON ABSOLUTENESS IN LINE WITH MODEL THEORETIC TERMINOLOGY AS IN YOUR MODEL THEORY NOTES. ALSO, MAKE IT MORE CLEAR WHAT ARE PARAMETERS, WHAT ARE VARIABLES, AND WHAT ARE INTERPRETATIONS OF VARIABLES IN A MODEL. RIGHT NOW, THIS IS SLOPPY, AND PROBABLY CAUSES A BIT OF CONFUSION FOR THE STUDENTS.

We begin with a more precise definition of absoluteness

Definition 6.2.1. Let ϕ be a formula with free variables only among $x_1, \ldots, x_n$.

1. If $\mathbf{M} \subset \mathbf{N}$, then ϕ is *absolute for* $\mathbf{M}$ *and* $\mathbf{N}$ iff

$$(\forall x_1, \ldots, x_n \in \mathbf{M})(\phi^{\mathbf{M}}(x_1, \ldots, x_n) \iff \phi^{\mathbf{N}}(x_1, \ldots, x_n)).$$

2. The formula ϕ is *absolute for* $\mathbf{M}$ iff ϕ is absolute for $\mathbf{M}$ and $\mathbf{V}$. Equivalently,

$$(\forall x_1, \ldots, x_n \in \mathbf{M})(\phi^{\mathbf{M}}(x_1, \ldots, x_n) \iff \phi(x_1, \ldots, x_n)).$$

Notice that if a formula ϕ is absolute for $\mathbf{M}$ and absolute for $\mathbf{N}$ and $\mathbf{N} \subseteq \mathbf{M}$, then ϕ is absolute for $\mathbf{M}$ and $\mathbf{N}$.

In this section we will look at methods that will show that certain formulas (but not all formulas) are absolute for many of the models we will look at.

Since one builds a first order formula inductively, we will make sure that our methods have an inductive character. An example is the following:

Lemma 6.2.2. *If* $\mathbf{M} \subset \mathbf{N}$ *and* ϕ *and* ψ *are absolute for* $\mathbf{M}$ *and* $\mathbf{N}$*, then so are* $\neg\phi$ *and* $\phi \wedge \psi$*.*

Since the atomic formulas $x \in y$ and $x = y$ are absolute for all $\mathbf{M}$ (remember, their relativized versions are just themselves!). Every quantifier-free formula is built from atomic formulas using just $\neg$ and $\wedge$, whence we get the following:

Corollary 6.2.3. *If ϕ is a quantifier-free formula, then ϕ is absolute for all* $\mathbf{M}$.

Unfortunately, very simple formulas, such as that which is abbreviated by $x \subseteq y$, have quantifiers, hence they need not be absolute. Fortunately, if $\mathbf{M}$ is transitive, $x \subseteq y$ is, in fact, absolute for $\mathbf{M}$.

Lemma 6.2.4. *If $\mathbf{M} \subseteq \mathbf{N}$ are both transitive classes and ϕ is absolute for $\mathbf{M}$ and $\mathbf{N}$, then the formula $(\exists x)(x \in y \wedge \phi)$ is absolute for $\mathbf{M}$ and $\mathbf{N}$ as well.*

Proof. Let ϕ be the formula $\phi(x, y, z_1, \ldots, z_n)$ where we are listing its free variables. Then, for any $y, z_1, \ldots, z_n \in \mathbf{M}$, the following formulas are equivalent

$$((\exists x)(x \in y \wedge \phi(y, z_1, \ldots, z_n)))^{\mathbf{M}};$$
$$(\exists x)(x \in y \wedge \phi^{\mathbf{M}}(y, z_1, \ldots, z_n));$$
$$(\exists x)(x \in y \wedge \phi^{\mathbf{N}}(y, z_1, \ldots, z_n));$$
$$((\exists x)(x \in y \wedge \phi(y, z_1, \ldots, z_n)))^{\mathbf{N}}.$$

The first and last equivalence use the transitivity of $\mathbf{M}$ and $\mathbf{N}$. The middle equivalence is obtained by applying the assumption that ϕ is absolute. $\square_{6.2.4}$

We call $\exists x \in y$ a *bounded quantifier*. A formula in which all quantifiers are bounded is called a Δ_0*-formula*. Formally:

Definition 6.2.5. The Δ_0 formulas are built inductively using the following rules:

1. $x \in y$ and $x = y$ are Δ_0;
2. If ϕ and ψ are Δ_0, then so are $\phi \wedge \psi$ and $\neg\phi$;
3. If ϕ is Δ_0, then $\exists x \in y$ is also Δ_0.

Corollary 6.2.6. *If $\mathbf{M}$ is transitive and ϕ is Δ_0, then ϕ is absolute for* $\mathbf{M}$.

The usefulness of this result is limited by the fact that one rarely sees Δ_0-formulas in practice. For example, as established before, $x \subseteq y$ is an abbreviation of $\forall z\,(z \in x \rightarrow z \in y)$, which is itself an abbreviation of $\neg(\exists z)\neg(z \in x \rightarrow z \in y)$, which is clearly not a Δ_0-formula. In practice, Lemma 6.2.6 is used in conjunction with the following:

Lemma 6.2.7. *Let $\mathbf{M} \subseteq \mathbf{N}$ and assume that both $\mathbf{M}$ and $\mathbf{N}$ are models for a set of sentences S such that*

$$S \vdash (\forall x_1, \ldots, x_n)(\phi(x_1, \ldots, x_n) \iff \psi(x_1, \ldots, x_n)).$$

Then ϕ is absolute for $\mathbf{M}$ and $\mathbf{N}$ iff ψ is absolute for $\mathbf{M}$ and $\mathbf{N}$.

Note that $\forall x \in y$ is essentially a bounded quantifier, since $\forall x \in y \phi$ is logically equivalent to $\neg\exists x \in y \neg\phi$.

———— HERE ENDED WINTER 2007 LECTURE 8 ————

If we apply Lemma 6.2.7 with $\mathbf{M}$ transitive and $\mathbf{N} = \mathbf{V}$, and S the empty set of sentences, we can see that $x \subset y$ is absolute for $\mathbf{M}$. We already showed this in the previous section that is, when we noticed that $(z \subset x)^{\mathbf{M}} \Leftrightarrow z \cap \mathbf{M} \subset x$ and for transitive $\mathbf{M}$, $z \cap \mathbf{M} = z$, in which case $(z \subset x)^{\mathbf{M}} \Leftrightarrow z \subset x$., however, using the method we have just demonstrated, we can establish more absoluteness results. As before, we need to be careful with defined functions.

Definition 6.2.8. If $\mathbf{M} \subset \mathbf{N}$ and $F(x_1, \ldots, x_n)$ is a defined function, then we say that F is *absolute* for $\mathbf{M}$ and $\mathbf{N}$ iff the formula $F(x_1, \ldots, x_n) = y$ is absolute for $\mathbf{M}$ and $\mathbf{N}$.

Theorem 6.2.9. *The following relations and functions were defined in ZF$^-$ − P− Inf using formulas that are equivalent to Δ_0-formulas in ZF$^-$ − P− Inf. Hence they are absolute for every transitive class $\mathbf{M}$ which is a model of ZF$^-$ − P− Inf.*

1. $x \in y$	6. $\langle x, y\rangle$	11. $S(x)$ (i.e. $x \cup \{x\}$)
2. $x = y$	7. $\emptyset$	12. x is transitive
3. $x \subseteq y$	8. $x \cup y$	13. $y = \bigcup x$
4. $\{x, y\}$	9. $x \cap y$	14. $y = \bigcap x$ (where $\bigcap \emptyset = \emptyset$)
5. $\{x\}$	10. $x \setminus y$	

Proof. All of these statements have been defined before in $\mathrm{ZF}^- - \mathrm{P} - \mathrm{Inf}$. However, we were not particularly careful about using Δ_0-formulas in the definitions. We will do that now.

We have already discussed 1,2, and 3.

For 4, notice that the expression $z = \{x, y\}$ is equivalent to the expression $(x \in z \wedge y \in z \wedge (\forall v \in z)(v = x \vee v = y))$, which is clearly Δ_0.

Cases 5 and 6 are done similarly to 4. For example for 6, $z = \langle x, y\rangle = \{\{x\}, \{x, y\}\}$ is equivalent to the expression

$$((\exists v \in z)(v = \{x\}) \wedge (\exists v \in z)(v = \{x, y\}) \wedge (\forall v \in z)(v = \{x\} \vee v = \{x, y\})).$$

This formula is equivalent to a Δ_0-formula obtained by replacing $v = \{x\}$ and $v = \{x, y\}$ with the Δ_0-formulas with which they are equivalent.

For cases 7, 8, 9, and 11, notice that the expressions $z = \emptyset$, $y = x \cup y$, $z = x \cap y$, and $z = S(x)$ are equivalent to the formulas

$$((\forall v \in z)\neg(v = v));$$
$$((\forall v \in z)(v \in x \vee v \in y) \wedge (x \subseteq z) \wedge (y \subseteq x));$$
$$((\forall v \in z)(v \in x \rightarrow v \in y) \wedge (z \subseteq x) \wedge (z \subseteq y));$$
$$((x \in z) \wedge (x \subseteq z) \wedge (\forall v \in z)(v = x \vee v \in x)).$$

Case 10 is similar to 9.

For 12, 13, and 14, we have the following equivalent expressions:

$$((\forall v \in x)(\forall z \in v)(z \in x));$$
$$((\forall v \in x)(v \subseteq y) \wedge (\forall z \in y)(\exists v \in x)(z \in v));$$
$$((\forall v \in x)(y \subseteq v) \wedge (\forall v \in x)(\forall z \in v)((\forall w \in x)(z \in w) \rightarrow z \in y) \wedge (x = \emptyset \rightarrow y = \emptyset)).$$

Note that in 14, $\bigcap \emptyset$ "should" be $\mathbf{V}$, but we have defined that *ad hoc* as $\bigcap \emptyset = \emptyset$. This way $\bigcap \emptyset$ is a set. $\Box_{6.2.9}$

The student who is still awake may perhaps have noticed that there is a quicker way of proving case 6 in Theorem 6.2.9. Once one is sure that the unordered pair means the same in a transitive $\mathbf{M}$ as it means in $\mathbf{V}$, the same must be true of any compositions of such an operation, so in particular of the ordered pair.

Lemma 6.2.10. *Absolute notions are closed under composition. More precisely, let* $\mathbf{M} \subseteq \mathbf{N}$ *and assume that* $\phi(x_1, \ldots, x_n)$*,* $F(x_1, \ldots, x_n)$*, and* $G_i(y_1, \ldots, y_m)$ *where* $i = 1, \ldots, n$ *are absolute for* $\mathbf{M}$ *and* $\mathbf{N}$*. Then so are the formula*

$$\phi(G_1(y_1, \ldots, y_m), \ldots, G_n(y_1, \ldots, y_m)),$$

and the function

$$F(G_1(y_1, \ldots, y_m), \ldots, G_n(y_1, \ldots, y_m)).$$

Proof. Out of laziness, we assume $n = m = 1$. This will make writing the proof easier.

If $y \in \mathbf{M}$, then

$$\phi(G(y))^{\mathbf{M}} \iff \phi^{\mathbf{M}}(G^{\mathbf{M}}(y)) \iff \phi^{\mathbf{N}}(G^{\mathbf{N}}(y)) \iff \phi(G(y))^{\mathbf{N}}$$

because $G^{\mathbf{M}}(x) = G^{\mathbf{N}}(x)$ and ϕ is absolute for $\mathbf{M}$ and $\mathbf{N}$.

Similarly, we have

$$F(G(y))^{\mathbf{M}} = F^{\mathbf{M}}(G^{\mathbf{M}}(y)) = F^{\mathbf{N}}(G^{\mathbf{N}}(y)) = F(G(y))^{\mathbf{N}}.$$

$\square_{6.2.10}$

So, using Lemma 6.2.10 makes the proof of Case 6 of Theorem 6.2.9 much easier. We just have to write that

$$\langle x, y \rangle = F(G_1(x, y), G_2(x, y)),$$

where $G_1(x, y) = \{x\}$ and $F(x, y) = G_2(x, y) = \{x, y\}$, and use the fact that G_1 and F are absolute (i.e. Cases 4 and 5 of the same theorem).

The longer proof we argued earlier for Case 6 did give us something more than the above little argument does: it shows that ordered pairing is a Δ_0 function. Note that it is NOT true in general that the compositions of Δ_0 functions is Δ_0. Examples of this, however, require the Axiom of Foundation.

Nevertheless, the functions and relations in the next Theorem are in effect provably Δ_0.

Theorem 6.2.11. *The following functions and relations are absolute in every transitive model of* $ZF^- - P - Inf$*:*

1. z *is an ordered pair;*
2. $A \times B$*;*
3. R *is a relation;*
4. $\mathrm{dom}(R)$*;*
5. $\mathrm{rng}(R)$*;*
6. R *is a function;*
7. $R(x)$*;*
8. R *is 1-1 function.*

Proof.

1. A set z is an ordered pair iff $((\exists x \in \bigcup z)(\exists y \in \bigcup z)(z = \langle x, y\rangle))$, and this formula is obtained by substituting an absolute function into an absolute relation. Therefore, it is absolute by Lemma 6.2.10. To see formally why this is the case, notice that

$$z \text{ is an ordered pair} \iff \phi(G_1(z), G_2(z), G_3(z)),$$

where $G_1(z) = G_2(z) = \bigcup z$, which is absolute by Theorem 6.2.9, and $G_3(z) = z$, and the formula $\phi(a, b, c) = (\exists x \in z)(\exists y \in b)(c = \langle x, y\rangle)$. The formula ϕ is absolute because it has only bounded quantification of the absolute formula $c = \langle x, y\rangle$.

The other cases are argued similarly. Notice that:

2. $C = A \times B$ iff

$$((\forall x \in A)(\forall y \in B)(\langle x, y\rangle \in C) \wedge (\forall z \in C)(\exists x \in A)(\exists y \in B)(z = \langle x, y\rangle));$$

3. R is a relation iff

$$((\forall z \in R)(z \text{ is an ordered pair}));$$

4. $A = \text{dom}(R)$ iff

$$((\forall x \in A)(\exists y \in \bigcup\bigcup R)(\langle x, y\rangle \in R) \wedge \\ (\forall x \in \bigcup\bigcup R)(\forall y \in \bigcup\bigcup R)(\langle x, y\rangle \in R \to x \in A));$$

5. This one is very similar to the previous case.

6. R is a function iff

$$(R \text{ is a relation } \wedge (\forall x \in \bigcup\bigcup R)(\forall y \in \bigcup\bigcup R)(\forall y' \in \bigcup\bigcup R) \\ (\langle x, y\rangle \in R \wedge \langle x, y'\rangle \in R \to y = y'));$$

7. $y = R(x)$ iff

$$((\phi(x) \wedge \langle x, y\rangle \in R) \vee (\neg\phi(x) \wedge x = \emptyset));$$

where $\phi(x)$ is the formula

$$((\exists v \in \bigcup\bigcup R)(\langle x, v\rangle \in R \wedge (\forall w \in \bigcup\bigcup R)(\langle x, w\rangle \in R \to v = w))).$$

8. R is a 1-1 function iff

$$(R \text{ is a function } \wedge \\ \wedge (\forall x \in \text{dom}(R))(\forall x' \in \text{dom}(R))(R(x) = R(x') \to x = x')).$$

Hence, the listed notions are obtained from absolute notions via substitutions, bounded quantification, and logical connective, and are thus absolute as we desired to show.

Notice further that in 7. $R(x)$ is really a defined function of *two* variables, R and x. To be truly formal, we should have notated this as $appl(R, x)$, where the function $appl(R, x)$ is the unique y such that $\langle x, y\rangle \in R$ if such a y exists, or $\emptyset$ otherwise. $\square_{6.2.11}$

Of course, there are lots of functions that are absolute. For example, "f maps A to A and has no fixed points". Instead of listing all possible function we may encounter in the previous theorem, which would be silly, we will just say that they are absolute by standard arguments.

6.2.2 Absoluteness and the Axioms

Now we can return to our discussion of models of the axioms.

The absoluteness methods of the previous section make it easy to check that the Axiom of Infinity is true in a model.

Lemma 6.2.12. *Let* $\mathbf{M}$ *be a transitive model of* $ZF^- - P - Inf$. *If* $\omega \in \mathbf{M}$*, then the Axiom of Infinity is true in* $\mathbf{M}$.

Proof. By the absoluteness of $\emptyset$ and the successor function S, the Axiom of Infinity relativized to $\mathbf{M}$ is equivalent to the sentence

$$(\exists x \in \mathbf{M})(\emptyset \in \mathbf{M} \wedge (\forall y \in x)(S(y) \in x)),$$

which is true if $x = \omega$. $\square_{6.2.12}$

The same argument can be used to show that the Axiom of Infinity fails in R_ω, since every $x \in \mathbf{WF}$ containing $\emptyset$ and closed under S has an infinite rank. The next theorem is our last word on R_ω.

Theorem 6.2.13 (ZF^-)**.** *The set* R_ω *is a model for* $ZFC - Infinity + \neg Infinity$.

Proof. By the above discussion and Lemma 6.1.13, we only need to check that the Axiom of Choice holds in R_ω. To do this, we must show that

$$((\forall A \in R_\omega)(\exists R \in R_\omega)(R \text{ well orders } A)^{R_\omega}).$$

Fix $A \in R_\omega$. We know, even without assuming the Axiom of Choice, that A is finite, and can thus be well-ordered. Let $R \subseteq A \times A$ be a well-ordering of A. Then, $R \in R_\omega$. The fact that $(R \text{ well orders } A)^{R_\omega})$ follows from the next lemma.

Lemma 6.2.14 (ZF^-)**.** *Suppose that* $\mathbf{M}$ *is a transitive model of* $ZF^- - P - Inf$. *Let* $A, R \in \mathbf{M}$ *and assume that* R *well-orders* A. *Then* $(R \textit{ well orders } A)^{\mathbf{M}}$.

Proof. That $(R \text{ linearly orders } A)^{\mathbf{M}}$ we get by standard arguments, since this is a statement expressed using basic properties of pairs and using (bounded) quantification over A.

To check well-ordering, we have to check that $((\forall X)(\phi(X, A, R))^{\mathbf{M}}$, where $\phi(X, A, R)$ is the formula

$$X \subseteq A \wedge X \neq \emptyset \rightarrow (\exists y \in X)(\forall z \in X)(\langle z, y\rangle \notin R).$$

The formula ϕ is absolute for $\mathbf{M}$ by standard arguments. Thus, it is sufficient to check that $(\forall X \in \mathbf{M})(\phi(X, A, R))$. This holds, because R well-ordered A. In fact, we have the stronger statement $(\forall X)\phi(X, A, R)$. $\square_{6.2.14}$

$\square_{6.2.13}$

Lemma 6.2.14 shows that universal quantification of an absolute formula relativizes *downward* from $\mathbf{V}$ to $\mathbf{M}$. However, it may not relativize *upward.* For example, well-ordering is however absolute if we assume Foundation. Other important notions, such as being a cardinal, do not relativize upward.

By Lemma 6.1.13 and Theorem 6.2.13 we have the following relative consistency result:

Corollary 6.2.15. $\mathrm{Con}(ZF^-) \to \mathrm{Con}(ZFC - \mathit{Infinity} + \neg\mathit{Infinity})$.

The next four results conclude our discussion of the class **WF**. We will then assume the Axiom of Foundation, and our axiomatic system will be either ZF or ZFC.

Theorem 6.2.16 (ZF$^-$)**.** *All of the axioms of ZF are true in* ***WF****.*

Proof. This follows by Lemmas 6.1.13 and 6.2.12 $\square_{6.2.16}$

Let us now look at the Axiom of Choice. The next lemma will not have any other applications.

Lemma 6.2.17. *Let $A \in$ **WF**. Then A can be well-ordered if and only if $(A$ can be well ordered$)^{\mathbf{WF}}$.*

Proof. Assume first that A can be well-ordered and that $R \subseteq A \times A$ well-orders A. Since $A \in \mathbf{WF}$, $A \times A \in \mathbf{WF}$, and consequently $R \in \mathbf{WF}$. Now, by Lemma 6.2.14, $(R$ well-orders $A)^{\mathbf{WF}}$, hence $(A$ can be well ordered$)^{\mathbf{WF}}$.

For the opposite implication, if $(A$ can be well ordered$)^{\mathbf{WF}}$, then fix $R \in \mathbf{WF}$ such that $(R$ well-orders $A)^{\mathbf{WF}}$. Then, as in the proof of Lemma 6.2.14, R linearly orders A and every non-empty subset of A *that is in* **WF** has an R-minimal element. However, every subset of A is actually in **WF** (5.1.10), thus R well-orders A. $\square_{6.2.17}$

Corollary 6.2.18 (ZF$^-$)**.** $AC \to (AC)^{\mathbf{WF}}$.

The converse of Corollary 6.2.18 need not hold, since it is consistent that every well-founded set can be well-ordered but certain non-well-founded sets cannot be.

Since in ZF$^-$ one can prove that **WF** is a model of ZF and in ZFC$^-$ once can prove that **WF** is a model of ZFC, we have the following corollary:

Corollary 6.2.19.

$$\mathrm{Con}(ZF^-) \to \mathrm{Con}(ZF)$$

and

$$\mathrm{Con}(ZFC^-) \to \mathrm{Con}(ZFC)$$

Corollary 6.2.19 fives a formal justification for the adoption of the Axiom of Foundation. The assumption of this axiom is a great convenience from a technical point of view since it allows us to establish the absoluteness of many more notions.

———— HERE ENDED WINTER 2006 LECTURE 9 ————

6.2.3 Absoluteness assuming Foundation

Theorem 6.2.20. *The following relations and functions were defined in ZF − P using formulas that are equivalent in ZF − P to Δ_0-formulas. They are thus absolute for all transitive models of ZF − P.*

1. x is an ordinal	*6.* 0
2. x is a limit ordinal	*7.* 1
3. x is a successor ordinal	*8.* 2
4. x is a finite ordinal	*9.* 3
5. ω	......

Proof.

1. Recall that assuming ZF − P, x is an ordinal iff x is transitive and linearly ordered by $\in$. Further, recall that the statement "x is transitive" is equivalent to a Δ_0-formula by Theorem 6.2.9. The statement "x is linearly ordered by $\in$" is expressed via quantification over the elements of x:

$$(\forall y \in x)(\forall z \in x)(y \in z \vee y = z \vee z \in y) \wedge \text{ etc. } \dots,$$

which is also a Δ_0-formula.

2. By definition, "x is a limit ordinal" iff "x is an ordinal and $(\forall y \in x)(\exists z \in x)(y \in z)$ and $(x \neq \emptyset)$". The first two parts of the latter statement are clearly Δ_0-formulas, the third part is the negation of the Δ_0-formula $x = \emptyset$ (Theorem 6.2.9).

3. Again, by definition "x is a successor ordinal" iff "x is an ordinal and x is not a limit ordinal and $(x \neq \emptyset)$". All of these have been established as Δ_0-formulas.

4. By definition, "x is a finite ordinal" iff "x is a successor ordinal and $(\forall y \in x)(y$ is a successor ordinal). These are clearly Δ_0-formulas. Note that this example 4 says that the predicate $x \in \omega$ is expressible using Δ_0-formulas. Compare this with Example 5 below.

5. Here we wish to show that $x = \omega$ is expressible using Δ_0-formulas. Notice that "$x = \omega$" iff "x is a limit ordinal and $(\forall y \in x)\neg(y$ is a limit ordinal), the latter are clearly Δ_0-formulas.

6. This was shown in the proof of Theorem 6.2.9

7-etc. Notice that the formula $y = S(x)$ is, by Theorem 6.2.9, a Δ_0 formula. Furthermore,

$$x = 1 \text{ iff } (\exists y \in x)(y = 0 \wedge S(y) = x),$$
$$x = 2 \text{ iff } (\exists y \in x)(y = 1 \wedge S(y) = x),$$
$$\dots\dots\dots\dots$$
$$x = 200 \text{ iff } (\exists y \in x)(y = 199 \wedge S(y) = x),$$
$$\dots\dots\dots\dots$$

Finally, we draw attention to the fact that **M** satisfies the Axiom of Infinity was only used in the proof of 5 where we needed the existence of ω. $\square_{6.2.20}$

Lemma 6.2.21. *If* **M** *is a transitive model of ZF* $-$ *P, then every finite subset of* **M** *is an element of* **M**.

Proof. We show by induction on n that

$$(\forall x \subseteq \mathbf{M})(|x| = n \to x \in \mathbf{M}).$$

For $n = 0$, this is just the absoluteness of $\emptyset$.

For the inductive step, let us assume that we have shown the above for some n. Let $x \subseteq \mathbf{M}$ have $n+1$ elements. Fix $y \in x$. Then, $y \in \mathbf{M}$, and $(x \setminus \{y\}) \subseteq \mathbf{M}$ has n elements, whence by the inductive hypothesis $(x \setminus \{y\}) \in \mathbf{M}$. We now apply Theorem 6.2.9, since $x = \{y\} \cup (x \setminus \{y\})$. Thus $x \in \mathbf{M}$. $\square_{6.2.21}$

Theorem 6.2.22. *The following notions are absolute for transitive models* **M** *of ZF* $-$ *P.*

1. *x is finite;*
2. *A^n;*
3. *$A^{<\omega}$ $(= \bigcup\{A^n : n \in \omega\})$.*

Proof.

1. By Theorems 6.2.9 and 6.2.20, we have that, assuming ZF $-$ P, "x is finite" iff $(\exists f)\, \phi(x, f)$, where $\phi(x, f)$ states that

$$(f \text{ is a function}) \wedge (\mathrm{dom}(f) = x) \wedge (\mathrm{rng}(f) \subseteq \omega) \wedge (f \text{ is 1-1}),$$

is absolute. Therefore, it suffices to show that for $x \in \mathbf{M}$,

$$(\exists f \in \mathbf{M})\phi(x, f) \Leftrightarrow (\exists f)\, \phi(x, f).$$

The implication $\Rightarrow$ is obvious.

The implication $\Leftarrow$ follows from the fact that for $x \in \mathbf{M}$, we have that $\phi(x, f) \to (f \in \mathbf{M})$. To see this, notice that $\phi(x, f)$ implies that f is a finite set of ordered pairs of elements of **M**. Recall that **M** is closed under pairing by the absoluteness of pairing. By Lemma 6.2.21, **M** is closed under finite subsets, and so $f \in \mathbf{M}$.

2. & 3 To prove the last two cases, notice that we can look at A^n as a function of two variable s $F(A, x)$, where $F(A, x) = \emptyset$ when $x \notin \omega$. Then we can define $A^{<\omega}$ as a function of one variable $G(A)$. As was explained in the discussion surrounding Definition 3.4.5, the above functions are defined in ZF $-$ P.

We concentrate on the proof of 2. The proof of 3 is similar. We need to check that for $A, x \in \mathbf{M}$, we have $F(A, x) = F^{\mathbf{M}}(A, x)$. From the absoluteness of ω, we see that $F^{\mathbf{M}}(A, x) = \emptyset$, unless $x \in \omega$. The absoluteness of notions involving functions and $n \in \omega$ implies that

$$F^{\mathbf{M}}(A, n) = \{f \in \mathbf{M} : (f \text{ is a function}) \wedge (\mathrm{dom}(f) = n) \wedge (\mathrm{rng}(f) \subseteq A),$$

which equals $F(A, n)$ as in part 1.

$\square_{6.2.22}$

Theorem 6.2.23. *The following notions are absolute for transitive models* $\mathbf{M}$ *of ZF − P.*

1. *R well orders A;*
2. $\mathrm{type}(A, R)$*;*

Proof.

1. It suffices to show that if $A, R \in \mathbf{M}$, then

   $$(R \text{ well-orders } A)^{\mathbf{M}} \rightarrow (R \text{ well-orders } A),$$

   since the opposite implication was demonstrated in Lemma 6.2.14.

   Recall that Theorem 3.2.6, which stated that every well-ordering is isomorphic to an ordinal, is a theorem in ZF−P. Therefore, if $(R \text{ well-orders } A)^{\mathbf{M}}$, then there exist $f, \alpha \in \mathbf{M}$ such that

   $$((\alpha \text{ is an ordinal}) \wedge (f : \langle A, R\rangle \longrightarrow \alpha \text{ is an isomorphism}))^{\mathbf{M}}.$$

   However, by Theorem 6.2.20 and absoluteness, this above formula is absolute for $\mathbf{M}$. Hence, α is genuinely an ordinal, and f is genuinely an isomorphism that is, as far as $\mathbf{V}$ is concerned. Since α is well-ordered by $\in$, A is well ordered by R (with order type α).

2. A similar argument to that above shows the absoluteness of $\mathrm{type}(A, R)$.

$\square_{6.2.23}$

Most of arithmetic is absolute. For example:

Theorem 6.2.24. *The following notions are absolute for transitive models* $\mathbf{M}$ *of ZF − P.*

1. $\alpha + 1$*;*
2. $\alpha - 1$*;*
3. $\alpha \cdot \beta$*.*
4. $\alpha + \beta$*;*

Proof.

1. $\alpha + 1$ is simply $S(\alpha)$, the absoluteness of which we have already shown.
2. $x = \alpha - 1$ is equivalent to the statement

   $$(\alpha \text{ is a successor ordinal } \wedge S(\alpha)) \vee (\alpha \text{ is not a successor ordinal } \wedge \alpha = x),$$

   which is clearly absolute.

3. Recall the definition that $\alpha \cdot \beta$ is equal to $\mathrm{type}(\beta \times \alpha, R)$, where R is the lexicographic ordering on $\beta \times \alpha$. This is absolute by standard arguments.

4. This proof is similar to that of 3.

$\square_{6.2.24}$

If instead we think consider ordinal addition $+$ and multiplication $\cdot$ to be defined by transfinite recursion, then their absoluteness can be proven using a general result about the absoluteness of notions defined by transfinite recursion.

Because our theorem (Theorem 5.4.6) about transfinitely recursive definitions were formulated in the language of classes, we need to think about what relativization and absoluteness really mean for classes.

Formally, a class $\mathbf{A}$ is a formula $\mathbf{A}(x)$, but intuitively, we think about it as $\mathbf{A} = \{x : \mathbf{A}(x)\}$. Thus, by $\mathbf{A}^{\mathbf{M}}$, we mean $\{x \in \mathbf{M} : \mathbf{A}^{\mathbf{M}}(x)\}$. Thus we can say that $\mathbf{A}$ is absolute for the class $\mathbf{M}$ iff $\mathbf{A}^{\mathbf{M}} = \mathbf{A} \cap \mathbf{M}$.

For example, $\mathbf{V}(x)$ is the formula $x = x$, which is always absolute, and $\mathbf{V}^{\mathbf{M}} = \mathbf{M}$. Similarly, Theorem 6.2.20.1 can be stated: if $\mathbf{M}$ is a transitive model of ZF $-$ P, then $\mathbf{ON}^{\mathbf{M}} = \mathbf{ON} \cap \mathbf{M}$.

Classes which are relations of more than one variable are treated similarly. Thus, if $\mathbf{R} \subseteq \mathbf{V} \times \mathbf{V}$ (by which we mean that $\mathbf{R}(x, y)$ is a formula and we have in mind $\mathbf{R} = \{\langle x, y\rangle : \mathbf{R}(x, y)\}$), then both $\mathbf{R}^{\mathbf{M}} = \{\langle x, y\rangle \in \mathbf{M} \times \mathbf{M} : \mathbf{R}^{\mathbf{M}}(x, y)\}$ and $\mathbf{R}$ are absolute for $\mathbf{M}$ iff $\mathbf{R}^{\mathbf{M}} = \mathbf{R} \cap (\mathbf{M} \times \mathbf{M})$.

Now, we turn our attention to the relativization of classes which are functions. As usual, this is a bit more fiddly than the relational case. Let $\mathbf{G} : \mathbf{V} \longrightarrow \mathbf{V}$ – that is $\mathbf{G}(x, y)$ is a formula and $(\forall x \exists!\, y \mathbf{G}(x, y))$. We could approach this in two ways: either treat the function as a collection of pairs, and therefore simply as a relation and possibly lose some of the functional nature upon relativization, or we can treat the function as a function.

More specifically, if we treat $\mathbf{G}$ as a relation, and as such as a collection of ordered pairs $\{\langle x, y\rangle : \mathbf{G}(x, y)\}$, then we can do as before and treat the relativization of $\mathbf{G}$ as the relativization of a relation. However, the absoluteness of $\mathbf{G}$ as a relation would require only that $\mathbf{G}^{\mathbf{M}} = \mathbf{G} \cap (\mathbf{M} \times \mathbf{M})$, and would *not* require that, for example, $\mathrm{dom}(\mathbf{G}^{\mathbf{M}}) = \mathbf{M}$.

On the other hand, if we treat $\mathbf{G}$ as a function that will remain a function when relativized, we have to be a bit careful. Note that if the use the functional notation $\mathbf{G}(x)$ is equivalent to using the formula $\mathbf{G}(x, y)$ to introduce a defined operation. Thus we have to follow the conventions given in the discussion given after Lemma 6.1.9. Thus, we can only talk about a *function* $\mathbf{G}^{\mathbf{M}}$ if $(\forall x \exists!\, y \mathbf{G}(x, y))^{\mathbf{M}}$ holds. In this case, $\mathbf{G}^{\mathbf{M}} : \mathbf{M} \longrightarrow \mathbf{M}$, and we say that $\mathbf{G}$ is absolute for $\mathbf{M}$ iff $\mathbf{G}^{\mathbf{M}} = \mathbf{G} \restriction \mathbf{M}$. Thus, it must be clear (from the context) whether we are looking at $\mathbf{G}$ as a relation or as a function.

Theorem 6.2.25 (Absoluteness for recursive classes)**.** *Let $\mathbf{R}$ be a relation which is well-founded and set-like on $\mathbf{A}$. Let $\mathbf{F} : \mathbf{A} \times \mathbf{V} \longrightarrow \mathbf{V}$. Let $\mathbf{G} : \mathbf{A} \longrightarrow \mathbf{V}$ be defined so that*

$$(\forall x \in \mathbf{A})(\mathbf{G}(x) = \mathbf{F}(x, \mathbf{G} \restriction \mathrm{pred}(\mathbf{A}, x, \mathbf{R}))).$$

Note that $\mathbf{G}$ is the function given by Theorem 5.4.6 on Transfinite Recursion on Well-Founded Relations. Let $\mathbf{M}$ be a transitive model of ZF $-$ P. Assume further that

1. *$\mathbf{F}$ is absolute for $\mathbf{M}$as a function;*

2. $\mathbf{R}$ *and* $\mathbf{A}$ *are absolute for* $\mathbf{M}$*;*

3. $(\mathbf{R}$ *is set-like*$)^{\mathbf{M}}$*; and*

4. $(\forall x \in \mathbf{M})(\mathrm{pred}(\mathbf{A}, x, \mathbf{R}) \subseteq \mathbf{M})$*.*

Then, $\mathbf{G}$ *is absolute for* $\mathbf{M}$*.*

Proof. First, notice that ($\mathbf{R}$ is well-founded on $\mathbf{A})^{\mathbf{M}}$ because $\mathbf{R}^{\mathbf{M}} = \mathbf{R} \cap (\mathbf{M} \times \mathbf{M})$ is well-founded on $\mathbf{A}^{\mathbf{M}} = \mathbf{A} \cap \mathbf{M}$. And thus, every non-empty subset of $\mathbf{A}^{\mathbf{M}}$ in $\mathbf{M}$ has an $\mathbf{R}^{\mathbf{M}}$-minimal element. We can thus utilize transfinite recursion inside $\mathbf{M}$ to define $\mathbf{G}^{\mathbf{M}} : \mathbf{A}^{\mathbf{M}} \longrightarrow \mathbf{M}$ such that

$$(\forall x \in \mathbf{A}^{\mathbf{M}})(\mathbf{G}^{\mathbf{M}}(x) = \mathbf{F}^{\mathbf{M}}(x, \mathbf{G}^{\mathbf{M}} \restriction \mathrm{pred}(\mathbf{A}^{\mathbf{M}}, x, \mathbf{R}^{\mathbf{M}}))).$$

But then, by transfinite induction, we have $\mathbf{G}^{\mathbf{M}} = \mathbf{G} \restriction \mathbf{A}^{\mathbf{M}}$. This is because an $\mathbf{R}$-minimal element of $\{x \in \mathbf{A}^{\mathbf{M}} : \mathbf{G}^{\mathbf{M}}(x) \neq \mathbf{G}\}$ would, thanks to our absoluteness statements, lead to a contradiction. $\square_{6.2.25}$

The most important applications of this Theorem 6.2.25 are when $\mathbf{R}$ is $\in$ and $\mathbf{A}$ is either $\mathbf{V}$ or $\mathbf{ON}$. Then, assumptions 2,3, and 4 are easy to check.

Theorem 6.2.26. *The following notions are absolute for every transitive model of ZF* $-$ *P:*

1. α^{β} *(ordinal exponentiation);*

2. $\mathrm{rank}(x)$ *(*$= \mathrm{rank}(x, \mathbf{V}, \in)$*);*

3. $\mathrm{trcl}(x)$*.*

Proof.

1. Recall that α^{β} is defined by transfinite recursion on β.

2. Similarly, $\mathrm{rank}(x)$ is defined by transfinite recursion on x.

3. First, define $\bigcup^{n}(x)$ by recursion on n:

$$\bigcup{}^{y}(x) = \begin{cases} 0 & \text{if } y \notin \omega, \\ x & \text{if } y = \emptyset, \\ \bigcup(\bigcup^{y-1} x) & \text{if } 0 \in y \in \omega. \end{cases}$$

   Then, $\bigcup^{y}(x)$ is an absolute function of variables y and x, thus $\mathrm{trcl}(x) = \bigcup\{\bigcup^{n}(x) : n \in \omega\}$ is absolute.

$\square_{6.2.26}$

It is vital that in Theorem 6.2.26 we can think of $\mathrm{rank}(x)$ as a function defined recursively. Recall that our original definition was in terms of R_{α}. However, if $\mathbf{M}$ does not satisfy the Power Set Axiom, then $R_{\alpha}^{\mathbf{M}}$ are not defined. We have shown earlier that under the Power Set Axiom, the recursive definition and the von Neumann hierarchical definitions are equivalent.

If $\mathbf{M}$ satisfies the Power Set Axiom, then $\mathscr{P}(\cdot)^{\mathbf{M}}$ and $R_{\alpha}^{\mathbf{M}}$ are defined, but in general are not absolute.

Lemma 6.2.27. *Let* $\mathbf{M}$ *be a transitive model of ZF. Then*

1. $\mathscr{P}(x)^{\mathbf{M}} = \mathscr{P}(x) \cap \mathbf{M}$ *for* $x \in \mathbf{M}$*;*

2. $R_\alpha^{\mathbf{M}} = R_\alpha \cap \mathbf{M}$ *for* $\alpha \in \mathbf{M}$.

Proof.

1. This results from the absoluteness of $\subseteq$.

2. this results from the absoluteness of the rank function and from the fact that $R_\alpha = \{x : \text{rank}(x) < \alpha\}$.

$\square_{6.2.27}$

——————— HERE ENDED WINTER 2007 LECTURE 9 ———————

6.3 Sets hereditarily of cardinality $< \kappa$.

We turn to an important method of construction, in ZFC, of transitive models of ZFC $-$ P.

Definition 6.3.1. For every infinite cardinal κ, let

$$H_\kappa = \{x : |\operatorname{trcl}(x)| < \kappa\}.$$

The Axiom of Choice is not necessary for this definition, since $|y| < \kappa$ means that y is well-ordered and $|y| < \kappa$, however, Choice is necessary for the development of the properties of H_κ.

The elements of H_κ are called sets *hereditarily of cardinality* $< \kappa$. In particular, H_ω is the family of *hereditarily finite* sets, while H_{ω_1} is the family of *hereditarily countable* sets.

The fact that every H_κ is a set, and not a proper class, results from the following lemma:

Lemma 6.3.2. *For every infinite cardinal* κ*,* $H_\kappa \subseteq R_\kappa$.

Proof. Fix $x \in H_\kappa$. We will show that $\text{rank}(x) < \kappa$.

Let $t = \operatorname{trcl}(x)$ and let $S = \{\text{rank}(y) : y \in t\}$.

Clearly, $S \subseteq \mathbf{ON}$. We will show that S is an ordinal: It is clear that S is a set, and not a proper class. Thus, $S \neq \mathbf{ON}$. Let α be the first ordinal that is not an element of S. Then, $\alpha \subseteq S$. If $\alpha \neq S$, then let β be the first element of S that is larger than α. Fix $y \in t$ such that $\text{rank}(y) = \beta$.Then, since t is transitive, $(\forall z \in y)(\text{rank}(z) \leqslant \alpha)$. However, since there is no element of t having rank α, we thus have $(\forall z \in y)(\text{rank}(z) < \alpha)$. Now, $\text{rank}(y) = \sup\{\text{rank}(x)+1 : z \in y\} \leqslant \alpha$, a contradiction. Thus, $\alpha = S$.

Since $|t| < \kappa$, we have $|\alpha| < \kappa$, whence $x \subseteq t \subseteq R_\alpha$. Thus $\text{rank}(x) \leqslant \alpha < \kappa$. $\square_{6.3.2}$

Note: In most cases, H_κ is a proper subset of R_κ. For example, $\mathscr{P}(\omega) \in R_{\omega_1} \setminus H_{\omega_1}$.

More generally, we have the following fact:

Lemma 6.3.3. *For regular cardinals κ, $H_\kappa = R_\kappa$ iff $\kappa = \omega$ or κ is strongly inaccessible.*

Proof. If $\kappa = \omega$ or κ is strongly inaccessible, then via easy induction on $\alpha < \kappa$, we have that $(\forall \alpha < \kappa)(|R_\alpha| < \kappa)$. Net, if $\operatorname{rank}(x) = \alpha < \kappa$, then $\operatorname{trcl}(x) \in R_\kappa$, hence $|\operatorname{trcl}(x)| < \kappa$. Thus, by Lemma 6.3.2, $R_\kappa = H_\kappa$. $\square_{6.3.3}$

Note: If $\kappa > \omega$ is not strongly inaccessible, then fix $\lambda < \kappa$, such that $2^\lambda \geqslant \kappa$. Then $\mathscr{P}(\lambda) \in R_\kappa \setminus H_\kappa$.

I dunno if there is time for this, but maybe exercise 5 from Kunen, p 147, about H_κ for singular κ??

Now, a few properties of H_κ:

Lemma 6.3.4. *For infinite cardinals κ:*

1. *H_κ is transitive;*
2. *$H_\kappa \cap \boldsymbol{ON} = \kappa$;*
3. *If $x \in H_\kappa$, then $\bigcup x \in H_\kappa$;*
4. *If $x, y \in H_\kappa$, then $\{x, y\} \in H_\kappa$;*
5. *If $x \in H_\kappa$ and $y \subseteq x$, then $y \in H_\kappa$;*
6. *(AC) If κ is regular, then $(\forall x)(x \in H_\kappa) \iff x \subseteq H_\kappa \wedge |x| < \kappa)$.*

Proof.

1. Follows from the fact that $x \in y$ implies $\operatorname{trcl}(x) \subseteq \operatorname{trcl}(y)$.
2. Follows from the fact that $\operatorname{trcl}(\alpha) = \alpha$.

3.-5. Similar.

6. If $x \subseteq H_\kappa$, and $|x| < \kappa$, then since $\operatorname{trcl}(x) = x \cup \bigcup\{\operatorname{trcl}(y) : y \in x\}$, $\operatorname{trcl}(x)$ is the sum of $< \kappa$ sets each of cardinality $< \kappa$. With the assumption of the Axiom of Choice, this means that $|\operatorname{trcl}(x)| < \kappa$.

$\square_{6.3.4}$

Theorem 6.3.5 (ZFC)**.** *If κ is regular and $\kappa > \omega$, then H_κ is a model of $ZFC - P$.*

Proof. That H_κ satisfies Extensionality results from the transitivity of H_κ. Regularity is satisfied in every model (see Lemma 6.1.12). The remaining axioms of $\mathrm{ZF} - \mathrm{P} - \mathrm{Inf}$ are checked just as in the proofs for R_ω $(= H_\omega)$ and **WF** using Lemma 6.3.4. In particular, Lemma 6.3.4.3 gives the Union Axiom; 4. , the Pairing Axiom; 5. gives the Separation Axiom; 6. gives Replacement.

Next, since H_κ is a model of $\mathrm{ZF} - \mathrm{P} - \mathrm{Inf}$, from Lemma 6.3.4.2 we have that $\omega \in H_\kappa$, and therefore the Axiom of Infinity is satisfied by H_κ.

Finally, to show that the Axiom of Choice holds in H_κ, it suffices to check that

$$(\forall A \in H_\kappa)(\exists R \in H_\kappa)(R \text{ well-orders } A),$$

because well-ordering is absolute for H_κ (see Theorem 6.2.23). To do this, fix $A \in H_\kappa$, and let (using the Axiom of Choice in $\mathbf{V}$) $R \subseteq A \times A$ be a well-ordering of A. Then, $R \subseteq H_\kappa$ by Lemma 6.3.4.4, and thus $R \in H_\kappa$ by 6.3.4.6. $\square_{6.3.5}$

Theorem 6.3.6 (ZFC). *If κ is regular and $\kappa > \omega$, then the following are equivalent:*

1. *H_κ is a model of ZFC;*
2. *$H_\kappa = R_\kappa$;*
3. *κ is strongly inaccessible.*

Proof. That $2 \Leftrightarrow 3$ is contained in Lemma 6.3.3. For the equivalence of 1, notice that H_κ satisfies the Power Set Axiom iff $(\forall x \in H_\kappa)(\exists y \in H_\kappa)(\forall z \in H_\kappa)(z \subseteq x \to z \in y)$. Since $z \subseteq x \in H_\kappa$ implies that $z \in H_\kappa$ and H_κ satisfies Separation, the Power Set Axiom is satisfied by H_κ iff $(\forall x \in H_\kappa)(\mathscr{P}(x) \in H_\kappa)$. This is true if $H_\kappa = R_\kappa$, but false if for some $\lambda < \kappa$, we have that $2^\lambda \geqslant \kappa$, because then we would have that $\lambda \in H_\kappa$, but $\mathscr{P}(\lambda) \notin H_\kappa$. $\square_{6.3.6}$

——————— HERE ENDED WINTER 2006 LECTURE 10 ———————

In particular, if we take κ to be some cardinal that is not strongly inaccessible, we have that

$$\mathrm{Con}(ZFC) \to \mathrm{Con}(ZFC - P + \neg(P)).$$

And so, the Power Set Axiom cannot be proven from the other axioms of ZFC. Indeed, if we take $\kappa = \omega_1$, then we get something even stronger:

Corollary 6.3.7. $\mathrm{Con}(ZFC) \to \mathrm{Con}(ZFC - P + (\forall x)(x \textit{ is countable}))$.

Proof. We define H_{ω_1} in ZFC, so therefore it is a model of ZFC $-$ P. If $x \in H_{\omega_1}$, then x is countable, and any function from ω onto x is also in H_{ω_1}. Thus, (x is countable) is satisfied in H_{ω_1}. $\square_{6.3.7}$

If κ is strongly inaccessible, then basic cardinal arithmetic is absolute in H_κ. Let us make note also of the following fact:

Lemma 6.3.8 (ZFC). *Let κ be strongly inaccessible. Then the statement "α is strongly inaccessible" is absolute for H_κ.*

In particular, if κ is the first strongly inaccessible cardinal, then H_κ is a model of ZFC in which there are no strongly inaccessible cardinals.

Corollary 6.3.9.

$$\mathrm{Con}(ZFC) \to$$
$$\to \mathrm{Con}(ZFC + (\neg\exists\alpha)(\alpha \textit{ is strongly inaccessible})).$$

Proof. Let strginacc(κ) be an abbreviation of the statement "κ is strongly inaccessible". Formally, working in ZFC, one cannot prove that $\exists\kappa$ strginacc(κ) (thanks to our previous Corollary), and so, we cannot define the least such κ. Instead, we define

$$\mathbf{M} = \{x : \forall\kappa\, \mathrm{strginacc}(\kappa) \to x \in H_\kappa\}.$$

Thus, in ZFC one cannot prove if $\mathbf{M} = \mathbf{V}$ or if $\mathbf{M} = H_\kappa$. for the smallest inaccessible κ. However, in both of these cases ZFC + $(\neg\exists\alpha)(\alpha$ is strongly inaccessible holds in $\mathbf{M}$. $\square_{6.3.9}$

Note that, working in (only) ZFC, one cannot produce a model for ZFC + $(\exists\alpha)(\alpha$ is strongly inaccessible)!

6.4 Reflection Theorems

We now look at a general procedure for attempts to build *sets* that are models of ZFC. For those that know model theory: this is an application of the Downward Löwenheim-Skolem Theorem to **V**. While the Downward Löwenheim-Skolem Theorem says that every model of a first-order theory has a small elementary submodel, the Reflection Theorems say that for every *finite* number of formulas in the (first-order) language of set theory, there is a set M that is something akin to an elementary submodel of **V** with respect to the given formulas. If Choice is assumed, then one can find a countable such model.

The theorems presented in this section are often given together in one theorem, and called the *Reflection Principle*.

Definition 6.4.1. We say that a list of formulas $\phi_1, \ldots, \phi_n$ is *subformula closed* iff every subformula of a formula on this list is also on this list.

Since every formula has only finitely many subformulas, every finite list of formulas can be extended to a finite subformula closed list. The following will be known to those familiar with model theory as the *Tarski-Vaught criterion*.

——————— HERE ENDED WINTER 2007 LECTURE 10 (thanks Heike!!)———————

Lemma 6.4.2. *Let* $\mathbf{M}$ *and* $\mathbf{N}$ *be classes and* $\mathbf{M} \subseteq \mathbf{N}$*. Let* $\phi_1, \ldots, \phi_n$ *be a subformula closed list. The following are equivalent:*

1. $\phi_1, \ldots, \phi_n$ *are absolute for* $\mathbf{M}$ *and* $\mathbf{N}$*.*
2. *if* ϕ_i *is of the form* $(\exists x)(\phi_j(x, y_1, \ldots, y_m))$ *(where all free variables are listed), then*

$$(\forall y_1, \ldots, y_m \in \mathbf{M})((\exists x \in \mathbf{N})(\phi_j^{\mathbf{N}}(x, y_1, \ldots, y_m)) \rightarrow (\exists x \in \mathbf{M})(\phi_j^{\mathbf{N}}(x, y_1, \ldots, y_m))). \quad (6.2)$$

Proof.

(1⇒2) Fix $y_1, \ldots, y_m \in \mathbf{M}$ and assume that $(\exists x \in \mathbf{N})(\phi_j^{\mathbf{N}}(x, y_1, \ldots, y_m)$. Then (rewriting), $\phi_i^{\mathbf{N}}(x, y_1, \ldots, y_m)$, and so by absoluteness of ϕ_i, $\phi_i^{\mathbf{M}}(x, y_1, \ldots, y_m)$. In other terms, $(\exists x \in \mathbf{M})(\phi_j^{\mathbf{M}}(x, y_1, \ldots, y_m)$. By absoluteness of ϕ_j, we have $(\exists x \in \mathbf{M})(\phi_j^{\mathbf{N}}(x, y_1, \ldots, y_m)$.

(2⇒1) We check, by induction on the complexity of the formula ϕ_j, that ϕ_j is absolute for $\mathbf{M}$ and $\mathbf{N}$. Assume that all subformulas of ϕ_i are absolute. If ϕ_i is atomic, then its absoluteness is clear. The absoluteness of $\phi_i = \phi_j \wedge \phi_k$ or $\phi_i = \neg\phi_j$ follows from the inductive assumption. Now, let us assume that ϕ_i is $(\exists x)(\phi_j^{\mathbf{M}}(x, y_1, \ldots, y_m)$, and fix $y_1, \ldots, y_m \in \mathbf{M}$. Then

$$\begin{aligned}
\phi_i^{\mathbf{M}}(x, y_1, \ldots, y_m) &\iff (\exists x \in \mathbf{M})(\phi_j^{\mathbf{M}}(x, y_1, \ldots, y_m) \\
&\iff (\exists x \in \mathbf{M})(\phi_j^{\mathbf{N}}(x, y_1, \ldots, y_m) \\
&\iff (\exists x \in \mathbf{N})(\phi_j^{\mathbf{N}}(x, y_1, \ldots, y_m) \\
&\iff \phi_i^{\mathbf{N}}(x, y_1, \ldots, y_m)
\end{aligned}$$

The first and last equivalences are applications of the definition of relativization. The second equivalence uses the absoluteness of ϕ_j, the third uses assumption 2.

$\square_{6.4.2}$

The usefulness of Lemma 6.4.2 relies on the fact that statement *2* involves only truth of formulas with respect to the larger class, **N**. Thus, statement *2* can be considered a kind of closure requirement for **M**.

In our first application, we will take **N** to be **V**, and we will try to find a set $\mathbf{M} = R_\alpha$ such that $\phi_1, \ldots, \phi_n$ are absolute for **M**

Theorem 6.4.3. *[Reflection Theorem] For given formulas $\phi_1, \ldots, \phi_n$,*

$$ZF \vdash (\forall\alpha)(\exists\beta > \alpha)(\phi_1, \ldots, \phi_n \textit{ are absolute for } R_\beta).$$

The proof of this theorem uses very little of the particular structure of R_α. To emphasize this, we will instead prove a much more general theorem, a special case of which will be Theorem 6.4.3 for $\mathbf{Z} = \mathbf{V}$ and $Z_\alpha = R_\alpha$.

Theorem 6.4.4. *Suppose **Z** is a class, and for every α, the set Z_α has the following properties:*

1. $\alpha < \beta \rightarrow Z_\alpha \subseteq Z_\beta$;
2. *if γ is a limit ordinal, then $Z_\gamma = \{Z_\alpha : \alpha < \gamma\}$;*
3. $\mathbf{Z} = \{Z_\alpha : \alpha \in \boldsymbol{ON}\}$.

Then, for arbitrary formulas $\phi_1, \ldots, \phi_n$,

$$(\forall\alpha)(\exists\beta > \alpha)(\phi_1, \ldots, \phi_n \textit{ are absolute for } Z_\beta \textit{ and } \mathbf{Z}).$$

Proof. We apply Lemma 6.4.2 for $\mathbf{N} = \mathbf{Z}$, and we try to find $\mathbf{M} = Z_\beta$ by condition 2 of that lemma. We can assume that $\phi_1, \ldots, \phi_n$ is subformula closed, for if it is not, we can extend it to a finite list that is.

For every $i = 1, \ldots n$, we define a function $\mathbf{F_i} : \mathbf{ON} \longrightarrow \mathbf{ON}$ as follows:

- If ϕ_i is not an existential quantification, set $\mathbf{F}_i(\xi) = 0$.
- If ϕ_i is of the form $\exists x\, \phi_j(x, y_1, \ldots, y_m)$, then let $\mathbf{G}_i(y_1, \ldots, y_m)$ be defined as follows:
  - $\mathbf{G}_i(y_1, \ldots, y_m) = 0$ if $\neg\exists x \in \mathbf{Z}\, \phi_j^{\mathbf{Z}}(x, y_1, \ldots, y_m)$;
  - $\mathbf{G}_i(y_1, \ldots, y_m)$ is the smallest μ such that $\exists x \in Z_\mu\, \phi_j^{\mathbf{Z}}(x, y_1, \ldots, y_m)$ if $\exists x \in \mathbf{Z}\, \phi_j^{\mathbf{Z}}(x, y_1, \ldots, y_m)$;

  Then, let

  $$\mathbf{F}_i(\xi) = \sup\{\mathbf{G}_i(y_1, \ldots, y_m) : y_1, \ldots, y_m \in Z_\xi\}.$$

  This supremum exists by the Replacement axiom.

By Lemma 6.4.2, if β is a limit ordinal, and if for each i, $\forall\xi < \beta\, (\mathbf{F}_i(\xi) < \beta)$, then $\phi_1, \ldots, \phi_n$ is absolute for every Z_β and **Z**. We fix α, and show that we can always find such a $\beta > \alpha$.

Let $\beta_0 = \alpha$, and let $\beta_{p+1} = \max\{\beta_p + 1, \mathbf{F}_1(\beta_p), \ldots, \mathbf{F}_n(\beta_p)\}$.. This is a good recursive definition of β_p for $p \in \omega$. Let $\beta = \{\beta_p : p \in \omega\}$. Because $\langle\beta_p\rangle$ is an increasing sequence, β is a limit ordinal $> \alpha$. Notice further that if $\xi > \xi'$, then $\mathbf{F}_i(\xi) \leqslant \mathbf{F}_i(\xi')$. Thus, if $\xi < \beta$, then $\xi < \beta_p$ for some p. This means that

$$\mathbf{F}_i(\xi) \leqslant \mathbf{F}_i(\beta_p) \leqslant \beta_{p+1} < \beta,$$

which completes the proof. $\square_{6.4.4}$

If, in Theorem 6.4.4, we take each ϕ_i to be a sentence, the we have

$$ZF \vdash (\forall \alpha)(\exists \beta > \alpha)\left(\bigwedge_{i=1}^{n}(\phi_i^{R_\beta} \iff \phi_i)\right).$$

In particular, if ϕ_i is an axiom of ZF, the obviously $ZF \vdash \phi_i$, and consequently,

$$ZF \vdash (\forall \alpha)(\exists \beta > \alpha)\left(\bigwedge_{i=1}^{n}(\phi_i^{R_\beta})\right). \tag{6.3}$$

If we want R_β to satisfy the Axiom of Choice, we have to make our arguments in ZFC so that the Axiom of Choice is true $\mathbf{V}$. However, even in (only) ZF, we can produce a β such that $AC^{R_\beta} \iff AC$. More generally, we have the following:

Corollary 6.4.5. *Let S be any set of axioms that contains ZF. Let $\phi_1, \ldots, \phi_n$ be a finite list of axioms from S. Then,*

$$S \vdash (\forall \alpha)(\exists \beta > \alpha)\left(\bigwedge_{i=1}^{n}(\phi_i^{R_\beta})\right).$$

Proof. Apply the previous theorem (specifically, equation 6.3), and the fact that $S \vdash \phi_i$ for each i. $\square_{6.4.5}$

The above corollary is a purely existential fact– we have no description of the kinds of β for which the hypothesis holds. One of the consequences of the corollary is that neither ZF nor ZFC is finitely axiomatizable.

Corollary 6.4.6. *Let S be any set of axioms that extends ZF, and let $\phi_1, \ldots, \phi_n$ be a finite list of sentences from S. If from $\phi_1, \ldots, \phi_n$ one can prove all axioms of S, then S is inconsistent.*

Proof. Assume that it is possible to prove all the axioms of S from $\phi_1, \ldots, \phi_n$. Let β be the smallest ordinal such that $\bigwedge_{i=1}^{n} \phi_i^{R_\beta}$. Then, all the axioms of S hold in R_β. Because S extends ZF, all of the basic results about absoluteness of formulas hold in R_β. In particular, if we take $\alpha \in R_\beta$, then $R_\alpha^{R_\beta} = R_\alpha \cap R_\beta = R_\alpha$. Thus, the function R_α is absolute for $\alpha \in R_\beta$. Since S proves $\exists \alpha \bigwedge_{i=1}^{n} \phi_i^{R_\alpha}$, this must also hold in R_β, therefore $\exists \alpha < \beta \bigwedge_{i=1}^{n} \phi_i^{R_\alpha}$, contradicting the minimality of β. $\square_{6.4.6}$

Corollary 6.4.6 shows that there cannot exist a finite list of axioms that would be equivalent to all of ZFC. In particular, for given axioms $\phi_1, \ldots, \phi_n$ of ZFC, the first R_β which is a model for $\bigwedge_{i=1}^{n} \phi_i$ is not a model of ZFC. The proof of Corollary 6.4.6 produces a theorem of ZFC, particularly $\exists \alpha \bigwedge_{i=1}^{n} \phi_i^{R_\alpha}$, which is false in R_β.

By a small modification of the proof of Theorem 6.4.3, one can get a countable set A for which a given list of formulas is absolute. Of course, such an A cannot be R_β. Neither can A be transitive, since $\mathscr{P}()$ cannot be absolute for a countable transitive model. Non-transitive models are not useful in themselves, but we can use the Mostowski collapse on them to get the respective transitive models. We state this generally:

Theorem 6.4.7 (AC). *Let* $\mathbf{Z}$ *be a class, and* $\phi_1, \ldots, \phi_n$ *be any formulas. Then,*

$$\forall X \subset \mathbf{Z}\, \exists A\, (X \subset A \subset \mathbf{Z} \wedge (\phi_1, \ldots, \phi_n \text{ are absolute for } A, \mathbf{Z}) \wedge \\ \wedge\, |A| \leqslant \max(\omega, |X|)).$$

Proof. Assume that the list $\phi_1, \ldots, \phi_n$ is subformula closed. Let $Z_\alpha = \mathbf{Z} \cap R_\alpha$. Note that $\mathbf{Z}$ and Z_α satisfy the assumptions of Theorem 6.4.4. Fix α such that $X \subseteq Z_\alpha$. By Theorem 6.4.4, fix $\beta > \alpha$ so that the formulas $\phi_1, \ldots, \phi_n$ are absolute for Z_β and $\mathbf{Z}$ We will build A as a subset of Z_β. By the assumption of the Axiom of Choice, fix a well-ordering $\lessdot$ of Z_β.

We define *Skolem functions* H_i for ϕ_i: By k_i denote the number of free variables $y_1, \ldots, y_{k_i}$ of ϕ_i. Define the function $H_i : Z_\beta^k \longrightarrow Z_\beta$:

- if ϕ is $(\exists x)(\phi_j(x, y_1, \ldots, y_{k_i})$ and $(\exists x \in Z_\beta)(\phi_j^{Z_\beta}(x, y_1, \ldots, y_{k_i})$ holds, then let $H_i(y_1, \ldots, y_{k_i})$ be the $\lessdot$-first such x;
- if $\neg(\exists x \in Z_\beta)(\phi_j^{Z_\beta}(x, y_1, \ldots, y_{k_i})$ or ϕ_i is not an existential formula, then let $H_i(y_1, \ldots, y_{k_i})$ be the $\lessdot$-first element of Z_β;
- if $k_i = 0$, then identify H_i with an element of Z_β.

By Lemma 6.4.2, if A is closed under each H_i, then each ϕ_i will be absolute for A and $\mathbf{Z}$. Hence, we can take A simply as the closure of X under the function $H_1, \ldots, H_n$. The fact that $|A| \leqslant \max(\omega, |X|)$ follows from Theorem 4.3.10 (which was the combinatorial version of the Downward Löwenheim-Skolem Theorem). $\square_{6.4.7}$

The proof of Theorem 6.4.7 is a bit inelegant since it uses the same argument twice: once to get Z_β and again for A. Unfortunately, an approach that starts with a well-ordering of $\mathbf{Z}$ cannot be argued in ZFC since $\mathbf{Z}$ might be a proper class. Even in set theories that allow quantification over classes, one cannot prove from the Axiom of Choice that every proper class can be well-ordered.

We now want to apply the Mostowski collapsing isomorphism to the set A from Theorem 6.4.7. Since isomorphisms preserve all properties, we can apply the following:

Lemma 6.4.8. *Let* G *be an injective map from* A *onto* M *which is an isomorphism for the* $\in$ *relation. Then, for each formula* $\phi(x_1, \ldots, x_n)$*,*

$$\forall x_1, \ldots, x_n \in A)(\phi(x_1, \ldots, x_n)^A \iff \phi(G(x_1), \ldots, G(x_n))^M).$$

Proof. Proceeds by induction on the complexity of ϕ. $\square_{6.4.8}$

In particular, for every sentence, we have $\phi^A \iff \phi^M$.

Corollary 6.4.9 (AC)**.** *Let* $\mathbf{Z}$ *be a transitive class, and* $\phi_1, \ldots, \phi_n$ *be any formulas. Then*

$$(\forall X \subset \mathbf{Z})(X \text{ is transitive } \rightarrow \\ \rightarrow (\exists M)((X \subseteq M) \wedge (\phi_1, \ldots, \phi_n \text{ are absolute for } M \text{ and } \mathbf{Z}) \wedge \\ \wedge (M \text{ is transitive}) \wedge (|M| \leqslant \max(\omega, |X|))).$$

Proof. We can assume that one of the ϕ_i is the Axiom of Extensionality, for if it is not, we can simply add it to the list.

Let A be as in the statement of Theorem 6.4.7. Then, for every i we have $\phi_i^A \iff \phi_i^{\mathbf{Z}}$. Since $\mathbf{Z}$ is a transitive class and the Axiom of Extensionality holds in $\mathbf{Z}$, it holds in A. By the Mostowski Collapsing Theorem, there exists an $\in$-isomorphism G from A onto a transitive set M. To see that $X \subseteq M$, notice that for $x \in X$, we have

$$G(x) = \{G(y) : (y \in A) \wedge (y \in x)\} = \{G(y) : y \in x\},$$

because of the transitivity of X. And so, $G(x) - x$ for $x \in X$ by $\in$-induction on x. $\square_{6.4.9}$

As a special case, we can take $\mathbf{Z} = \mathbf{V}$ and $X = \omega$. Then we get the following:

Corollary 6.4.10. *Let S be any set of axioms extending ZFC and let $\phi_1, \ldots, \phi_n$ be any axioms of S. Then*

$$S \vdash (\exists M)(|M| = \omega \wedge M \text{ is transitive } \wedge \bigwedge_{i=1}^{n} \phi_i^M).$$

In particular, in ZFC we can prove the existence of a countable transitive model M of any given finite fragment of ZFC. By listing enough axioms, we can ensure that all the basic absoluteness results hold for M, and that $\mathscr{P}^M(x)$ and ω_α^M are defined. However, these last two notions are not absolute. For example, ω_1^M is a countable ordinal which M "thinks" is uncountable. This only means that there is no function from ω onto ω_1^M *in* M. Similarly, $\mathscr{P}^M(x) = \mathscr{P}(x) \cap M$ is a countable set, but not as far as those "living inside" M are concerned. The fact that sets from M are truly (i.e. in $\mathbf{V}$) countable, but are uncountable from the point of view of M (i.e. that "countable" is not absolute) is called the *Skolem paradox.*

——————— HERE ENDED WINTER 2006 LECTURE 11 ———————

Chapter 7

Gödel's Constructible Universe L

7.1 The intuition for the notion of constructibility

In this chapter, we will work in ZF and define the class L of constructible sets. This class L is a proper class, is a transitive model of ZFC + GCH, and is the smallest model of ZF that contains all ordinal numbers. It also satisfies many useful combinatorial principles (though we will not discuss these in this lecture).

If we look at the axioms of set theory, we see that they postulate the existence of certain sets. Among these sets are ones that are "well-defined": pairs, unions, a subset separated using a first-order formula, etc.

Other axioms axioms give sets whose constructions are not directly given. Here we have both "individual" examples, and also "collective" examples.

First, let us look at axioms that give "individual" examples sets without explicit construction:

The weakest example of an axiom that gives us such examples is the Axiom of Regularity, which talks about the existence of an $\in$-minimal element. In this case, we simply believe that descending through the $\in$ relation to elements of elements we will eventually get to that $\in$-minimal element after a finite number of steps.

The case of the existence of an element differentiating two different sets (i.e. the Axiom of Extensionality) is more problematic. There is no way of constructing such an element, which in the easiest case – that of finding an element of a non-empty set – leads some mathematicians to the very false conviction that one needs to use here the Axiom of Choice. Intuitively, the Axiom of Choice says only that elements chosen from certain sets also form a set. For this reason, the Axiom of Choice (once we get used to the real role of the Axiom of Extensionality) is more reminiscent of the Axiom of Replacement.

The case of axioms giving "collective" examples is especially bad. Here this applies especially to the Power Set Axiom. That axiom says that for a given set X there exists a set $\mathscr{P}(X)$ composed of *all* subsets of X. However, this axiom does not make precise exactly *how* one is supposed to find these subsets.

This is where a new axiom can be useful – the Axiom of Constructibility. Speaking very simply and imprecisely, we take only those subsets which we can "construct", that is, that we can define using the Replacement Axiom.

So, this gives a natural construction of constructible sets in the following manner. We form a hierarchy of sets L_α, $\alpha \in \mathbf{ON}$ (NOTE THAT THIS IS NOT A FORMAL DEFINITION!!!!!):

$L_0 = \emptyset$;

$L_{\alpha+1} = \{x \subseteq L_\alpha : \text{ for some formula } \phi,\ x = \{u \in L_\alpha : \langle L_\alpha, \in, c\rangle_{c\in L_\alpha} \models \phi(u)\}\}$;

$L_\gamma = \bigcup\{L_\alpha : \alpha < \gamma\}$, for γ a limit;

$L = \bigcup\{L_\alpha : \alpha \in \mathbf{ON}\}$, the constructible universe.

In other words, the constructible subsets of L_α are those that have an individual definition ϕ in the relational system $\langle L_\alpha, \in, c\rangle_{c\in L_\alpha}$. This hierarchy is reminiscent of the von Neumann Hierarchy of R_αs. The only difference is that in the definition of $R_{\alpha+1}$, we take *all* subsets of R_α; for $L_{\alpha+1}$ we take only the "definable" subsets of L_α.

The main shortcoming in the above approach to the L_α hierarchy is the notion of definability, which in the above case is an external notion. One would wish to to "mathematicize" this notion if we are to use it formally. This shortcoming can be rectified using the so-called Gödel operations, which will be discussed in detail in the next section.

Generally speaking, the idea here is to find some "very absolute" operations (in this lecture, there will be 10 of these) $\mathbf{F}_1, \ldots, \mathbf{F}_{10}$. Then, we need to take some function

$$n : \mathbf{ON} \longrightarrow \{1, \ldots, 10\} \times \mathbf{ON} \times \mathbf{ON}.$$

This function ought to also be absolute, and be an onto mapping. It should also have a further property: if $n(\alpha) = \langle i(\alpha), m(\alpha), k(\alpha)\rangle$, then $m(\alpha), k(\alpha) < \alpha$ (this is to ensure we are only building things using components that have already been constructed). Such a function is relatively easy to build using the canonical mapping from $\mathbf{ON}^2$ onto $\mathbf{ON}$. Finally, we can then define the sets

$$F_\alpha = \mathbf{F}_{i(\alpha)}(F_{m(\alpha)}, F_{k(\alpha)}).$$

Constructible sets, (that is, elements of the class L) are then defined as sets of the form F_α. This construction does not help much in forming a good intuition for constructible sets. For example, it is not immediately clear how we can get certain constructible sets (for example, ordinals). On the other hand, certain facts will be well clear – for example, the existence of a well-ordering of the constructible universe L.

In this presentation, we will take a middle route - when passing from L_α to $L_{\alpha+1}$, we will use Gödel functions, not definability.

Finally, a word about the last paper Cohen wrote before his discovery of forcing: In that paper, Cohen built a hierarchy he called the T_α based not on the Axiom of Separation (as in the case of L), but rather on the Axiom of Replacement. In this way, he found a model of ZFC minimal in the sense that

1. if there exists a set $\langle M, \in\rangle$ that is a transitive model of ZF such that $T \subseteq M$, where $T = \bigcup\{T_\alpha : \alpha \in \mathbf{ON}\}$, then T is countable and transitive, and $\langle T, \in\rangle$ is a model of ZFC + $V = L$; or

2. Such a set $\langle M, \in \rangle$ does not exist, and then $L = T = \bigcup\{T_\alpha : \alpha \in \mathbf{ON}\}$.

7.2 Gödel Operations

The Axiom of Separation says that for every formula $\phi(x)$ and every set X, there exists a set $Y = \{u \in X : \phi(u)\}$. It turns out that for Δ_0-formulas, the construction of Y from X can be carried out explicitly using the so-called *Gödel operations*:

$$\begin{aligned}
&\mathbf{F}_1(X,Y) = \{X,Y\};\\
&\mathbf{F}_2(X,Y) = X \times Y;\\
&\mathbf{F}_3(X,Y) = \epsilon(X,Y) = \{(u,v) : u \in X \wedge v \in Y \wedge u \in v\};\\
&\mathbf{F}_4(X,Y) = X \setminus Y;\\
&\mathbf{F}_5(X,Y) = X \cap Y;\\
&\mathbf{F}_6(X,Y) = \bigcap X;\\
&\mathbf{F}_7(X,Y) = \mathrm{dom}(X);\\
&\mathbf{F}_8(X,Y) = \{(u,v) : (v,u) \in X\};\\
&\mathbf{F}_9(X,Y) = \{(u,v,w) : (u,w,v) \in X\};\\
&\mathbf{F}_{10}(X,Y) = \{(u,v,w) : (v,w,u) \in X\}.
\end{aligned}$$

Definition 7.2.1. A class (not necessarily proper) C is called *closed* if it is closed with respect to all of the Gödel operations. That is, $\mathbf{F}_i(x,y) \in C$ for all $x, y \in C$ and $i = 1, \dots, 10$.

Theorem 7.2.2 (Gödel's Normal Form Theorem)**.** *For every Δ_0-formula $\phi(u_1, \dots, u_n)$ there exists such a composition of Gödel operations $\mathbf{F}$ that for every $X_1, \dots, X_n$ we have*

$$\mathbf{F}(X_1, \dots, X_n) = \{(u_1, \dots, u_n) : u \in X_1 \wedge \cdots \wedge u \in X_n \wedge \phi(u_1, \dots, u_n)\}.$$

To make the proof of this theorem a little simpler, we proceed by induction on the complexity of formulas, and also assume that all formulas are of the following form:

Definition 7.2.3. A formula ϕ is *normal* if

1. the only logical symbols in ϕ are $\neg$, $\wedge$, and the bounded quantifier $\exists$;
2. $=$ does not occur in ϕ;
3. the only occurrence of $\in$ is $u_i \in u_j$, where $i \neq j$;
4. the only occurrence of $\exists$ is of the form $(\exists u_{m+1} \in u_i)\psi(u_1, \dots, u_{m+1})$, where $i \leqslant m$.

We can make the assumption that all formulas are normal for the purposes of the proof of the theorem based on the following lemma.

Lemma 7.2.4. *Every Δ_0-formula can be written as a normal formula. More exactly, for every Δ_0-formula there exists a normal formula ϕ' such that $\phi \iff \phi'$ is a theorem of (a very weak) set theory containing Extensionality.*

Proof. First, notice that exclusively on the basis of the laws of logic, we can exchange logical symbols occurring in ϕ with those occurring in requirement 1 of the normal formula definition. Similarly, we can exchange the indices of those variables over which a quantifier acts in ϕ so that the quantifier acts over the variable with the highest index. Finally, the expression $x \in x$ can be replaced with $(\exists u \in x)(u = x)$. All of those changes can be carried out on the logical level.

What remains is the problem of what to do with $=$. Every occurrence of $x = y$ can be exchanged, on the basis of the Axiom of Extensionality (and the appropriate theorems of logic), with the formula

$$(\forall u \in x)(u \in y) \wedge (\forall v \in y)(v \in x).$$

This completes the proof. $\square_{7.2.4}$

Proof. By Lemma 7.2.4, we can assume that ϕ is a normal function. We proceed by induction on the complexity of ϕ, and assume that the hypothesis has been shown true for all subformulas of ϕ. The rest of this proof goes via a terrifyingly long series of cases.

1. Assume $\phi(u_1, \ldots, u_n)$ is an atomic formula. That is, ϕ is $u_i \in u_j$, $(i \neq j)$. We proceed by induction on n.

   (a) Let $n = 2$. Then, we have

   $$\{(u_1, u_2) : u_1 \in X_1 \wedge u_2 \in X_2 \wedge u_1 \in u_2\} = \\ = \epsilon(X_1, X_2) = \mathbf{F}_3(X_1, X_2)$$

   and

   $$\{(u_1, u_2) : u_1 \in X_1 \wedge u_2 \in X_2 \wedge u_2 \in u_1\} = \\ = \mathbf{F}_8(\epsilon(X_1, X_2), X_2) = \mathbf{F}_8(\mathbf{F}_3(X_1, X_2), X_2).$$

   (b) Let $n > 2$ and $i, j \neq n$. By the inductive hypothesis, there exists an operation $\mathbf{F}$ such that

   $$\{(u_1, \ldots, u_{n-1}) : u_1 \in X_1 \wedge \cdots \wedge u_{n-1} \in X_{n-1} \wedge u_i \in u_j\} = \\ = \mathbf{F}(X_1, \ldots, X_{n-1}).$$

   So, we have

   $$\{(u_1, \ldots, u_n) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge u_i \in u_j\} = \\ = \mathbf{F}(X_1, \ldots, X_{n-1}) \times X_n = \\ = \mathbf{F}_2(\mathbf{F}(X_1, \ldots, X_{n-1}), X_n).$$

   (c) Assume $n > 2$ and $i, j \neq n - 1$. Using case 1(b), we have an $\mathbf{F}$ such that

   $$\{(u_1, \ldots, u_{n-2}, u_n, u_{n-1}) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge u_i \in u_j\} = \\ = \mathbf{F}(X_1, \ldots, X_n).$$

Notice further that

$$(u_1, \dots, u_{n-2}, u_n, u_{n-1}) = ((u_1, \dots, u_{n-2}), u_n, u_{n-1}),$$

and so

$$\{(u_1, \dots, u_{n-2}, u_n, u_{n-1}) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge u_i \in u_j\} = \\ = \mathbf{F}_9(\mathbf{F}(X_1, \dots, X_n), X_1).$$

(d) Let $i = n-1$, $j = n$. Then, using case 1(a), we have

$$\{(u_{n-1}, u_n) : u_{n-1} \in X_{n-1} \wedge u_n \in X_n \wedge u_{n-1} \in u_n\} = \\ = \epsilon(X_{n-1}, X_n) = \mathbf{F}_3(X_{n-1}, X_n).$$

And so

$$\{((u_{n-1}, u_n), (u_1, \dots, u_{n-2})) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge u_{n-1} \in u_n\} = \\ = \epsilon(X_{n-1}, X_n) \times (X_1 \times \cdots \times X_{n-2}) = \\ = \mathbf{F}(X_1, \dots, X_n).$$

Next, notice that

$$((u_{n-1}, u_n), (u_1, \dots, u_{n-2})) = (u_{n-1}, u_n, (u_1, \dots, u_{n-2}))$$

and

$$(u_1, \dots, u_n) = ((u_1, \dots, u_{n-2}), u_{n-1}, u_n).$$

Therefore

$$\{(u_1, \dots, u_n) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge u_{n-1} \in u_n\} = \\ = \mathbf{F}_{10}(\mathbf{F}(X_1, \dots, X_n)).$$

(e) Here, $i = n$, $j = n-1$. This case is similar to case 1(d).

2. Assume $\phi(u_1, \dots, u_n)$ is a negation. That is, $\phi(u_1, \dots, u_n)$ is a formula of the form $\neg\psi(u_1, \dots, u_n)$. From the inductive hypothesis, there exists an $\mathbf{F}$ such that

$$\{(u_1, \dots, u_n) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge \psi(u_1, \dots, u_n)\} = \\ = \mathbf{F}(X_1, \dots, X_n).$$

Of course,

$$\{(u_1, \dots, u_n) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge \phi(u_1, \dots, u_n)\} = \\ = (X_1 \times \cdots \times X_n) \setminus \mathbf{F}(X_1, \dots, X_n).$$

3. Assume $\phi(u_1, \dots, u_n)$ is the conjunction $\psi_1(u_1, \dots, u_n) \wedge \psi_2(u_1, \dots, u_n)$. By the inductive hypothesis, there exists function $\mathbf{G}_1$ and $\mathbf{G}_2$ such that

$$\{(u_1, \dots, u_n) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge \psi_i(u_1, \dots, u_n)\} = \\ = \mathbf{G}_i(X_1, \dots, X_n) \text{ for } i = 1, 2.$$

But then,

$$\begin{aligned}\{(u_1,\dots,u_n) : u_1 \in X_1 \wedge \cdots \wedge u_n \in X_n \wedge \phi(u_1,\dots,u_n)\} = \\ = \mathbf{G}_1(X_1,\dots,X_n) \cap \mathbf{G}_2(X_1,\dots,X_n) = \\ = \mathbf{F}_5(\mathbf{G}_1(X_1,\dots,X_n), \mathbf{G}_2(X_1,\dots,X_n)).\end{aligned}$$

4. Assume $\phi(u_1,\dots,u_n)$ is the formula $(\exists u_{n+1} \in u_i)\psi(u_1,\dots,u_{n+1})$. Let $\chi(u_1,\dots,u_n)$ be the formula $\psi(u_1,\dots,u_{n+1}) \wedge (u_{n+1} \in u_i)$. We consider such a formula χ to be less complex than ϕ. By the inductive hypothesis, there exists an $\mathbf{F}$ such that

$$\begin{aligned}\{(u_1,\dots,u_{n+1}) : u_1 \in X_1 \wedge \cdots \wedge u_{n+1} \in X_{n+1} \wedge \chi(u_1,\dots,u_{n+1})\} \\ = \mathbf{F}(X_1,\dots,X_{n+1}),\end{aligned}$$

for all $X_1,\dots,X_{n+1}$. We claim that

$$\begin{aligned}\{(u_1,\dots,u_{n+1}) : u_1 \in X_1 \wedge \cdots \wedge u_{n+1} \in X_{n+1} \wedge \chi(u_1,\dots,u_{n+1})\} = \\ = X_1 \times \cdots \times X_n \cap \operatorname{dom}(\mathbf{F}(X_1,\dots,X_n,\bigcup X_i)).\end{aligned}$$

For ease of notation, let $u = (u_1,\dots,u_{n+1})$ and $X = X_1 \times \cdots \times X_n$. Then, for all $u \in X$ we have

$$\begin{aligned}\phi(u) &\iff \\ &\iff (\exists v \in u_i)\psi(u,v) \\ &\iff \exists v\,(v \in u_i \wedge \psi(u,v) \wedge v \in \bigcup X_i) \\ &\iff u \in \operatorname{dom}(\{(u,v) \in X \times \bigcup X : \chi(u,v)\}).\end{aligned}$$

That completes the proof of case 4, and with it, the proof of the theorem.

$\square_{7.2.2}$

Corollary 7.2.5. *In* $\mathbf{M}$ *is a closed transitive class, then for every* Δ_0*-formula* $\phi(u,p_1,\dots,p_n)$ *and* $X,p_1,\dots,p_n \in \mathbf{M}$ *we have*

$$Y = \{u \in X : \phi(u,p_1,\dots,p_n)\} \in \mathbf{M}.$$

In other terms: $\mathbf{M}$ *satisfies the Axiom of Separation for* Δ_0*-formulas.*

Proof. Let $\phi(u,p_1,\dots,p_n)$ be a closed formula and $X,p_1,\dots,p_n \in \mathbf{M}$. By the previous Theorem 7.2.2, there exists such an operation $\mathbf{F}$ that

$$\mathbf{F}(X,\{p_1\},\dots,\{p_n\}) = \{(u,p_1,\dots,p_n) : u \in X \wedge \phi(u,p_1,\dots,p_n)\}.$$

Now,

$$\begin{aligned}Y = \{u : (\exists u_1 \dots \exists u_n)(u,p_1,\dots,p_n) \in \mathbf{F}(X,\{p_1\},\dots,\{p_n\}) = \\ = \underbrace{\operatorname{dom}\dots\operatorname{dom}}_{n} \mathbf{F}(X,\{p_1\},\dots,\{p_n\}).\end{aligned}$$

Because $\{x,y\}$ and $\operatorname{dom}(x)$ are both Gödel operations, and $\mathbf{M}$ is closed, we thus have that $Y \in \mathbf{M}$. $\square_{7.2.5}$

In fact, the proof of Corollary 7.2.5 is the proof of a somewhat stronger corollary:

Corollary 7.2.6. *In* $\mathbf{M}$ *is a closed transitive class, then for every* Δ_0*-formula* $\phi(u, p_1, \ldots, p_n)$ *and* $X \subseteq \mathbf{M}$, $p_1, \ldots, p_n \in \mathbf{M}$, *there exists an operation* $\mathbf{F}$ *such that*

$$Y = \{u \in X : \phi(u, p_1, \ldots, p_n)\} = \mathbf{F}(X, p_1, \ldots, p_n).$$

For every set M, there exists a smallest closed $W \supseteq M$. In particular, such a set can be built in the following manner:
Set $W = M_0$. Let $W_{n+1} = W_n \cup \{\mathbf{F}(X, Y) : X, Y \in W_n, i = 1, \ldots, 10\}$. Define $W = \bigcup\{W_n : n < \omega\}$.

We denote $W = \mathrm{cl}(M)$. We call W the *Gödel closure* of the set M.

Definition 7.2.7. We say that a class $\mathbf{M}$ is *almost universal* if every subset $x \subseteq \mathbf{M}$ is contained in a certain set $Y \in \mathbf{M}$.

note that above, $x \subseteq \mathbf{M}$ means that x is a set in $\mathbf{V}$, *but not necessarily in* $\mathbf{M}$, and all of the elements of x are also elements of $\mathbf{M}$. Note that an almost universal class must be a proper class. Examples of almost universal classes are $\mathbf{ON}$ and $\mathbf{V}$.

We will prove that a transitive closed almost universal class is a model of ZF. To do this, we will need the following technical lemma:

Lemma 7.2.8 (Reduction Lemma)**.** *For every formula* $\phi(u_1, \ldots, u_n)$ *with* k *quantifiers, let* $\bar{\phi}(u_1, \ldots, u_n, Y_1, \ldots, Y_k)$ *denote the* Δ_0*-formula obtained by substituting all quantifiers* $\exists x$ *and* $\forall x$ *in* ϕ *with* $\exists x \in Y_j$ *or* $\forall x \in Y_j$, $j = 1, \ldots, k$. *Let* $\mathbf{M}$ *be an almost universal transitive class. Then, for every* $X \in \mathbf{M}$ *there exist* $Y_1, \ldots, Y_k \in \mathbf{M}$ *such that*

$$\mathbf{M} \models \phi(u_1, \ldots, u_n) \text{ iff } \bar{\phi}(u_1, \ldots, u_n, Y_1, \ldots, Y_k)$$

for all $u_1, \ldots, u_n \in X$.

Proof. We proceed by induction on the complexity of the formula ϕ. Of course, we can assume that ϕ has only existential quantifiers. Let us notice that

- if ϕ has no quantifiers, then we can set $\bar{\phi} = \phi$;
- if ϕ is $\neg\psi$ or $\psi \wedge \chi$ and the Lemma holds for ϕ and χ, then it holds for ϕ as well.

  Thus, we are left only with the case where ϕ is of the form $\exists v \psi(u, v)$. Let us assume that the Lemma holds for ψ. Assuming that ψ has k quantifiers, we will show that $\bar{\phi}$ is $(\exists v \in Y_{k+1})(\bar{\psi}(u, v, Y_1, \ldots, Y_k)$.

  Let $X \in \mathbf{M}$. We look for $Y_1, \ldots, Y_k, Y_{k+1} \in \mathbf{M}$ such that for every $u \in X$:

  $$\mathbf{M} \models \exists v \psi(u, v) \text{ iff } (\exists v \in Y_{k+1})(\bar{\psi}(u, v, Y_1, \ldots, Y_k).$$

  With this goal in mind, we apply the Collection Principle[1] to the formula $v \in \mathbf{M} \wedge \mathbf{M} \models \psi(u, v)$. Thus, there exists a set M such that $X \subseteq M \subseteq \mathbf{M}$

[1] The *Collection Principle* is the schema of formulas:

$$\forall X\, \exists Y\, (\forall u \in X)(\exists v\, \phi(u, v, p) \to (\exists v \in Y)\phi(u, v, p)).$$

Here, p is a parameter. The intuition here is that if each of the classes $C_u = \{v : \phi(u, v, p)\}$, $u \in X$ is non-empty, then there exists a set Y which intersects each of these classes. This Principle is occasionally given as an axiom of ZF. The Collection Principle implies the Replacement Schema, and can be proven from the axioms of ZF as given in this lecture. For a proof of this, see Jech's Set Theory, p. 65.

and for each $u \in X$ we have

$$\mathbf{M} \models \exists v \psi(u, v) \text{ iff } (\exists v \in Y_{k+1})(\bar{\psi}(u, v, Y_1, \ldots, Y_k).$$

By the inductive hypothesis, for a given $Y \in \mathbf{M}$ there exist $Y_1, \ldots, Y_k \in \mathbf{M}$ such that for all $u, v \in Y$,

$$\mathbf{M} \models \psi(u, v) \text{ iff } \bar{\psi}(u, v, Y_1, \ldots, Y_k).$$

Let us thus set $Y_{k+1} = Y$, since $X \subseteq Y$ and for all $u \in X$ we have

$$\begin{aligned} \mathbf{M} \models \exists v\, \psi(u, v) \quad & \text{iff } (\exists v \in \mathbf{M})\mathbf{M} \models \psi(u, v) \\ & \text{iff } (\exists v \in Y)\mathbf{M} \models \psi(u, v) \\ & \text{iff } (\exists v \in Y)\bar{\psi}(u, v, Y_1, \ldots, Y_k). \end{aligned}$$

$\square_{7.2.8}$

——————— HERE ENDED SPRING 2007 LECTURE 1 (135 min) ———-

Theorem 7.2.9. *Let* $\mathbf{M}$ *be a closed almost universal transitive class. Then* $(\mathbf{M}, \in)$ *is a model of ZF.*

Proof. We need to check the axioms. Of course, we will take advantage of the fact that all axioms hold in $\mathbf{V}$. Thus, we will be proving that the axioms also hold when we interpret them in $\mathbf{M}$.

Extensionality: Every transitive class satisfies Extensionality (6.1.4).

Foundation: Let $S \in \mathbf{M}$ be a non-empty set. Because V satisfies Foundation, there is $x \in S$ which is the $\in$-minimal element of S, i.e. $S \cap x = \emptyset$. Since $\mathbf{M}$ is transitive, $x \in \mathbf{M}$ and $\mathbf{M} \models S \cap x = \emptyset$.

Separation: Let $\phi(u, p)$ be a formula. We wish to show that for all $X, p \in \mathbf{M}$ the set

$$Y = \{u \in X : \mathbf{M} \models \phi(u, p)\}$$

is in $\mathbf{M}$. To this end, we use the Reduction Lemma. Thus, there exists $Y_1, \ldots, Y_k \in \mathbf{M}$ such that

$$Y = \{u \in X : \bar{\phi}(u, p, Y_1, \ldots, Y_k)\}.$$

Since $\bar{\phi}$ is a Δ_0-formula, thus $Y \in \mathbf{M}$ by Corollary 7.2.5.

Pairing and **Union:** Both result from the closure of $\mathbf{M}$ under the Gödel operations, by their transitivity.

Replacement: Let ϕ be a formula, and let us assume that

$$\mathbf{M} \models (\forall x, y, z)(\phi(x, y) \wedge \phi(x, z) \rightarrow x = z).$$

Let $\mathbf{F}$ be a function $\mathbf{F} = \{(x, y) \in \mathbf{M} : \mathbf{M} \models \phi(x, y)\}$. We want to show that $\mathbf{F}[X] \in \mathbf{M}$ for each $X \in \mathbf{M}$. We start by showing that $\mathbf{F}[X] \subseteq \mathbf{M}$ for each $X \in \mathbf{M}$. Because we have assumed that the axioms, particularly Replacement, hold in $\mathbf{V}$, we know that $\mathbf{F}[X]$ is a set in $\mathbf{V}$. So, by transitivity, $y \in \mathbf{F}[X] \rightarrow y \in \mathbf{M}$. Therefore, $\mathbf{F}[X] \subseteq \mathbf{M}$ for each $X \in \mathbf{M}$. Now, $\mathbf{F}[X]$ is a sub*set* of $\mathbf{M}$ in $\mathbf{V}$, so by almost universality, there is $Y \in \mathbf{M}$ such that $\mathbf{F}[X] \subseteq Y$. Now, $\mathbf{F}[X]$ is a subclass of a set (in $\mathbf{M}$). Thus, using Separation in $\mathbf{M}$, we have that $\mathbf{F}[X] \in \mathbf{M}$.

Infinity: By the closure of $\mathbf{M}$, using induction, we can show that every natural number is in $\mathbf{M}$, and so $\omega \subseteq \mathbf{M}$. Using almost universality, we have such $Y \in \mathbf{M}$ that $\omega \subseteq Y$. Separation gives us immediately that $\omega \in \mathbf{M}$.

Powerset: The formula $z \subseteq x$ is Δ_0, and hence the fact that

$$\mathbf{M} \models \forall x\, \exists y\, \forall z\, (z \subseteq x \to z \in y)$$

is equivalent to $(\forall x \in \mathbf{M})(\exists Y \in \mathbf{M})(\mathscr{P}(x) \cap \mathbf{M} \subseteq Y)$. But this results from the almost universality of $\mathbf{M}$. $\square_{7.2.9}$

Lemma 7.2.10. *The Gödel operations are absolute for transitive classes.*

Proof. The absoluteness of $\mathbf{F}_1, \mathbf{F}_4, \mathbf{F}_5$,and $\mathbf{F}_6$ was established by Theorem 6.2.9. The absoluteness of $\mathbf{F}_2$ and $\mathbf{F}_7$ was established by Theorem 6.2.11.

By Lemma 6.2.6, we need to show that $\mathbf{F}_3, \mathbf{F}_8, \mathbf{F}_9$ and $\mathbf{F}_{10}$ can be described using Δ_0-formulas.

For $\mathbf{F}_3$: $Z =\in (X, Y)$ iff $(\forall z \in Z)(\exists x \in X)(\exists y \in Y)(x \in y \wedge z = (x, y)) \wedge (\forall x \in X)(\forall y \in Y)((x \in y) \to (\exists z \in Z)(z = (x, y)))$.

For $\mathbf{F}_8$: $Z = \mathbf{F}_8$ iff $\forall z \in Z)(\exists u \in \mathrm{rng}(X))(\exists v \in \mathrm{dom}(X))(z = (u, v)) \wedge (\forall u \in \mathrm{rng}(X))(\forall v \in \mathrm{dom}(X))(\exists z \in Z)(z = (u, v))$.

The rest are similar. $\square_{7.2.10}$

7.3 Constructible Sets

Now we come to the definition of constructible sets.

Recall that by $\mathrm{cl}(X)$ we mean the Gödel closure of X.

Definition 7.3.1. $\mathrm{def}(U) = \mathrm{cl}(U \cup \{U\}) \cap \mathscr{P}(U)$.

In other terms, $\mathrm{def}(U)$ is the family of all subsets $X \subseteq U$ which can be obtained from U and the elements of U using the Gödel operations.

Corollary 7.3.2. $U \in \mathrm{def}(U)$

Lemma 7.3.3. *If U is a transitive set, then so is* $\mathrm{def}(U)$*.*

Proof. Let $A = \mathrm{cl}(U \cup \{U\}) \cap \mathscr{P}(U)$. By the transitivity of U, we have that $U \subseteq \mathscr{P}(U)$. Since of course $U \subseteq \mathrm{cl}(U \cup \{U\})$, we have that $U \subseteq A$. Now, if $x \in A$, then $x \in \mathscr{P}(U)$, and so $x \subseteq U \subseteq A$, which gives the transitivity of A. $\square_{7.3.3}$

Definition 7.3.4. We can now define the *hierarchy of constructible sets*:

$$\begin{aligned} L_0 &= \emptyset \\ L_{\alpha+1} &= \mathrm{def}(L_\alpha) \\ L_\gamma &= \bigcup\{L_\alpha : \alpha < \gamma\} \text{ for } \gamma \text{ a limit} \\ \mathbf{L} &= \bigcup\{L_\alpha : \alpha \in \mathbf{ON}\} \end{aligned}$$

Definition 7.3.5. A set is called *constructible* if and only if it is an element of the class $\mathbf{L}$.

Notice that each L_α is transitive and that $L_\alpha \subseteq R_\alpha$. Furthermore, the class $\mathbf{L}$ is almost universal and closed.

Corollary 7.3.6. $\mathbf{L}$ *is a model of ZF.*

Axiom (Axiom of Constructibility)**.**

$$\mathbf{V} = \mathbf{L}$$

(or: Every set is constructible.)

Theorem 7.3.7 (Gödel's Theorem)**.**

1. $\mathbf{L}$ *is a model of ZF;*
2. $\mathbf{L}$ *satisfies the Axiom of Constructibility;*
3. *If* $\mathbf{M}$ *is a transitive model of ZF containing all the ordinals, then* $\mathbf{L} \subseteq \mathbf{M}$*.*

Before we prove this theorem, we need a couple of lemmas.

Lemma 7.3.8. *The Gödel Operations* $\mathbf{F}_1, \ldots, \mathbf{F}_{10}$ *are absolute for transitive classes. Furthermore, compositions of Gödel Operations are absolute for transitive classes.*

Proof. We have already established that the Gödel Operations $\mathbf{F}_1, \ldots, \mathbf{F}_{10}$ are absolute for transitive classes in Lemma 7.2.10. What remains is that compositions of these functions are absolute. This fact can be easily deduced from the discussion preceding Theorem 6.2.25 about the Absoluteness of Recursive Classes. $\square_{7.3.8}$

The following lemma is key in showing the consistency of the Axiom of Constructibility.

Lemma 7.3.9.

1. *The function* $\alpha \longrightarrow L_\alpha$ *is absolute for transitive models of ZF.*
2. *If* $\mathbf{M}$ *is a transitive model of ZF containing all the ordinals, then the formula "x is constructible" is absolute for* $\mathbf{M}$ *and* $\mathbf{L}^{\mathbf{M}} = \mathbf{L}$*.*

Proof. First of all, **2** is a consequence of 1 because, if $\mathbf{M}$ is a model of ZF, then the function $\alpha \longrightarrow L_\alpha^{\mathbf{M}}$ is defined in $\mathbf{M}$. By absoluteness, we have for all x:

$$x \in \mathbf{L} \text{ iff } \exists\alpha\,(x \in L_\alpha) \text{ iff } \exists\alpha\,(x \in L_\alpha^{\mathbf{M}}) \text{ iff } x \in \mathbf{L}^{\mathbf{M}}.$$

To prove **1**, let $\mathbf{M}$ be a transitive model of ZF. Recall that our function $\alpha \longrightarrow L_\alpha$ is defined by induction:

$$\begin{aligned} L_0 &= \emptyset \\ L_{\alpha+1} &= \operatorname{def}(L_\alpha) \\ L_\gamma &= \bigcup\{L_\alpha : \alpha < \gamma\} \text{ for } \gamma \text{ a limit} \end{aligned}$$

where $u \in \mathrm{def}(U)$ iff $u \in \mathrm{cl}(U \cup \{U\}) \wedge u \subseteq U$, and $\mathrm{cl}(M) = \bigcup \mathrm{rng}(W)$, where W is defined from M using induction:

$$\begin{aligned} W(0) &= M \\ W_{n+1} &= W_n \cup \{\mathbf{F}_i(X, Y) : X, Y \in W_n, i = 1, \ldots, 10\} \end{aligned}$$

First of all, notice that $W(n+1)$ is obtained from $W(n)$ by an absolute operation, since for all U and Z, we have $Z = \{\mathbf{F}_i(x, y) : x \in U \wedge y \in U, i = 1, \ldots, 10\}$ if and only if

$$(\forall z \in Z)(\exists x, y \in U)(z = \mathbf{F}_1(x, y) \wedge \ldots \wedge z = \mathbf{F}_{10}(x, y)) \wedge$$
$$\wedge (\forall x, y \in U)(\exists z_1, \ldots, z_{10} \in Z)(z_1 = \mathbf{F}_1(x, y) \wedge \ldots \wedge z_{10} = \mathbf{F}_{10}(x, y)).$$

Since $z = \mathbf{F}_i(x, y)$ is a Δ_0-formula (by Lemma 7.2.10) for $i = 1, \ldots, 10$, hence the expression above is also Δ_0, and hence absolute. By Theorem 6.2.25 about the absoluteness of recursively defined classes, we see that the function W is also absolute. Thus we immediately have that $\mathrm{cl}^{\mathbf{M}}(M) = \mathrm{cl}(M)$, for all $M \in \mathbf{M}$. Since the formula $u \in \mathrm{def}(U)$ is absolute for $\mathbf{M}$ and $\mathrm{def}(U) \subseteq \mathbf{M}$ for all $U \in \mathbf{M}$, the operation $\mathrm{def}(U)$ is absolute, whence $\mathrm{def}^{\mathbf{M}}(U) = \mathrm{def}(U)$ for all $U \in \mathbf{M}$.

Again we apply Theorem 6.2.25: since the operations $\bigcup$ and def are absolute we have that $L_\alpha^{\mathbf{M}} = L_\alpha$ $\square_{7.3.9}$

Proof. This results directly from Lemma 7.3.9. $\square_{7.3.7}$

7.4 The Axiom of Choice in L

The next theorem establishes that the Axiom of Choice holds in **L**.

Theorem 7.4.1 (Gödel)**.** *There exists a well-ordering of the class* **L**.

Proof. We will show that **L** has a definable well-ordering. That is, we will define a formula $<$ which is a (class-sized) well-ordering of **L**.

To this end, we define a well-ordering $<_\alpha$ on L_α for each α. We will do this in such a way that for $\alpha < \beta$, $<_\beta$ is an extension of the ordering $<_\alpha$; that is:

1. if $x <_\alpha y$, then $x <_\beta y$;
2. if $x \in L_\alpha$, and $y \in L_\beta$, then $x <_\beta y$.

Notice that this immediately gives

3. if $x \in y \in L_\alpha$ then $x <_\alpha y$.

For limit ordinals γ, we put $<_\gamma = \bigcup\{<_\alpha : \alpha < \gamma\}$. That is, if $x, y \in L_\gamma$, then $x <_\gamma y$ iff $(\exists \alpha < \gamma)(x <_\alpha y)$.

Now, we still have to define $<_{\alpha+1}$ if given $<_\alpha$. Let us first recall the definition of $L_{\alpha+1}$:

$$L_{\alpha+1} = \mathscr{P}(L_\alpha) \cap \mathrm{cl}(L_\alpha \cup \{L_\alpha\}) = \mathscr{P}(L_\alpha) \cap \bigcup \{W_\alpha(n) : n < \omega\},$$

where $W_\alpha(0) = L_\alpha \cup \{L_\alpha\}$ and $W_\alpha(n+1) = \{\mathbf{F}_i(X, Y) : X, Y \in W_\alpha(n), i = 1, \ldots, 10\}$.

The intuition behind the construction of $<_{\alpha+1}$ is the following: first we take elements of L_α in their ordering $<_\alpha$, next L_α itself, then the remaining elements $W_\alpha(1)$, then the remaining elements of $W_\alpha(2)$, and so on.

To order the new elements of $W_\alpha(n+1)$ we will used the defined well-ordering of $W_\alpha(n)$, since every element $x \in W_\alpha(n+1)$ is of the form $\mathbf{F}(u,v)$ for some $i = 1, \ldots, 10$ and certain $u, v \in W_\alpha(n)$.

More precisely, we:

Let $<^0_{\alpha+1}$ be a well-ordering of $L_\alpha \cup \{L_\alpha\}$ extending $<_\alpha$ so that L_α is the last element.

Let $<^{n+1}_{\alpha+1}$ be the following ordering of $W_\alpha(n+1)$: $x <^{n+1}_{\alpha+1} y$ iff $x <^n_{\alpha+1} y$, or $x \in W_\alpha(n)$ but $y \notin W_\alpha(n)$, or both $x, y \in W_\alpha(n+1) \setminus W_\alpha(n)$ but

- The smallest i such that $(\exists u, v \in W_\alpha(n))(x = \mathbf{F}_i(u,v))$ is smaller than the smallest j such that $(\exists u, v \in W_\alpha(n))(y = \mathbf{F}_j(u,v))$; or
- or the smallest i and the smallest j defined above are equal, but the $<^n_{\alpha+1}$-smallest $u \in W_\alpha(n)$ such that $(\exists v \in W_\alpha(n))(x = \mathbf{F}_j(u,v))$ is smaller than the $<^n_{\alpha+1}$-smallest $s \in W_\alpha(n)$ such that $(\exists v \in W_\alpha(n))(y = \mathbf{F}_j(s,v))$; or
- or the same as above, for the second parameter.

We then put $<_{\alpha+1} = \bigcup\{<^n_{\alpha+1} : n < \omega\}$. This completes the construction.

Now, we can define the well-ordering $<_{\mathbf{L}}$ on $\mathbf{L}$ putting $x <_{\mathbf{L}} y$ iff $\exists\alpha\,(x <_\alpha y)$. $\square_{7.4.1}$

We call $x <_{\mathbf{L}} y$ the *canonical well-ordering of* $\mathbf{L}$.

——————— HERE ENDED SPRING 2007 LECTURE 2 (135 min) ————-

7.5 The Generalized Continuum Hypothesis in L

In this section, we will prove that the Axiom of Constructibility implies the Generalized Continuum Hypothesis. To this end, we will need some tools that are somewhat stronger than those we have been using thus far.

Definition 7.5.1. A transitive set M is called *adequate* if M is closed, $\mathrm{cl}(U) \in M$ for every $U \in M$, and for every $\alpha \in M$ we have $\langle L_\xi : \xi < \alpha \rangle \in M$.

Our key tool is the following fact, which is reminiscent of the absoluteness Lemma 7.3.9 that we used to show the consistency of the Axiom of Constructibility.

Theorem 7.5.2. *The function $\alpha \longrightarrow L_\alpha$ is absolute for transitive adequate sets. If M is transitive and adequate, then M satisfies the Axiom of Constructibility iff $M = L_\alpha$ for some α.*

Proof. This proof is roughly similar to that of the mentioned Lemma. Since we are not assuming that M is a model of ZF, we need to prove not only that certain operations are absolute for M, but also that these operations are defined in M in the first place.

Since each of the formulas $z = \mathbf{F}_i(x, y)$ is Δ_0, and since M is closed under Gödel operations, each Gödel formula is defined in M and is absolute for M.

Next, the formula $Z = \{\mathbf{F}_i(x, y) : x, y \in U, i = 1, \dots, 10\}$ is a Δ_0-formula, hence is also absolute for M. Similarly, if $U \in M$, then $Z \in M$, since if we denote $C = \mathrm{cl}(U)$, then $Z = \{u \in C : \exists x, y \in U : (u = \mathbf{F}_1(x, y) \vee \dots \vee \mathbf{F}_{10}(x, y))\}$. Thus, from the Normal Form Theorem, we have $Z = \mathbf{F}(U, C)$. where $\mathbf{F}$ is a composition of Gödel operations.

Now, notice that the formula $x \in \mathrm{cl}(U)$ is absolute for M. To show this, fix $U \in M$ Let W be the function defined by

$$W(0) = U, \text{ and}$$
$$W(n+1) = W(n) \cup \{\mathbf{F}_i(x, y) : x, y \in W(n), i = 1, \dots, 10\}.$$

Since M is closed, if contains all the natural numbers. Thus, by induction we can see that $W \restriction n \in M$, for all n. Thus, the function W is absolute and and defined in M despite the fact that it does not have to be an element of M. Since, however, $x \in \mathrm{cl}(U)$ is equivalent with $(\exists n)(x \in W(n))$, we thus have the required absoluteness of $x \in \mathrm{cl}(U)$.

As a consequence, we get the absoluteness of the operation $\mathrm{def}(U)$, which is defined in M since $\mathrm{def}(U) = \{X \in D : X \subseteq U\} = \mathbf{F}(D, U)$, where $D = \mathrm{cl}(U) \cup \{U\}$ and $\mathbf{F}$ is some composition of Gödel operations.

Now we are ready to show that the function $\alpha \longrightarrow L_\alpha$ is absolute for M. Since this function is defined inductively, we will be simply copying the proofs of Lemma 7.3.8 and 7.3.9. Let us therefore begin with the statement that $y = L_\alpha$ iff α is an ordinal and $\exists f$, such that

- f is a function on $\alpha + 1$, and
- $(\forall \xi \leqslant \alpha)(\xi + 1 \leqslant \alpha + 1 \rightarrow f(\xi + 1) = \mathrm{def}(f(\xi)))$, and
- $(\forall \xi \leqslant \alpha)(\xi \text{ is a limit } \rightarrow f(\xi) = \bigcup \mathrm{rng}(f \restriction \xi))$, and
- $f(\alpha) = y$.

Since the operations $\mathrm{def}(X)$, $\bigcup X$, $\mathrm{rng}(f)$ and $f \restriction \xi$ are absolute for M and defined in M, the adequateness of M implies that the formula $y = L_\alpha$ is absolute for M.

To finish this proof of the absoluteness of the function $\alpha \mapsto L_\alpha$, it is enough to notice that

1. if $\alpha \in M$, then $L_\alpha \in M$. This is so, because $\alpha + 1 \in M$ and L_α is in the transitive closure of $\langle L_\beta : \beta < \alpha + 1 \rangle$ and M is transitive;
2. the formula $x \in L_\alpha$ is also absolute for M. This is so because $x \in L_\alpha$ iff $(\exists y)(x \in y \wedge y = L_\alpha)$.

   Let us now concentrate on the second statement of this theorem. Let M be an adequate transitive set. If $M = L_\alpha$ for some α, then α must be a limit because M is closed and $\mathbf{ON} \cap L_\alpha = \alpha$. Thus we have

   $$(\forall x \in M)(\exists \beta \in M)(x \in L_\beta).$$

   Since the formula $x \in L_\beta$ is absolute for M, we have

   $$M \models \forall x \, \exists \beta \, (x \in L_\beta).$$

This therefore means that M satisfies $\mathbf{V} = \mathbf{L}$.

If M satisfies $\mathbf{V} = \mathbf{L}$, then by the absoluteness of $x \in L_\beta$ we have

$$(\forall x \in M)(\exists \beta \in M)(x \in L_\beta).$$

Consequently, $M = \bigcup\{L_\beta : \beta \in \mathbf{ON}^M\} = L_\alpha$, where $\alpha = \mathbf{ON}^M$.

$\square_{7.5.2}$

As a consequence of this proof, we have the following fact.

Remark 7.5.3. *Let Σ be the conjunction for the following sentences:*

1. $(\forall x, y)(\exists z)(z = \mathbf{F}_i(x, y))$, $i = 1, \ldots, 10$;
2. $(\forall U)(\exists C)(C = \mathrm{cl}(U))$;
3. $(\forall \alpha)(\exists f)(f = \langle L_\beta : \beta < \alpha\rangle)$.

Then, for every transitive set M, M is adequate if and only if $M \models \Sigma$.

Proof. Of course, if M is and adequate transitive set, then the above operation are absolute for M and definable for M. Thus $M \models \Sigma$.

In the other direction, if M is transitive and $M \models \Sigma$, then we can simply repeat the proof of the previous theorem. $\square_{7.5.3}$

Key to the proof that GCH holds in $\mathbf{L}$ is the following lemma.

Lemma 7.5.4. *If $X \in \mathscr{P}(\omega_\alpha) \cap \mathbf{L}$, then there exists $\gamma < \omega_{\alpha+1}$ such that $X \in L_\gamma$.*

Proof. Since X is constructible, we know that for some β, $X \in L_\beta$. By the Reflection Theorems 6.4.3 and 6.4.4, there exists a set S such that :

$$\begin{aligned}
S &\models \text{ Axiom of Extensionality;} \\
S &\models \Sigma \\
\omega_\alpha &\subseteq S \\
X &\in S \\
\beta &\in S \\
|S| &= \aleph_\alpha \\
S &\models \text{“ } \beta \text{ is and ordinal} \wedge X \in L_\beta\text{”}.
\end{aligned}$$

and hence we have the Mostowski Collapse

Since S is extensional, there exists an isomorphism π of the set S onto a transitive set $M = \pi[S]$, and we have the following:

$$\begin{aligned}
M &\text{ is adequate} \\
\pi(\xi) &= \xi, \text{ for all } \xi < \omega_\alpha \\
\pi(X) &= X \\
|M| &= \aleph_\alpha \\
M &\models \text{ “}\pi(\beta) \text{ is an ordinal} \wedge \pi(X) \in L_{\pi(\beta)}\text{”}.
\end{aligned}$$

The ordinals are absolute, thus $\gamma = \pi(\beta)$ is an ordinal and $M \models X \in L_\gamma$. Since M is adequate, the formula $X \in L_\gamma$ is absolute for M, thus $X \in L_\gamma$. Finally, $\gamma \in M$, M is transitive, and $|M| = \aleph_\alpha$. This gives us that $\gamma < \omega_{\alpha+1}$. $\square_{7.5.4}$

Theorem 7.5.5 (Gödel). *If* $\mathbf{V} = \mathbf{L}$, *then* $2^{\aleph_\alpha} = \aleph_{\alpha+1}$, *for every* α.

Proof. Assume that $\mathbf{V} = \mathbf{L}$. By induction it is clear that for $\gamma \geqslant \omega$, we have $|L_\gamma| = |\gamma|$. Now, by Lemma 7.5.4, we see that every subset of ω_α is constructible before the $\omega_{\alpha+1}$-th step, and so $\mathscr{P}^{\mathbf{L}}(\omega_\alpha) \subseteq L_{\omega_{\alpha+1}}$. Thus, $|\mathscr{P}^{\mathbf{L}}(\omega_\alpha)| \leqslant |L_{\omega_{\alpha+1}}| = \aleph_{\alpha+1}$. $\square_{7.5.5}$

Chapter 8

The Independence of the Axiom of Choice from ZFU

In the previous chapter, we established the consistency of the Axiom of Choice with the rest of the axioms of ZF. In a following chapter, we will show its independence using the method of forcing.

In the present chapter, we will study an older method, that of Permutation Models and attributed to Fraenkel, Mostowski, and Specker but we usually forget about Specker... that shows the independence of the Axiom of Choice from set theory with *urelemente*, or *atoms*. We present this older method because the concepts behind the two methods are quite similar, but here we will not have the added complicating fog of forcing.

The very vague intuition here is that we wish to "confuse" the model so that it cannot tell the difference between various sets, and so has no basis on which a "choice" can be made.

8.1 Set theory with urelemente

Set theory with urelemente ZFU (sometimes called set theory with atoms and abbreviated ZFA) is a modified version of ZF set theory. This modified ZFU differs from the usual kind in that it admits objects other than sets – *urelemente* or *atoms*. Urelemente are objects that do not contain any elements.

The language of ZFU consists of $=$ and $\in$ and has two constant symbols $\emptyset$ and A, which will be the set of all urelemente. So, one difference here is that we have the constants...

The axioms of ZFU are as in ZF, but with a few modifications to take into account the existence of urelemente:

Axiom (Empty Set)**.**

$$\neg\exists x\,(x \in \emptyset).$$

Recall that in our presentation of the axioms of ZF, the existence of an empty set was a consequence of Set Existence, Comprehension, and Extensionality. Here, we have added a symbol for the empty set to our language, and we need this axiom to differentiate it from an urelement.

We also need to say something about our other constant:

Axiom (Urelemente)**.**

$$\forall z\,(z \in A \iff z \neq \emptyset \wedge \neg\exists x\,(x \in z)).$$

So, *urelemente* are the elements of A. On the other hand, *sets* are all objects which are not urelemente.

We need to modify a couple of older axioms. We will write $(\forall_{set} X)$ as shorthand for $\forall X\,(X \notin A)$ i.e. "for all sets X".

Axiom (Extensionality^{ZFU})**.**

$$(\forall_{set} X)(\forall_{set} Y)(\forall u\,(u \in X \iff u \in Y) \iff X = Y).$$

Axiom (Foundation^{ZFU})**.**

$$(\forall_{set} S \neq \emptyset)(\exists x \in S)(x \cap S = \emptyset).$$

The other axioms are modified similarly – by restricting the objects they govern to sets.

Some operations that we defined earlier only make sense for sets, such as $\bigcup X$ or $\mathscr{P}(X)$. Some also make sense for atoms, like $\{x, y\}$.

If we were to add to ZFU the axiom $A = \emptyset$, then we would get ZF.

The development of set theory with urelemente is a lot like that of ZF. Again, we have to make a couple of minor modifications. In the definition of the *ordinals*, we have to insert the statement that an ordinal does not have urelemente among its elements. One can also define the *rank* of sets and build a *hierarchy* $\mathcal{U}_\alpha$ analagous to the van Neumann R_α.

For any set S, let $\mathscr{P}^\alpha(S)$ be defined as follows:

$$\begin{aligned}
\mathscr{P}^0(S) &= S;\\
\mathscr{P}^{\alpha+1}(S) &= \mathscr{P}^\alpha(S) \cup \mathscr{P}(\mathscr{P}^\alpha(S));\\
\mathscr{P}^\gamma(S) &= \bigcup_{\alpha<\gamma} \mathscr{P}^\alpha(S) \text{ for } \gamma \text{ a limit};\\
\mathscr{P}^\infty(S) &= \bigcup_{\alpha\in\mathbf{ON}} \mathscr{P}^\alpha(S).
\end{aligned}$$

Then,

$$\mathcal{U}_\alpha = \mathscr{P}^\alpha(A)$$

and

$$\mathcal{U} = \mathscr{P}^\infty(A).$$

We call the class $\mathscr{P}^\infty(\emptyset)$ the *kernel*. The kernel is a model of ZF, and all the ordinals are in the kernel.

A *transitive* set in this context does not necessarily contain $\emptyset$, and most importantly for our purposes, may have nontrivial automorphisms. For example, the set $\{a_1, a_2\}$ where $a_1, a_2 \in A$ is transitive in the sense of ZFU and admits an automorphism (that respects $\in$ and $=$) that switches a_1 and a_2.

A transitive class which is almost universal and closed under the Gödel operations is a model of ZFU. If the class contains $\emptyset$, we can interpret $\in$ as $\in$, $\emptyset$ as $\emptyset$, and urelemente as urelemente.

Theorem 8.1.1. $\mathrm{Con}(ZFC) \to \mathrm{Con}(ZFU + AC)$.

Proof. We will only sketch this proof.

Assume that we work in $\mathbf{V}$ where all the axioms of ZFC are satisfied.

Let C be a (countably) infinite set of infinite subsets of ω. We will build a model using this set.

Choose $a_0 \in C$. In the model $\amalg$ we build, we will interpret this a_0 as $\emptyset$, and the set $C \setminus \{a_0\}$ as the set of urelemente A.

Let

$$\amalg = \bigcup_{\alpha \in \mathbf{ON}} \amalg_\alpha,$$

where

$$\begin{aligned}
\amalg_0 &= C; \\
\amalg_{\alpha+1} &= \amalg_\alpha \cup \mathscr{P}(\amalg_\alpha) \setminus \{\emptyset\}; \\
\amalg_\gamma &= \bigcup_{\alpha<\gamma} \amalg_\alpha \text{ for } \gamma \text{ a limit ordinal.}
\end{aligned}$$

First, we show that $\amalg$ is a transitive class in the sense of ZFU, with $\amalg_0 = C$ interpreted as above.

We then check all the axioms: the Empty Set and Urelemente Axioms clearly hold with the above interpretation. The ZFU versions of Extensionality, Separation, Powerset, Replacement, Pairing and Union, Foundation, and Choice hold by arguments similar to those of Lemma 6.1.4, Corollary 6.1.6, Lemmas 6.1.9, 6.1.11, 6.1.10, 6.1.12, and 6.2.17, respectively. The Axiom of Infinity holds by arguments similar to those of Lemma 6.2.12. Note here that the “real” ω (i.e. the set that $\mathbf{V}$ considers to be ω) is not in $\amalg$ because we removed the empty set at every stage $\amalg_\alpha$ of its construction. However, with our interpretation of a_0 as the empty set, there is another set in $\amalg$ that can play the role of ω. $\square_{8.1.1}$

8.2 Fraenkel-Mostowski-Specker Permutation Models

The inspiring idea for Fraenkel-Mostowski-Specker permutation models (or *FM-models* for short) is the fact that the axioms of ZFU do not distinguish between the urelemente. We use this method to construct models in which the set A of urelemente has no well-ordering.

We work in in ZFCU.

8.2.1 $\in$-automorphisms of the universe

As mentioned earlier, we can define an $\in$-automorphism of the universe: Let π be a permutation of the set A. We use the $\mathcal{U}_\alpha$ hierarchy to define πx for every $x \in \mathcal{U}$ as follows. We put

$$\pi(\emptyset) = \emptyset$$

and

$$\pi(x) = \pi'' x = \{\pi(y) : y \in x\},$$

using either $\in$-recursion or by recursion on the rank of x. We will say that this is an automorphism of the universe induced by the permutation π of A.

Remark 8.2.1. *The function π, defined as above, is an $\in$-automorphism of the universe, and so the following facts hold:*

1. $x \in y \iff \pi x \in \pi y$
2. $\phi(x_1, \ldots, x_n) \iff \phi(\pi x_1, \ldots, \pi x_n)$.
3. $\mathrm{rank}(x) = \mathrm{rank}(\pi x)$.
4. $\pi\{x, y\} = \{\pi x, \pi y\}$ *and* $\pi(x, y) = (\pi x, \pi y)$.
5. *If R is a relation, then πR is a relation and* $(x, y) \in R \iff (\pi x, \pi y) \in R$.
6. *If f is a function on X, then πf is a function on πX and* $(\pi f)(\pi x) = \pi(f(x))$.
7. $\pi x = x$ *for every x in the kernel.*
8. *If ρ is another $\in$-automorphism, then* $(\pi \cdot \rho)x = \pi(\rho(x))$.

I leave the proof of the above as an exercise.

8.2.2 A few reminders from group theory

Now, a couple of definitions out of group theory:

Definition 8.2.2. By $\leq$ we denote the subgroup relation.

Let G be a group of permutations of some set S. The *setwise stabilizer* of $x \subseteq S$ is

$$G_{\{x\}} = \{\pi \in G : \pi x = x\}.$$

Clearly $G_{\{x\}} \leq G$.

The *pointwise stabilizer* of $x \subseteq S$ is

$$G_{(x)} = \{\pi \in G : \pi y = y \text{ for all} y \in x\}.$$

Again, it is clear that $G_{(x)} \leq G$.

Recall some easy facts:

Remark 8.2.3. 1. *Let $x \subseteq S$. Then,*$G_{(x)} \leq G_{\{x\}}$.

2. $\pi G_{(x)} \pi^{-1} = G_{(\pi x)}$.

8.2.3 The definition of the model

Definition 8.2.4. Let G be a group of $\in$-automorphisms induced by a group of permutations of our set of urelemente A. A set $\mathcal{F}$ of subgroups of G is called a *normal filter* on G if for all subgroups $H, K \leq G$ we have

1. $G \in \mathcal{G}$;
2. if $H \in \mathcal{F}$ and $H \leq K$, then $K \in \mathcal{F}$;
3. if $H \in \mathcal{F}$ and $K \in \mathcal{F}$, then $H \cap K \in \mathcal{F}$;
4. $\mathcal{F}$ is closed under conjugacy: if $\pi \in G$ and $H \in \mathcal{F}$, then $\pi H \pi^{-1} \in \mathcal{F}$;

5. for each $a \in A$, the stabilizer of a $G_{(\{a\})} = \{\pi \in G : \pi a = a\} \in \mathcal{F}$.

Now we have the ingredients for the model. Let us fix G and $\mathcal{F}$.

Definition 8.2.5. We say that x is *symmetric* if $G_{\{x\}} \in \mathcal{F}$.
The class

$$\mathfrak{U} = \{x : (\forall y \in x)(y \text{ is symmetric}) \wedge (x \text{ is symmetric})\},$$

defined by $\in$-recursion or by the rank of x, is composed of all *hereditarily symmetric* objects.
The class $\mathfrak{U}$ is called a *permutation model*.

For most practical applications of this model construction method, it is more helpful to think of the normal filter as being generated in a specific way – using a *normal ideal of supports.*

Definition 8.2.6. Let G be our fixed group of permutations of our set of urelemente A. A family $\mathcal{I}$ of subsets of A is called a *normal ideal* if for all subsets $E, F \subseteq A$,

1. $\emptyset \in \mathcal{I}$;
2. if $E \in \mathcal{I}$ and $F \subseteq E$, then $F \in \mathcal{I}$;
3. if $E, F \in \mathcal{I}$, then$E \cup F \in \mathcal{I}$;
4. if $\pi \in G$ and $E \in \mathcal{I}$, then $\pi'' E \in \mathcal{I}$;
5. for each $a \in A$, $\{a\} \in \mathcal{I}$.

We can then define $\mathcal{F}$ to be the filter on G generated by the pointwise stabilizer subgroups $G_{(E)}$, where $E \in \mathcal{I}$. Clearly, $\mathcal{F}$ generated in this way is a normal filter, and we get a permutation model as described above. However, with this particular way of generating the filter, we have another way of describing when a set is in the model. Namely, this happens if it has a support:

Definition 8.2.7. A set x is symmetric if and only if there exists $E \in \mathcal{I}$ such that $G_{(E)} \leq G_{\{x\}}$. In this case, we say that E is a *support* of x.

Theorem 8.2.8. *The class $\mathfrak{U}$ is a transitive model of ZFU. Furthermore, it contains all of the elements of the kernel, and $A \in \mathfrak{U}$.*

Proof. Clearly, $\mathfrak{U}$ is transitive.
To show that $\mathfrak{U}$ is closed under the Gödel operations, it is enough to show
exercise! that for all x and y,

$$G_{\mathbf{F}_i(x,y)} \geq G_{\{x\}} \cap G_{\{y\}}, \text{ for } i = 1, \ldots, 10.$$

We must still show that $\mathfrak{U}$ is almost universal. To do this, we show that the set $\mathcal{U}_\alpha \cap \mathfrak{U}$ is symmetric. In particular, we will show that $G_{\{\mathcal{U}_\alpha \cap \mathfrak{U}\}} = G$. By one of the easy facts about permutation groups above, we can see that if x is symmetric and $\pi \in G$, then πx is symmetric. Thus, by induction, if $x \in \mathfrak{U}$, then $\pi x \in \mathfrak{U}$ for all x and all $\pi \in G$. Since $\text{rank}(\pi x) = \text{rank}(x)$, we have that $\pi(\mathcal{U}_\alpha \cap \mathfrak{U}) = \mathcal{U}_\alpha \cap \mathfrak{U}$ for all α and $\pi \in G$.
The kernel $\mathscr{P}^\infty(\emptyset) \subseteq \mathfrak{U}$ because $G_{\{x\}} = G$ for all $x \in \mathscr{P}^\infty(\emptyset)$.
The set A of urelemente of $\mathcal{U}$ is also in $\mathfrak{U}$ because $G_{\{A\}} = G$, and $G_{(\{a\})} \in \mathcal{F}$ for each $a \in A$. $\square_{8.2.8}$

We have established that the elements of the kernel $\mathscr{P}^\infty(\emptyset)$ are all in $\mathfrak{U}$, and since we assumed that $\mathcal{U}$ is a model of ZFCU, thus, the Axiom of Choice holds in the kernel. So, by the equivalence of the Axiom of Choice to the Well-Ordering Principle, every $x \in \mathscr{P}^\infty(\emptyset)$ can be well-ordered. Thus, any $x \in \mathfrak{U}$ can be well-ordered if and only if there is a one-to-one mapping f of x into the kernel. However, such an f must also be in the model $\mathfrak{U}$, and as such, must be symmetric. Note, however, that

$$\pi \in G_{\{f\}} \iff \pi f = f \iff \pi \in G_{(x)}.$$

Thus,

$$\mathfrak{U} \models (x \text{ can be well-ordered } \iff G_{(x)} \in \mathcal{F}$$

.

8.2.4 An example: the basic Fraenkel model

The following example of a permutation model does not satisfy the Axiom of Choice.

Let A be an infinite countable set. Let G be the group of all permutations of A, and let $\mathcal{I}$ be the normal ideal of supports generated by finite subsets of A. Let $\mathfrak{U}$ be the resulting permutation model. We show that A cannot be well-ordered in $\mathfrak{U}$.

By comments above, we need to show that $G_{(A)}$ is not in the normal filter $\mathcal{F}$ generated by $\mathcal{I}$. That is, we would need to show that there is some finite set $E \subset A$ such that $G_{(E)} \leq G_{(A)}$. However, for each finite $E \subset A$, one can easily find $\pi \in G$ such that $\pi \in G_{(E)}$ but $\pi \notin G_{(A)}$. Thus, A cannot be well-ordered in $\mathfrak{U}$. Therefore, we get that:

Theorem 8.2.9. *The Axiom of Choice is independent from ZFU.*

Another example might be a good idea.

——————— HERE ENDED SPRING 2007 LECTURE 3 (135 min) ———-

Some remarks

The student may find that permutation model construction is not, on the surface, a very "logical" construction. The arguments above are on the basis of groups and filters, and not formulas. We make a few remarks that may lead the student to have some intuition about the logical mechanisms behind the construction.

First, we need a bit of model theory.

Let A be a countably infinite structure in some countable language L. Let $G = \mathrm{Aut}(A)$ be a group of automorphisms of A. Let $X \subseteq A$.

We define

$$DCL(X) = \{a \in A : \pi \in G_{(X)} \to \pi a = a\}.$$

So, $a \in DCL(X)$ if the orbit of a under $G_{(X)}$ is a singleton, or in other terms, $G_{(X)} < G_{(X \cup \{a\})}$, which implies that $G_{(X)} = G_{(X \cup \{a\})}$.

We define

$$ACL(X) = \{a \in A : \text{ the orbit of } a \text{ under } G_{(X)} \text{ is finite}\}.$$

Thus, for $a \in ACL(X)$, if E is the (finite) orbit of a under $G_{(X)}$, then $G_{(X)} < G_{\{E\}}$.

On the other hand, we say an element $a \in A$ is *L-definable over X* if there is a formula $\phi(x)$ in the language L with parameters from X such that $A \models \phi(a) \wedge \exists! x\, \phi(x)$. We say an element $a \in A$ is *L-algebraic over X* if there is a formula $\phi(x)$ in the language L with parameters from X such that $A \models \phi(a) \wedge \exists_{\leqslant n} x\, \phi(x)$ for some $n \in \omega$.

Also, a little reminder about infinitary languages: Often, we abuse notation and write L for the *signature*, or set of constant, relation, and function symbols of a language L. A language $L_{\omega_1 \omega}$ consists of formulas which can have at most $< \omega$ (so, finitely) many quantifiers in a row, and $< \omega_1$ (so, countably) many formulas can be joined together with $\wedge$ or $\vee$. Using this convention, first-order languages can be denoted $L_{\omega\omega}$.

The following lemma shows how these notions can sometimes be connected. The proof of the lemma is outside the scope of this lecture.

Lemma 8.2.10. *Let L be $L_{\omega_1 \omega}$ with a countable signature. Let A be a countable L-structure. Then, for all sets X of elements of A,*

$$DCL(X) = \{a \in A : a \text{ is } L\text{-definable over } X\}.$$

Similarly,

$$ACL(X) = \{a \in A : a \text{ is } L\text{-algebraic over } X\}.$$

Chapter 9

Forcing

———- SPRING 2008 LECTURE 1 consisted of the presentation "Forcing in a Nutshell"———-

———- SPRING 2007 LECTURE 4 consisted of the presentation "Forcing in a Nutshell"———-

9.1 A few remarks on the metamathematics of forcing

The method of forcing is a powerful general technique for producing a wide variety of models satisfying diverse mathematical properties. Forcing was discovered by Paul Cohen in the early 1960's for the purpose of showing the independence of GCH and AC from ZF set theory.

We point out a metamathematical difficulty that was glossed over in the "Nutshell" presentation. Suppose we wish to demonstrate the consistency of $ZF + (\mathbf{V} \neq \mathbf{L})$ (or indeed, anything stronger that that). We cannot simply continue with similar methods as those employed up until now, i.e. work in ZF or ZFC and define a transitive model for the appropriate axioms.

Why not? Well, suppose we, working within ZFC (i.e. in $\mathbf{V}$), were to define a transitive *proper class* $\mathbf{N}$. Suppose it were possible that we could prove that each axiom of $ZF + (\mathbf{V} \neq \mathbf{L})$ is true in $\mathbf{N}$. By Gödel's Theorem 7.3.7 that $\mathbf{L}$ is minimal, we would have that $\mathbf{L} \subset \mathbf{N}$. Also, this inclusion would be proper, $\mathbf{L} \neq \mathbf{N}$, since $\mathbf{V} = \mathbf{L}$ is true in $\mathbf{L}$, but false in $\mathbf{N}$. So, working in ZFC, we would be able to prove that there is a proper extension of $\mathbf{L}$. In other terms, we would demonstrate that $ZFC \vdash \mathbf{V} \neq \mathbf{L}$. This is impossible if we assume that ZFC is consistent because we have shown that $\text{Con}(ZFC) \to \text{Con}(ZFC + (\mathbf{V} = \mathbf{L}))$.

The obvious way of getting around this would be to produce a set sized model. However, as a result of the Gödel Incompleteness Theorem, one cannot argue within ZFC and produce a set model for ZFC. What we *are* able to do is produce a countable transitive model M for any desired *finite* list of axioms of ZFC, as we established in the section on Reflection Theorems. We will then produce an extension N of M which is a model for a finite list of axioms $ZFC + (\mathbf{V} \neq \mathbf{L})$. The finite lists will be chosen so that our arguments can be carried out.

Formally, our proof of $\text{Con}(ZFC) \to \text{Con}(ZFC + (\mathbf{V} \neq \mathbf{L}))$ will be as follows: Assume that ZFC is consistent. Assume that $ZFC + (\mathbf{V} \neq \mathbf{L})$ is inconsistent – that we can derive a contradiction from it. Then, there is a finite list of axioms

$\phi_1, \ldots, \phi_n$ of $ZFC + (\mathbf{V} \neq \mathbf{L})$ such that

$$\phi_1, \ldots, \phi_n \vdash \psi \wedge \neg\psi.$$

Using forcing, we will show that $ZFC \vdash \exists N\,(\phi_1^N \wedge \ldots \phi_n^N)$, and so $ZFC \vdash \exists N\,(\psi \wedge \neg\psi)$, which implies that ZFC is inconsistent, contrary to our assumption.

Practically, we will thus say "let M be a countable transitive model of ZFC", by which we will formally mean "let M be a countable transitive model of enough axioms of ZFC so that we can carry out the argument at hand".

9.2 Partial orders

Let M be a countable transitive model of ZFC. We will see that if $\langle \mathbb{P}, \leq \rangle$ is a partial order, and $\langle \mathbb{P}, \leq \rangle \in M$, then $\langle \mathbb{P}, \leq \rangle$ will give us a method of getting a so-called *generic extension* N of M, which will also be a model of ZFC.

Definition 9.2.1. A *partial order* is a pair $\langle \mathbb{P}, \leq \rangle$ such that $\mathbb{P} \neq \emptyset$ and $\leq$ is a relation on $\mathbb{P}$ which is transitive and reflexive. We read $p \leq q$ as "*p extends q*". Elements of $\mathbb{P}$ are called *conditions*.

A *chain* in $\mathbb{P}$ is a set $C \subset \mathbb{P}$ such that $\forall p, q \in C\,(p \leq q \vee q \leq p)$. We say that p and q are *compatible* iff

$$\exists r \in \mathbb{P}\,(r \leq p \wedge r \leq q);$$

they are *incompatible* (symbolically: $p \perp q$) iff $\neg\exists r \in \mathbb{P}\,(r \leq p \wedge r \leq q)$. An *antichain* in $\mathbb{P}$ is a subset $A \subset \mathbb{P}$ such that $\forall p, q \in A\,(p \neq q \rightarrow p \perp q)$.

We will be interested in partial orderings $\mathbb{P}$ having a *maximal element* $\mathbb{1}$. By this we mean that $\forall p \in \mathbb{P}\,(p \leq \mathbb{1})$. We do not lose any generality with this restriction: one can still force with partial orders lacking a maximal element, but the procedure is messier and one will not produce any more consistency results than one would using a partial order with a maximal element. To be specific about the maximal element, we will write $\langle \mathbb{P}, \leq, \mathbb{1} \rangle$ to denote the partial order. We need to be specific in this way because, of course the set $\mathbb{P}$ does not determine its ordering $\leq$. Furthermore, because our partial ordering relation is not strict, there could be more than one maximal element, thus $\langle \mathbb{P}, \leq \rangle$ does not determine the maximal element $\mathbb{1}$.

We will abuse notation a bit: we will talk about "the partial order $\mathbb{P}$" or "the partial order $\leq$". We write $\mathbb{P} \in M$ and mean $\mathbb{P} \in M$, $\leq \in M$, and $\mathbb{1} \in M$ (the last follows in any case from $\mathbb{P} \in M$ and the transitivity of M). If we are talking about more than one partial order, we will notate the difference so: $\langle \mathbb{P}, \leq_{\mathbb{P}}, \mathbb{1}_{\mathbb{P}} \rangle$ and $\langle \mathbb{Q}, \leq_{\mathbb{Q}}, \mathbb{1}_{\mathbb{Q}} \rangle$.

Definition 9.2.2. Let $\langle \mathbb{P}, \leq \rangle$ be a partial order. We say that $D \subset \mathbb{P}$ is *dense* in $\mathbb{P}$ iff $\forall p \in \mathbb{P}\,\exists q \leq p\,(q \in D)$.

We say that $G \subset \mathbb{P}$ is a *filter* on $\mathbb{P}$ iff

1. $\forall p, q \in G\,\exists r \in G\,(r \leq p \wedge r \leq q)$, i.e. all elements of G are compatible. and
2. $\forall p \in G\,\forall q \in \mathbb{P}\,(p \leq q \rightarrow q \in G)$ i.e. is closed upwards.

Definition 9.2.3. Let $\mathbb{P}$ be a partial order. We say that G is $\mathbb{P}$-*generic* over M iff G is a filter on $\mathbb{P}$ and for all dense $D \subset \mathbb{P}$, $D \in M \to G \cap D \neq \emptyset$.

Lemma 9.2.4. *If M is countable and $p \in \mathbb{P}$, then there is a G (in $\mathbf{V}$) which is $\mathbb{P}$-generic over M such that $p \in G$.*

Proof. Let D_n, $n \in \omega$, be an enumeration of all the dense subsets of $\mathbb{P}$ which are in M. Recall that M is countable, so we will have at most countably many such dense sets. By induction, we choose a sequence q_n, $n \in \omega$, so that $p = q_0 \geq q_1 \geq \ldots$ and $q_{n+1} \in D_n$. This is possible because each D_n is dense. Let G be the filter generated by $\{q_n : n \in \omega\}$, that is, let $G = \{p \in \mathbb{P} : \exists n\, (p \geq q_n)\}$. Then, G is a filter and $G \cap D_n \neq \emptyset$ for each $n \in \omega$, and so is generic. $\square_{9.2.4}$

We will need to keep track of which of our notions are absolute for M, and which notions are not. Recall that we (will) assume that M is a countable transitive model for ZFC, and that $\langle \mathbb{P}, \leq, \mathbb{1} \rangle$ is in M. By standard arguments, one can find that the notions "is a partial order" and "is dense" are absolute for M. On the other hand, the enumeration of the dense sets D_n happens outside of M. By absoluteness,

$$\{D \in M : D \text{ is dense in } \mathbb{P}\} = \{D : D \text{ is dense in } \mathbb{P}\}^M,$$

but this set will not usually be countable in M.

Our definition of a generic and Lemma 9.2.4 did not require that M is a model. However, it will become important as we develop the machinery of forcing that M satisfy at least certain of the axioms of ZFC to ensure that various dense sets that we will construct do lie in M. That M satisfies at least part of ZFC is also important for the next lemma:

Lemma 9.2.5. *If M is a transitive model of $ZF - P$, $\mathbb{P} \in M$ is a partial order such that*

$$\forall p \in \mathbb{P}\, \exists q, r \in \mathbb{P}\, (q \leq p \wedge r \leq p \wedge q \perp r),$$

and G is $\mathbb{P}$-generic over M, then $G \notin M$.

We call a partial order $\mathbb{P}$ *separable* if it satisfies the condition above that

$$\forall p \in \mathbb{P}\, \exists q, r \in \mathbb{P}\, (q \leq p \wedge r \leq p \wedge q \perp r).$$

Proof. Assume, to the contrary, that $G \in M$. Then, $D = \mathbb{P} \setminus G \in M$, because set-theoretic difference is absolute. Furthermore, D is dense: if $p \in \mathbb{P}$ and q, r are as above (i.e such that $q \leq p \wedge r \leq p \wedge q \perp r$), then q and r cannot both be in G because G is a filter. Thus, p has an extension in D.

However, $G \cap D = \emptyset$, contradicting the definition of generic. $\square_{9.2.5}$

The above proof only required M to satisfy a very weak (finite) fragment of $ZF - P$. We won't keep track of exactly which finite subset of axioms of ZFC M will have to satisfy. There will be a finite number of steps where a finite fragment is needed, so keeping exact track isn't worth the extra effort.

Note also that if $\mathbb{P}$ fails the extra condition in the statement of Lemma 9.2.5, then there is a filter G on $\mathbb{P}$ which intersects all dense sets of $\mathbb{P}$, and if $\mathbb{P} \in M$, then $G \in M$. Thus, any application of forcing based on such a partial order will be trivial. So, all partial orders found in our practical applications of forcing will also satisfy this special condition.

——— HERE ENDED SPRING 2008 LECTURE 2 (90 min) ———

9.3 Generic extensions

Let M be a countable transitive model for ZFC, $\mathbb{P}$ a partial order in M, and G a $\mathbb{P}$-generic over M. In this section, we will show the method with which we can construct another countable transitive model for ZFC, which we will call $M[G]$. This new model will be such that $M \subset M[G]$, the two models will have the same ordinals (that is, $o(M) = o(M[G])$), $G \in M[G]$, and $M[G]$ will be the least extension of M to a countable transitive model for ZFC containing G. Lemma 9.2.5 implies that in most of our cases, $M \neq M[G]$.

Note that the axioms beyond ZFC that $M[G]$ satisfies are very dependent on the combinatorial properties satisfied by $\mathbb{P}$ *in* M, and most of the time, these properties are not absolute!!

The construction may seem complicated, but once it is understood, then the problem of finding a partial order $\mathbb{P}$ with which to produce a desired consistency result will reduce to a problem in the combinatoric of partial orders.

The first step is to define $M[G]$. Roughly speaking, this will be the set of all sets that can be constructed from G using set-theoretic processes that are definable in M (and here it may seem that these arguments are a very distant cousin of the definition of L...). Each element in $M[G]$ will have a *name* in M (and here recall the analogy with field extensions), which tells how it has been constructed from G. We will use letters τ, σ, and π to denote names. Inhabitants of the universe M will be able to comprehend a name τ, for an object in $M[G]$, but often, they will not have a sense of the object τ_G that τ names. Knowing the object named would require a knowledge of G.

Definition 9.3.1. We say that τ is a $\mathbb{P}$-name iff τ is a relation and

$$\forall \langle \sigma, p\rangle \in \tau\, (\sigma \text{ is a } \mathbb{P}\text{-name } \wedge p \in \mathbb{P}).$$

Note that the definition of a name does not mention models, or any order on $\mathbb{P}$. Note also that the collection of $\mathbb{P}$-names will be a proper class if $\mathbb{P} \neq \emptyset$.

Definition 9.3.1 has to be viewed as a definition by transfinite recursion. Seeing a formal definition of a name may be helpful in this: to this end, we define the characteristic function of the $\mathbb{P}$-names, $\mathbf{H}(\mathbb{P}, \tau)$, by

$$\mathbf{H}(\mathbb{P}, \tau) = \begin{cases} 1 & \text{iff } \tau \text{ is a relation } \wedge \forall \langle \sigma, p\rangle \in \tau\, (\mathbf{H}(\mathbb{P}, \sigma) = 1 \wedge p \in \mathbb{P}); \\ 0 & \text{otherwise.} \end{cases}$$

Then, τ is a $\mathbb{P}$-name iff $\mathbf{H}(\mathbb{P}, \tau) = 1$. For a fixed partial order $\mathbb{P}$, the function $\mathbf{H}(\mathbb{P}, \tau)$ is defined from $\mathbf{H} \restriction \text{trcl}(\tau)$ using concepts absolute for transitive models of $ZF - P$, so $\mathbf{H}$ is absolute for transitive models of $ZF - P$. (think about Theorem 6.2.25 with the relation $x\mathbf{R}y$ iff $x \in \text{trcl}(y)$ in mind.) Therefore the notion "τ is a name" is absolute for transitive models of $ZF - P$.

Definition 9.3.2. We define $\mathbf{V}^{\mathbb{P}}$ to be the class of $\mathbb{P}$-names. If M is a transitive model of ZFC and $\mathbb{P} \in M$, then $M^{\mathbb{P}} = \mathbf{V}^{\mathbb{P}} \cap M$. Or, by absoluteness,

$$M^{\mathbb{P}} = \{\tau \in M : (\tau \text{ is a } \mathbb{P}\text{-name})^M\}.$$

When forcing over M, we only use the $\mathbb{P}$-names in $M^{\mathbb{P}}$, which can be thought of as having been defined in M.

Definition 9.3.3.

$$\tau_G = \{\sigma_G : \exists p \in G\,(\langle \sigma, p\rangle \in \tau)\}.$$

Note that, like the definition of $\mathbb{P}$-name, τ_G is defined by transfinite recursion on τ. One can think G as a great big dictionary, or parser, and about τ_G as the translation, or meaning, of the name τ according to G.

Definition 9.3.4. if M is a transitive model for ZFC, $\mathbb{P} \in M$, and $G \subset \mathbb{P}$, then

$$M[G] = \{\tau_G : \tau \in M^{\mathbb{P}}\}.$$

We define $\mathrm{dom}(\tau) = \{\sigma : \exists p\,(\langle \sigma, p\rangle \in \tau)\}$. This looks like the usual definition of domain, though note that τ is usually not a function. By absoluteness, the inhabitants of M know $\mathrm{dom}(\tau)$, and they might think of $\mathrm{dom}(\tau)$ as a set of names for objects which may possibly be in τ_G.

Since τ_G was defined by transfinite recursion, it is absolute for transitive models of $ZF-P$ for similar reasons as for $\mathbb{P}$-name. However, the absoluteness of τ_G says nothing for M unless $G \in M$, which will usually be false. Nevertheless, we do have the following fact:

Lemma 9.3.5. *If N is a transitive model for ZFC with $M \subset N$ and $G \in N$, then $M[G] \subset N$.*

Proof. For each $\tau \in M^{\mathbb{P}}$, $\tau \in M$ and hence $\tau \in N$. By the assumption that $G \in N$ and the absoluteness of τ_G in such cases, $\tau_G = \tau_G^N \in N$. $\square_{9.3.5}$

The above lemma will guarantee, once we check that $M[G]$ is in fact a transitive extension of M containing G and satisfying ZFC, that $M[G]$ is the least such extension.

Example 3. Some examples of $\mathbb{P}$-names might be helpful for visualizing the above. To this end, let M be a countable transitive model for ZFC and let $\mathbb{P}$ be a partial order in M.

Note that $\emptyset$ is a $\mathbb{P}$-name since it trivially satisfies Definition 9.3.1. By Definition 9.3.3, $\emptyset_G = \emptyset$ for any G.

If $p \in \mathbb{P}$, then $\{\langle \emptyset, p\rangle\} \in M^{\mathbb{P}}$, and

$$\{\langle \emptyset, p\rangle\}_G = \begin{cases} \{\emptyset\} & \text{if } p \in G, \\ \emptyset & \text{if } p \notin G. \end{cases}$$

By Lemma 9.2.4, there is always a generic G with $p \in G$ and, assuming $\exists q \in \mathbb{P}\,(q \perp p)$, there will be be a generic G with $p \notin G$. Thus, τ_G can depend on the choice of G. However, there are some cases when τ_G is independent of the choice of G. Such a case was given in the example: $\emptyset_G = \emptyset$. Furthermore $\{\langle \emptyset, \mathbb{1}\rangle\}_G = \{\emptyset\}$ for all generic G because any non-empty filter contains $\mathbb{1}$. We can state this more generally:

$$\{\langle \sigma_i, \mathbb{1}\rangle : i \in I\}_G = \{\sigma_{iG} : i \in I\}.$$

So, we see that any element $x \in M$ is represented in a canonical way by a name, which we will call $\check{x}$.

Definition 9.3.6. If $\mathbb{P}$ is a partial order, then we define the (canonical) $\mathbb{P}$-name $\check{x}$ recursively by

$$\check{x} = \{\langle \check{y}, \mathbb{1}_{\mathbb{P}} \rangle : y \in x\}.$$

Formally, the definition of $\check{x}$ depends both on x and on $\mathbb{1}_{\mathbb{P}}$. Nevertheless, the partial order $\langle \mathbb{P}, \leq_{\mathbb{P}}, \mathbb{1}_{\mathbb{P}} \rangle$ will always be clear from context. The definition of the canonical name is another definition by recursion and can be seen to be absolute for transitive models of ZFC. Thus, if $x \in M$ then $\check{x} \in M$.

Examples 1.

- $\check{\emptyset} = \emptyset$,
- $\check{1} = \{\check{\emptyset}\} = \{\langle \emptyset, \mathbb{1} \rangle\}$,
- $\check{2} = \{\langle \check{\emptyset}, \mathbb{1} \rangle, \langle \check{1}, \mathbb{1} \rangle\}$,

As we saw, $\check{\emptyset}_G = \emptyset$ and $\check{1}_G = 1$

——————— HERE ENDED SPRING 2008 LECTURE 3 (90 min) ————

Lemma 9.3.7. *If M is a transitive model for ZFC, $\mathbb{P}$ is a partial order in M, and G is a non-empty filter on $\mathbb{P}$, then*

1. $\forall x \in M\, (\check{x} \in M^{\mathbb{P}} \wedge \check{x}_G = x)$.
2. $M \subset M[G]$.

Proof. For **1**, the absoluteness of ˇ implies $\check{x} \in M^{\mathbb{P}}$. That $\check{x}_G = x$ is proved by induction on x, using

$$\check{x}_G = \{\check{y}_G : y \in x\}.$$

Part **2** follows immediately from 1. $\square_{9.3.7}$

We will now show that $G \in M[G]$ by finding a name that represents G.

Definition 9.3.8. If $\mathbb{P}$ is a partial order, let $\Gamma = \{\langle \check{p}, p \rangle : p \in \mathbb{P}\}$.

Again, Γ depends on $\mathbb{P}$, but this will be clear from context. Unlike the canonical names of the form $\check{x}$, the object named by Γ depends on the choice of G. By absoluteness, Γ is in M if $\mathbb{P}$ is in M.

Lemma 9.3.9. *If M is a transitive model for ZFC, $\mathbb{P}$ is a partial order in M, and G is a non-empty filter on $\mathbb{P}$, then $\Gamma_G = G$. Hence, $G \in M[G]$.*

Proof.

$$\Gamma_G = \{(\check{p})_G : p \in G\} = \{p : p \in G\} = G.$$

$\square_{9.3.9}$

——————— HERE ENDED SPRING 2007 LECTURE 5 (135 min) ————-

Lemma 9.3.10. *If M is a transitive model for ZFC, $\mathbb{P}$ is a partial order in M, and G is a non-empty filter on $\mathbb{P}$, then $M[G]$ is transitive.*

Proof. This follows immediately from Definitions 9.3.3 and 9.3.4. $\square_{9.3.10}$

Lemma 9.3.11. *If M is a transitive model for ZFC, $\mathbb{P}$ is a partial order in M, and G is a non-empty filter on $\mathbb{P}$, then*

1. $\forall \tau \in M^{\mathbb{P}}\,(\mathrm{rank}(\tau_G) \leqslant \mathrm{rank}(\tau))$.
2. $\mathrm{o}(M[G]) = \mathrm{o}(M)$.

Proof. Statement **1** is proved by induction on τ.

For **2**, note that 1 along with the fact that $\mathrm{rank}(\tau) \in M$ for all $\tau \in M$ gives us $M[G] \cap \mathbf{ON} \subset M \cap \mathbf{ON}$. Therefore, $M[G] \cap \mathbf{ON} = M \cap \mathbf{ON}$ since $M \subset M[G]$. $\square_{9.3.11}$

To give some more examples of building names, we will check that $M[G]$ satisfies some of the simpler axioms of ZFC. We start with the Pairing Axiom. To show that Pairing holds, we show that for given $\sigma, \tau \in M^{\mathbb{P}}$, we can define a name $\mathrm{up}(\sigma, \tau) \in M^{\mathbb{P}}$ which always names $\{\sigma_G, \tau_G\}$.

Definition 9.3.12.

1. $\mathrm{up}(\sigma, \tau) = \{\langle \sigma, \mathbb{1}\rangle, \langle \tau, \mathbb{1}\rangle\}$.
2. $\mathrm{op}(\sigma, \tau) = \mathrm{up}(\mathrm{up}(\sigma, \sigma), \mathrm{up}(\sigma, \tau))$.

Lemma 9.3.13. *If M is a transitive model for ZFC, $\mathbb{P}$ is a partial order in M, and G is a non-empty filter on $\mathbb{P}$, then*

1. $\mathrm{up}(\sigma, \tau) \in M^{\mathbb{P}}$ *and* $\mathrm{up}(\sigma, \tau)_G = \{\sigma_G, \tau_G\}$.
2. $\mathrm{op}(\sigma, \tau) \in M^{\mathbb{P}}$ *and* $\mathrm{op}(\sigma, \tau)_G = \langle \sigma_G, \tau_G\rangle$.

Lemma 9.3.14. *If M is a transitive model for ZFC, $\mathbb{P}$ is a partial order in M, and G is a non-empty filter on $\mathbb{P}$, then the Axioms of Extensionality, Foundation, Pairing, and Union are true in $M[G]$.*

Proof. Extensionality holds because $M[G]$ is transitive. Foundation is true relativized to any class. Pairing is immediate by the previous lemma.

For Union, it is enough to show that if $a \in M[G]$, then there is a $b \in M[G]$ such that $\bigcup a \subseteq b$. To this end, fix $\tau \in M^{\mathbb{P}}$ such that $a = \tau_G$. Let $\pi = \bigcup \mathrm{dom}(\tau)$. Then, $\pi \in M^{\mathbb{P}}$, so $b = \pi_G \in M[G]$. If c is any element of a, $c = \sigma_G$ for some $\sigma \in \mathrm{dom}(\tau)$. Since $\sigma \subseteq \pi$, $c = \sigma_G \subset \pi_G = b$. Thus, $\bigcup a \subseteq b$. $\square_{9.3.14}$

In the above proof we did not show that $\bigcup a \in M[G]$. This will follow once we show that $M[G]$ satisfies the Comprehension Axiom. So far, we did not use any of the particular properties of the generic; we will need those later when we define the notion of forcing in the next section.

The following are a couple of facts that will be useful later.

Definition 9.3.15. If $E \subset \mathbb{P}$, and $p \in \mathbb{P}$, then E is *dense below* p iff

$$\forall q \leq p\, \exists r \leq q\, (r \in E).$$

Lemma 9.3.16. *Assume that M is a transitive model for ZFC, $\mathbb{P} \in M$, $E \subset \mathbb{P}$, and $E \in M$. Let G be $\mathbb{P}$-generic over M. Then,*

1. *Either $G \cap E \neq \emptyset$, or $\exists q \in G\, \forall r \in E\, (r \perp q)$.*

2. *If $p \in G$ and E is dense below p, then $G \cap E \neq \emptyset$.*

Proof. For statement **1**, let

$$D = \{p : \exists r \in E\,(p \leq r)\} \cup \{q : \forall r \in E\,(r \perp q)\}.$$

The set D is dense, since if $q \in \mathbb{P}$ and $q \notin D$, then we can fix $r \in E$ such that r and q are compatible. If $p \leq r$ and $p \leq q$, then p is an extension of q in D. Thus, $G \cap D \neq \emptyset$, which implies statement 1.

For statement **2**: if $G \cap E = \emptyset$, then by 1 we can fix $q \in G$ such that $\forall r \in E\,(r \perp q)$. Let $q' \in G$ be such that $q' \leq q$ and $q' \leq p$. Then, since E is dense below p, let $r \in E$ be such that $r \leq q'$. Then $r \leq q$, which contradicts $r \perp q$. $\square_{9.3.16}$

——————— HERE ENDED SPRING 2008 LECTURE 4 (90 min) ————

9.4 Forcing

We gave a name Γ for G. We start this section by demonstrating a name for an object constructed from G, for a specific partial order.

Example 4. Fix a countable transitive model M for ZFC, and let $\mathbb{P}$ be the set of finite partial functions from ω to $\{0, 1\}$ ordered by reverse inclusion (so, the usual crazy backward ordering of finite bits of functions). Thus, $\mathbb{1}_{\mathbb{P}}$ is the empty function. The partial ordering $\langle \mathbb{P}, \leq, \mathbb{1} \rangle \in M$, since its definition is absolute for transitive models of $ZF - P$.

If G is a filter on $\mathbb{P}$, $f^G = \bigcup G$ is a function with $\mathrm{dom}(f^G) \subseteq \omega$. For each n, we let $D_n = \{p \in \mathbb{P} : n \in \mathrm{dom}(p)\}$. Then, D_n is dense and $D_n \in M$ by the absoluteness of its definition. Thus, if G is $\mathbb{P}$-generic over M, then $G \cap D_n \neq \emptyset$ for all n. So, $\mathrm{dom}(f^G) = \omega$.

To show that $f^G \in M[G]$. This could be done in a couple of ways. Firstly, since $G \in M[G]$, and $f^G = \bigcup G$, $f^G \in M[G]$ would follow immediately from the absoluteness of $\bigcup$ for transitive models of ZF once we determine that $M[G]$ is a model of enough of $ZF - P$ to determine the absoluteness of $\bigcup$. The second way would be to find a name for f^G. Let

$$\Phi = \{\langle \langle n, m\rangle\check{}, p \rangle : p \in \mathbb{P} \wedge n \in \mathrm{dom}(p) \wedge p(n) = m\}.$$

Since $\langle n, m\rangle\check{}_G = \langle n, m \rangle$, we have

$$\Phi_G = \{\langle n, m \rangle : \exists p \in G\,(n \in \mathrm{dom}(p) \wedge p(n) = m)\} = f^G.$$

Thus, $f^G \in M[G]$.

In the previous section, we had the intuitive idea that elements $p \in \mathbb{P}$ were conditions which say something about G, or some object that we plan to construct from G. The inhabitants of M cannot construct a G which is $\mathbb{P}$-generic over M. They could have some superstition featuring a great mythical being to whom their universe M is countable. This mythical being will have a generic G and a function $f^G = \bigcup G$. The inhabitants of M do not know what G and f^G are, but they do have names of them: Γ and Φ. The superstitious inhabitants of M might also be able to deduce some of the properties of G and f^G discussed in the past couple of paragraphs: for example, that f^G is a function from ω to $\{0, 1\}$. They will not be able to figure out what $f^G(0)$ is because that value depends on the choice of G. However, they would be able to figure out that $f^G(0) = 0$ if $\{\langle 0, 0\rangle\} \in G$ and $f^G(0) = 1$ if $\{\langle 0, 1\rangle\} \in G$.

More generally, the superstitious inhabitants of M can construct a *forcing language* to discuss their myths: a sentence ψ of the forcing language uses the names in $M^{\mathbb{P}}$ to assert something about $M[G]$. An example of such a ψ is $\Phi(\check{0}) = \check{1}$. An inhabitant of M can state such a sentence, but cannot know if a given ψ is true in $M[G]$, since the truth or falsity of ψ generally depends on G.

We will write $p \Vdash \psi$ (in words: *p forces* ψ) to mean that for all G which are $\mathbb{P}$-generic over M, if $p \in G$ then ψ is true in $M[G]$. For example,

$$\{\langle 0, 0\rangle\} \Vdash \Phi(\check{0}) = \check{0},$$

and

$$\{\langle 0, 1\rangle\} \Vdash \Phi(\check{0}) = \check{1}.$$

Also,

$$\mathbb{1} \Vdash \Phi \text{ is a function from } \check{\omega} \text{ into } \check{2},$$

and

$$\mathbb{1} \Vdash \Phi = \bigcup \Gamma;$$

that is, these last two sentences are true for all generic G.

The superstitious inhabitants of M can figure out all the above forcing facts without seeing a generic G. This illustrates the following fact:

Fact. It may be decided within M whether or not $p \Vdash \psi$.

This fact will be very important not only for proving that $M[G]$ satisfies ZFC, but also for applying forcing later, since the inhabitants of M will have to be able to apply their combinatorial techniques to construct various complicated $\mathbb{P}$ for which the desired axioms of set theory beyond ZFC are forced to be true in $M[G]$.

The fact may seem surprising, since the notion $p \Vdash \psi$ seems to require a knowledge of all generic G. However, a superstitious M inhabitant can always decide whether $p \Vdash \psi$ by going through the same kind of analysis as was presented in our examples.

It is immediate from the definition of $\Vdash$ that if G is $\mathbb{P}$-generic over M and $p \Vdash \psi$ for some $p \in G$, then ψ is true in $M[G]$. We shall also show this converse:

Fact. If G is $\mathbb{P}$-generic over M and ψ is true in $M[G]$, then for some $p \in G$, $p \Vdash \psi$.

For example, if ψ is the sentence $\Phi(\check{0}) = \check{0}$ and ψ is true (that is, $f^G(0) = 0$), then $p(0) = 0$ for some $p \in G$. Thus, if $p \in H$, where H is some other generic filter, then $f^H(0) = 0$ also – that is, ψ will be true in $M[H]$. Thus $p \Vdash \psi$.

Now we turn away from the specific example and back to the general and rigorous. Theorem 9.4.6 will express the two facts, and is the main theorem for the forcing techniques.

Definition 9.4.1. Let $\phi(x_1, \ldots, x_n)$ be a formula with all free variables listed. Let M be a countable transitive model for ZFC, $\mathbb{P}$ a partial order in M, $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$, and $p \in \mathbb{P}$. Then, $p \Vdash_{\mathbb{P},M} \phi(\tau_1, \ldots, \tau_n)$ iff

$$\forall G\,((G \text{ is } \mathbb{P}\text{-generic over } M \wedge p \in G) \to \phi^{M[G]}(\tau_{1G}, \ldots \tau_{nG})). \tag{9.1}$$

We will usually just write $\Vdash$ when there is only one partial order and one ground model M under consideration.

Intuitively, the $\phi(\tau_1, \ldots, \tau_n)$ in the above definition is a sentence of the forcing language. The idea of the forcing language could be made rigorous by formalizing logic within set theory and defining the forcing language to be the first-order language whose one binary relation symbol is $\in$, and whose constant symbols are the elements of $M^{\mathbb{P}}$. However, we will take a different approach and won't define a forcing language. Instead, note that the above definition is a definition schema in the metatheory. For each formula $\phi(x_1, \ldots, x_n)$ with free variable listed, we can define another formula $Force_\phi(\tau_1, \ldots, \tau_n, \mathbb{P}, \leq, \mathbb{1}, M, p)$, which asserts 9.1, along with the statements $\langle \mathbb{P}, \leq, \mathbb{1} \rangle \in M$, $p \in \mathbb{P}$, and $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$.

To illustrate:

Lemma 9.4.2.

1. $(p \Vdash \phi(\tau_1, \ldots, \tau_n) \wedge q \leq p) \rightarrow q \Vdash \phi(\tau_1, \ldots, \tau_n)$.
2. $(p \Vdash \phi(\tau_1, \ldots, \tau_n) \wedge p \Vdash \psi(\tau_1, \ldots, \tau_n)) \leftrightarrow p \Vdash (\phi(\tau_1, \ldots, \tau_n) \wedge \psi(\tau_1, \ldots, \tau_n))$

——————— HERE ENDED SPRING 2007 LECTURE 6 (135 min) ———-

Notice that the notion "$p \Vdash \phi(\tau_1, \ldots, \tau_n)$" has been defined in $\mathbf{V}$ and not in M, and involves a knowledge of all possible generic G. Our first fact tells us that we should be able to decide within M whether $p \Vdash \phi(\tau_1, \ldots, \tau_n)$: we do this rigorously be defining another relation $p \Vdash^* \phi(\tau_1, \ldots, \tau_n)$ and showing that for all ϕ,

$$p \Vdash \phi(\tau_1, \ldots, \tau_n) \leftrightarrow (p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M.$$

Thus, $p \Vdash \phi(\tau_1, \ldots, \tau_n)$ will be equivalent to some statement relativized to M.

After this section, we will rarely refer back to the details of the definition of $\Vdash^*$, although we will frequently use its results: the two facts (Theorem 9.4.6) and Corollary following. I also mention that there are heaps of different but equivalent definitions of $\Vdash^*$. The one we have is as defined in Kunen's book (the source for this exposition of forcing).

We define, in $\mathbf{V}$, the notion $p \Vdash^* \phi(\tau_1, \ldots, \tau_n)$. This definition does not mention any model. In practice, we will only consider the relativized notion $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M$, where M is the ground model.

Definition 9.4.3. Fix a partial order $\mathbb{P}$. The following clauses define the notion $p \Vdash^* \phi(\tau_1, \ldots, \tau_n)$ where $\phi(x_1, \ldots, x_n)$ is a formula with all free variables listed, $p \in \mathbb{P}$, and $\tau_1, \ldots, \tau_n \in \mathbf{V}^{\mathbb{P}}$.

1. $p \Vdash^* (\tau_1 = \tau_2)$ iff
   (a) for all $\langle \pi_1, s_1 \rangle \in \tau_1$,
   $$\{q \leq p : q \leq s_1 \rightarrow \exists \langle \pi_2, s_2 \rangle \in \tau_2\, (q \leq s_2 \wedge q \Vdash^* (\pi_1 = \pi_2))\}$$
   is dense below p, and
   (b) for all $\langle \pi_2, s_2 \rangle \in \tau_2$,
   $$\{q \leq p : q \leq s_2 \rightarrow \exists \langle \pi_1, s_1 \rangle \in \tau_1\, (q \leq s_1 \wedge q \Vdash^* (\pi_1 = \pi_2))\}$$
   is dense below p.

2. $p \Vdash^* (\tau_1 \in \tau_2)$ iff

$$\{q \leq p : \exists \langle \pi, s \rangle \in \tau_2 \, (q \leq s \wedge q \Vdash^* (\pi = \tau_1))\}$$

is dense below p.

3. $p \Vdash^* (\phi(\tau_1, \ldots, \tau_n) \wedge \psi(\tau_1, \ldots, \tau_n))$ iff

$$p \Vdash^* \phi(\tau_1, \ldots, \tau_n) \text{ and } p \Vdash^* \psi(\tau_1, \ldots, \tau_n).$$

4. $p \Vdash^* \neg\phi(\tau_1, \ldots, \tau_n)$ iff there is no $q \leq p$ such that $q \Vdash^* \phi(\tau_1, \ldots, \tau_n)$.

5. $p \Vdash^* \exists x \phi(x, \tau_1, \ldots, \tau_n)$ iff

$$\{r \leq p : \exists \sigma \in \mathbf{V}^{\mathbb{P}} \, (r \Vdash^* \phi(\sigma, \tau_1, \ldots, \tau_n))\}$$

is dense below p.

Oddly, the most difficult part of the above definition is clause 1: when $\phi(\tau_1, \tau_2)$ is $\tau_1 = \tau_2$. To give an idea of the motivation behind the definition, let us consider an example of a forcing situation, and what an M person can deduce about the situation in his terms, with view to his limitations. To this end, suppose that $\tau_1 = \{\langle \pi_1, s \rangle\}$ and $\tau_2 = \{\langle \pi_2, s \rangle\}$, and we are trying to tell an inhabitant of M which p force $\tau_1 = \tau_2$. There are a couple of possibilities. If $p \perp s$, then $p \Vdash \tau_1 = \tau_2$, since whenever $p \in G$, $s \notin G$, so $\tau_{1G} = \emptyset = \tau_{2G}$. On the other hand, if $p \leq s$, then whenever $p \in G$, $\tau_{1G} = \{\pi_{1G}\}$ and $\tau_{2G} = \{\pi_{2G}\}$, so $p \Vdash \tau_1 = \tau_2$ iff $p \Vdash \pi_1 = \pi_2$. One can then check that for any p, $p \Vdash \tau_1 = \tau_2$ iff

$$\forall q \, (q \leq p \wedge q \leq s \rightarrow q \Vdash \pi_1 = \pi_2).$$

This helps to explain why clause 1 has the form it does. We emphasize that in the definition of $\Vdash^*$, the question of whether $p \Vdash^* \tau_1 = \tau_2$ has to depend on whether $q \Vdash^* \pi_1 = \pi_2$ for various $q \in \mathbb{P}$, $\pi_1 \in \text{dom}(\tau_1)$, $\pi_2 \in \text{dom}(\tau_2)$.

——————— HERE ENDED SPRING 2008 LECTURE 5 (90 min) ———

Definition 9.4.3 looks circular, so is a recursion. The intention is that clause 1 is applied first to define the notion $p \Vdash^* \tau_1 = \tau_2$. Formally, we are defining a function $\mathbf{F} : \mathbf{V}^{\mathbb{P}} \times \mathbf{V}^{\mathbb{P}} \longrightarrow \mathscr{P}(\mathbb{P})$, where $\mathbf{F}(\langle \tau_1, \tau_2 \rangle)$ is intended to be

$$\{p \in \mathbb{P} : p \Vdash^* \tau_1 = \tau_2\}.$$

This function $\mathbf{F}$ is defined by transfinite recursion on the relation $\mathbf{R}$, where

$$\langle \pi_1, \pi_2 \rangle \mathbf{R} \langle \tau_1, \tau_2 \rangle$$

iff $\pi_1 \in \text{dom}(\tau_1)$ and $\pi_2 \in \text{dom}(\tau_2)$. The relation $\mathbf{R}$ is clearly set-like, and $\mathbf{R}$ is well-founded because $\langle \pi_1, \pi_2 \rangle \mathbf{R} \langle \tau_1, \tau_2 \rangle$ implies $\text{rank}(\pi_1) < \text{rank}(\tau_1)$ and $\text{rank}(\pi_2) < \text{rank}(\tau_2)$.

The rest of the induction is straightforward, on the length of the formula. The induction takes place in the metatheory. Formally, as with $\Vdash$, for each formula $\phi(x_1, \ldots, x_n)$, we are defining a formula

$$Force^*_\phi(\tau_1, \ldots, \tau_n, \mathbb{P}, \leq, \mathbb{1}, M, p).$$

For atomic formulas, the recursion defining $\Vdash^*$ uses only absolute concepts and is thus absolute for transitive models of $ZF - P$. To be precise, we are using the absoluteness of the relation $\mathbf{R}$ and the absoluteness of $\{\langle\pi_1, \pi_2\rangle : \langle\pi_1, \pi_2\rangle \mathbf{R} \langle\tau_1, \tau_2\rangle\}$ to get the absoluteness of $\mathbf{F}$. However, for ϕ arbitrary (non-atomic), $\Vdash^*$ is not absolute: the $\exists\sigma \in \mathbf{V}^{\mathbb{P}}$ in clause 5 becomes $\exists\sigma \in M^{\mathbb{P}}$ when relativized to a model M. When we check the first Fact, we will only be interested in $\Vdash^*$ relativized to M.

Some words on the motivation for the details of clauses 1-5: One should think of $(p \Vdash^* \phi)^M$ as an attempt by an inhabitant of M to decide $\Vdash$. We will try to prove the first Fact, that $\Vdash$ is definable in M, by showing that $p \Vdash \phi$ iff $(p \Vdash^* \phi)^M$. Thus, we use, as in the inductive clauses in the definition of $\Vdash^*$, relations which $\Vdash$ itself satisfies. We can then try to prove the first Fact by induction on ϕ.

To see that $\Vdash$ satisfies analogous clauses to clauses 1-5 may need some argument. On the one hand, for clause 3 it is immediate from Lemma 9.4.2 that $p \Vdash (\phi \wedge \psi)$ iff $p \Vdash \phi$ and $p \Vdash \psi$. However, for clause 4, as an example, we have to work a bit harder: first assume that $\neg\exists q \leq p\,(q \Vdash \phi)$ and we wish to show that $p \Vdash \neg\phi$. Assume the contrary. Then there is a generic G with $p \in G$ and ϕ true in $M[G]$. By the second Fact, there is an $r \in G$ such that $r \Vdash \phi$. Let $q \in G$ with $q \leq r$ and $q \leq p$. Then, $q \Vdash \phi$ by Lemma 9.4.2, contradicting our assumption.

Clause 5, relativized to M, says $(p \Vdash^* \exists x\, \phi(x))^M$ iff

$$\{r \leq p : \exists\sigma \in M^{\mathbb{P}}\, (r \Vdash^* \phi(\sigma))^M\}$$

is dense below p. To check the analogous statement for $\Vdash$, suppose $D = \{r \leq p : \exists\sigma \in M^{\mathbb{P}}\, (r \Vdash^* \phi(\sigma))^M\}$ is dense below p. By the first Fact, $D \in M$. Thus, whenever G is generic over M and $p \in G$, $G \cap D \neq \emptyset$. So there is a $\sigma \in M^{\mathbb{P}}$ and $r \in G$ with $r \Vdash \phi(\sigma)$. Thus, $(\phi(\sigma_G))^{M[G]}$, so $(\exists x\, \phi(x))^{M[G]}$. Thus, $p \Vdash \exists x\, \phi(x)$.

The motivation above was very circular, it is true. However, hopefully it does give a sense of how these definitions are meant to fit together. We now return to our more rigorous proofs of the two Facts.

Lemma 9.4.4. *For p and $\phi(\tau_1, \ldots, \tau_n)$ as in Definition 9.4.3, the following are equivalent:*

1. $p \Vdash^* \phi(\tau_1, \ldots, \tau_n)$.
2. $\forall r \leq p\, (r \Vdash^* \phi(\tau_1, \ldots, \tau_n)$.
3. $\{r : r \Vdash^* \phi(\tau_1, \ldots, \tau_n)\}$ *is dense below* p.

Proof. **2 ⇒ 1:** trivial from the definition.

2 ⇒ 3: also trivial from the definition.

We first show the remaining implications for atomic formulas.

1 ⇒ 2 for atomic ϕ: Assume that $\phi(\tau_1, \tau_2)$ is either $\tau_1 = \tau_2$ or $\tau_1 \in \tau_2$. If D is dense below p and $r \leq p$, then D is dense below r as well. The rest follows from the definition.

3 ⇒ 1 for atomic ϕ: Note that for a set D if the set $E = \{r : D$ is dense below $r\}$ is dense below p, then D is dense below p.

The rest follows by induction using the definition of $\Vdash^*$.

$\square_{9.4.4}$

It should be noted that Lemma 9.4.4 gives properties that hold for $\Vdash^*$, but not for $\Vdash$. This is an important difference between the two notions.

Theorem 9.4.5. *Let $\phi(x_1, \ldots, x_n)$ be a formula with all free variables listed. Let M be a countable transitive model for ZFC, $\mathbb{P}$ a partial order in M, and $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$. Let G be $\mathbb{P}$-generic over M. Then*

1. *if $p \in G$ and $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M$, then $(\phi(\tau_{1G}, \ldots, \tau_{nG}))^{M[G]}$.*
2. *If $(\phi(\tau_{1G}, \ldots, \tau_{nG}))^{M[G]}$, then $\exists p \in G\,((p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M)$.*

please get comfortable. The proof is a loooong one.

Proof. We proceed by induction on the complexity of formulas.

- **Assume ϕ is atomic.**
  - **Assume $\phi(\tau_1, \tau_2)$ is $\tau_1 = \tau_2$.** We prove both statements of the theorem for this case using transfinite induction, using the definition of $\Vdash^*$ for formulas of the form $\tau_1 = \tau_2$. That this really is an induction on a well-founded relation can be established in exactly the same way that we justified the definition of $\Vdash^*$ for such ϕ thus it goes vaguely along the rank of the names. Since $\Vdash^*$ for atomic formulas is absolute for M, we will be lazy and not write the relativizations to M.
    * **Statement 1:** Assume $p \in G$ and $p \Vdash^* \tau_1 = \tau_2$, and assume we have established (1) already for names of "lower rank". We aim to show that $\tau_{1G} = \tau_{2G}$.
    We show that $\tau_{1G} \subseteq \tau_{2G}$ using clause 1a of Definition 9.4.3. Every element of τ_{1G} is of the form π_{1G}, where $\langle \pi_1, s_1 \rangle \in \tau_1$ for some $s_1 \in G$. We have to show that $\pi_{1G} \in \tau_{2G}$. Fix $r \in G$ such that $r \le p$ and $r \le s_1$. Then, by Lemma 9.4.4, $r \Vdash^* \tau_1 = \tau_2$, so by Lemma 9.3.16 (2), there is $q \in G$ such that $q \le r$, and such that if $q \le s_1$ then
    $$\exists \langle \pi_2, s_2 \rangle \in \tau_2\,(q \le s_2 \wedge q \Vdash^* \pi_1 = \pi_2). \tag{9.2}$$
    However, we know that $q \le r \le s_1$, so fix $\langle \pi_2, s_2 \rangle$ as in 9.2. Then, $s_2 \in G$, so $\pi_{2G} \in \tau_{2G}$. Now, by the inductive assumption as applied to $\pi_1 = \pi_2$, $q \Vdash^* \pi_1 = \pi_2$ implies $\pi_{1G} = \pi_{2G}$, thus $\pi_{1G} \in \tau_{2G}$.
    That $\tau_{2G} \subseteq \tau_{1G}$ is proved similarly, using clause 1b of Definition 9.4.3.

    ——— HERE ENDED SPRING 2008 LECTURE 6 (90 min) ———

    * **Statement 2:** Assume $\tau_{1G} = \tau_{2G}$. Let D be the set of all $r \in \mathbb{P}$ such that either
    $$r \Vdash^* \tau_1 = \tau_2, \tag{9.3}$$
    or
    $$\exists \langle \pi_1, s_1 \rangle \in \tau_1\,(r \le s_1 \wedge \forall \langle \pi_2, s_2 \rangle \in \tau_2\, \forall q \in \mathbb{P}\; ((q \le s_2 \wedge q \Vdash^* \pi_1 = \pi_2) \to q \perp r)), \tag{9.4}$$

or

$$\exists\langle\pi_2, s_2\rangle \in \tau_2\,(r \le s_2 \wedge \forall\langle\pi_1, s_1\rangle \in \tau_1\,\forall q \in \mathbb{P}\\ ((q \le s_1 \wedge q \Vdash^* \pi_1 = \pi_2) \to q \perp r)). \quad (9.5)$$

note the similarity between 9.4 and 9.5 and clauses 1a and 1b of Definition 9.4.3, respectively.

Note first that no $r \in G$ can satisfy 9.4 or 9.5: Suppose to the contrary, that $r \in G$ and, for example, $\langle\pi_1, s_1\rangle \in \tau_1$ as in 9.4. Then, $s_1 \in G$, so $\pi_{1G} \in \tau_{1G} = \tau_{2G}$. Thus fix $\langle\pi_2, s_2\rangle \in \tau_2$ with $s_2 \in G$ and $\pi_{1G} = \pi_{2G}$. Then, by the inductive hypothesis, (2) applies to $\pi_1 = \pi_2$. So, fix $q_0 \in G$ with $q_0 \Vdash^* \pi_1 = \pi_2$. Now, fix $q \in G$ with $q \le q_0$ and $q \le s_2$. Since by Lemma 9.4.4 $q \Vdash^* \pi_1 = \pi_2$, we have $q \perp r$ (by 9.4), $q \in G$, and $r \in G$, a contradiction.

Now, if $\neg\exists r \in G\,(r \Vdash^* \tau_1 = \tau_2)$, then $D \cap G = \emptyset$. Since $D \in M$ by absoluteness, once we check that D is dense, this case of the proof will be complete.

To show that D is dense, fix $p \in \mathbb{P}$. Then, either $p \Vdash^* \tau_1 = \tau_2$, or Definition 9.4.3 1a or 1b fails. If Definition 9.4.3 1a fails, then by the definition of "dense below p", fix $\langle\pi_1, s_1\rangle \in \tau_1$ and $r \le p$ such that

$$\forall q \le r\,(q \le s_1 \wedge \forall\langle\pi_2, s_2\rangle \in \tau_2\,(\neg(q \Vdash^* \pi_1 = \pi_2 \wedge q \le s_2))). \quad (9.6)$$

So, in particular, we can ensure $r \le s_1$ by improving our choice of r with an extension if necessary (here a picture might be of help). If $\langle\pi_2, s_2\rangle \in \tau_2$, $q \le s_2$, and $q \Vdash^* \pi_1 = \pi_2$, then $q \perp r$ since a common extension q' of q and r would contradict 9.6. Thus, $r \le p$ and r satisfies 9.4. Similarly, if Definition 9.4.3 1b fails, then there is $r \le p$ that satisfies 9.5.

———————— HERE ENDED SPRING 2007 LECTURE 7 (135 min) ————-

– **Assume $\phi(\tau_1, \tau_2)$ is $\tau_1 \in \tau_2$.** Again, remember that we are using induction here on the names.

  * **Statement 1:** Assume $p \in G$ and $p \Vdash^* \tau_1 \in \tau_2$. Then

    $$D = \{q : \exists\langle\pi, s\rangle \in \tau_2\,(q \le s \wedge q \Vdash^* \pi = \tau_1)\}$$

    is dense below p. So, fix $q \in G \cap D$, and fix $\langle\pi, s\rangle \in \tau_2$ so that $q \le s$ and $q \Vdash^* \pi = \tau_1$. Since $s \in G$ and $\langle\pi, s\rangle \in \tau_2$, $\pi_G \in \tau_{2G}$ by the definition of τ_{2G}. Since $q \in G$ and $q \Vdash^* \pi = \tau_1$, $\pi_G = \tau_{1G}$ by the inductive hypothesis that we can apply statement 1 to $\pi = \tau_1$. Thus $\tau_{1G} \in \tau_{2G}$.
  * **Statement 2:** Assume $\tau_{1G} \in \tau_{2G}$. By definition of τ_{2G}, there is a $\langle\pi, s\rangle \in \tau_2$ such that $s \in G$ and $\pi_G = \tau_{1G}$. By the inductive hypothesis that we can apply Statement 2 to $\pi = \tau_1$, there is an $r \in G$ such that $r \Vdash^* \pi = \tau_1$. Let $p \in G$ be such that $p \le s$ and $p \le r$. Then $\forall q \le p\,(q \le s \wedge q \Vdash^* \pi = \tau_1)$. Thus, $p \Vdash^* \tau_1 \in \tau_2$. (Note that we have proved more than necessary: a statement that is stronger than called for by Definition 9.4.3 2.)

- **Assume ϕ is not atomic:** We will assume that both statements 1 and 2 hold for ϕ (and ψ), and will show that these statements hold for formulas with ϕ as a subformula as shown. Formally, this induction takes place in the metatheory. Also, since $\Vdash^*$ is not absolute for formulas involving quantifiers, we must explicitly relativize to M. Also, out of laziness, we will not make explicit mention of $\tau_1, \ldots, \tau_n$. It should be clear where they should go.
  - $\neg\phi$:
    - **Statement 1:** Assume statements 1 and 2 hold for ϕ. We wish to show that statement 1 holds for $\neg\phi$.
      Assume $p \in G$ and $(p \Vdash^* \neg\phi)^M$. We wish to show that $\neg\phi^{M[G]}$. If $\phi^{M[G]}$, then since statement 2 holds for ϕ, there is $q \in G$ such that $(q \Vdash^* \phi)^M$. Let $r \in G$ such that $r \leq p$ and $r \leq q$. Then $(r \Vdash^* \phi)^M$, contradicting the definition of $p \Vdash^* \neg\phi$.
    - **Statement 2:** Assume $\neg\phi^{M[G]}$. Let
      $$D = \{p : (p \Vdash^* \phi)^M \vee (p \Vdash^* \neg\phi)^M\}.$$
      Then, $D \in M$, and D is dense by the definition of $\Vdash^*$ applied within M. Thus we can fix $p \in D \cap G$. There are two possibilities: either $(p \Vdash^* \neg\phi)^M$ (which is what we want), or $(p \Vdash^* \phi)^M$. If we are in the latter case, we have a contradiction, since statement 1 is assumed to hold for ϕ and so the latter case implies that $\phi^{M[G]}$.
  - $\phi \wedge \psi$:
    - **Statement 1:** Assume statements 1 and 2 hold for both ϕ and ψ. Assume $p \in G$ and $(p \Vdash^* (\phi \wedge \psi))^M$. Then, $(p \Vdash^* \phi)^M$ and $(p \Vdash^* \psi)^M$, so $\phi^{M[G]}$ and $\psi^{M[G]}$, and so $(\phi \wedge \psi)^{M[G]}$.
    - **Statement 2:** Assume $(\phi \wedge \psi)^{M[G]}$. By the assumption that statement 2 holds for ϕ and ψ, there are $p, q \in G$ such that $(p \Vdash^* \phi)^M$ and $(q \Vdash^* \psi)^M$. Let $r \in G$ be such that $r \leq p$ and $r \leq q$. Then $(r \Vdash^* \phi)^M$ and $(r \Vdash^* \psi)^M$, so $(r \Vdash^* (\phi \wedge \psi))^M$.
  - $\exists \mathbf{x}\, \phi(\mathbf{x})$:
    - **Statement 1:** Assume $p \in G$ and $(p \Vdash^* \exists x\, \phi(x))^M$. Then
      $$\{r : \exists \sigma \in M^{\mathbb{P}}\, (r \Vdash^* \phi(\sigma))^M\}$$
      is dense below p and in M. Thus, we can fix such an $r \in G$ and $\sigma \in M^{\mathbb{P}}$ with $(r \Vdash^* \phi(\sigma))^M$. By the assumption that statement 1 holds for ϕ, $(\phi(\sigma_G))^{M[G]}$, so $(\exists x\, \phi(x))^{M[G]}$.
    - **Statement 2:** Assume $(\exists x\, \phi(x))^{M[G]}$ and fix $\sigma \in M^{\mathbb{P}}$ such that $(\phi(\sigma_G))^{M[G]}$. By the assumption that statement 2 holds for ϕ, we can fix $p \in G$ such that $(p \Vdash^* \phi(\sigma))^M$. Then, $\forall r \leq p\, ((r \Vdash^* \phi(\sigma))^M)$, so $(p \Vdash^* \exists x\, \phi(x))^M)$.

$\square_{9.4.5}$

——————— HERE ENDED SPRING 2008 LECTURE 7 (90 min) ————

We can now formally state, and prove, the main forcing theorem:

Theorem 9.4.6. *[Main Forcing Theorem] Let M be a countable transitive model for ZFC and $\mathbb{P}$ a partial order in M. Let $\phi(x_1, \ldots, x_n)$ be a formula with all free variable shown. Let $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$.*

1. *For all $p \in \mathbb{P}$,*

$$p \Vdash \phi(\tau_1, \ldots, \tau_n) \iff (p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M.$$

2. *For all G which are $\mathbb{P}$-generic over M,*

$$(\phi(\tau_{1G}, \ldots, \tau_{nG}))^{M[G]} \iff \exists p \in G\,(p \Vdash \phi(\tau_1, \ldots, \tau_n)).$$

Proof.

- **Statement 1:**
  - $\Leftarrow$: This is immediate from Statement 1 of Theorem 9.4.5 and the definition of $\Vdash$.
  - $\Rightarrow$: Assume $p \Vdash \phi(\tau_1, \ldots, \tau_n)$. To show that $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M$, it is enough to show, by Lemma 9.4.4, to show that $D = \{r : (r \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M\}$ is dense below p. Assume otherwise. Let $q \leq p$ be such that $\neg\exists r \leq q\,(r \in D)$. Then, by the definition of $\Vdash^*$,

    $$(q \Vdash^* \neg\phi(\tau_1, \ldots, \tau_n))^M.$$

    Thus, by Statement 1 of this theorem, implication $\Leftarrow$, $q \Vdash \neg\phi(\tau_1, \ldots, \tau_n)$. Let G be $\mathbb{P}$-generic over M with $q \in G$. Then $(\neg\phi(\tau_{1G}, \ldots, \tau_{nG}))^{M[G]}$. But it is also true that $p \in G$, since $p \geq q$, so $\phi(\tau_{1G}, \ldots, \tau_{nG}))^{M[G]}$, a contradiction.
- **Statement 2:**
  - $\Rightarrow$: This follows from Statement 1 of this theorem, and from Statement 2 of Theorem 9.4.5, which says something analogous about $\Vdash^*$.
  - $\Leftarrow$: This is immediate from the definition of $\Vdash$.

$\square_{9.4.6}$

In practice, we will use the first statement of Theorem 9.4.6 to show that various sets defined using $\Vdash$ actually lie in M.

Example. 1. For fixed $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$,

$$\{p \in \mathbb{P} : p \Vdash \phi(\tau_1, \ldots, \tau_n)\}$$

is in M, since this set is equal to

$$\{p \in \mathbb{P} : (p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M\}.$$

This latter set is in M by the fact that the Axiom of Comprehension holds in M.

2. For fixed $\sigma, \tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$,

$$\{\langle p, \tau_1\rangle \in \mathbb{P} \times \mathrm{dom}(\sigma) : p \Vdash \phi(\tau_1, \ldots, \tau_n)\} \in M,$$

using a similar argument to that above.

The second statement of Theorem 9.4.6 is important because it relates truth in $M[G]$ to $\Vdash$. Some further facts about $\Vdash$ will be useful.

Corollary 9.4.7. *Let M be a countable transitive model for ZFC, $\mathbb{P}$ a partial order in M, and $\sigma, \tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$. Then*

1. $\{p \in \mathbb{P} : (p \Vdash \phi(\tau_1, \ldots, \tau_n)) \vee (p \Vdash \neg\phi(\tau_1, \ldots, \tau_n))\}$ *is dense.*
2. $p \Vdash \neg\phi(\tau_1, \ldots, \tau_n)$ *iff* $\neg\exists q \leq p\, (q \Vdash \phi(\tau_1, \ldots, \tau_n))$.
3. $p \Vdash \exists x\, \phi(x, \tau_1, \ldots, \tau_n)$ *iff*

$$\{r \leq p : \exists \sigma \in M^{\mathbb{P}}\, (r \Vdash \phi(\sigma, \tau_1, \ldots, \tau_n))\}$$

is dense below p.

4. *If* $p \Vdash \exists x\, (x \in \sigma \wedge \phi(x, \tau_1, \ldots, \tau_n))$, *then*

$$\exists q \leq p\, \exists \pi \in \mathrm{dom}(\sigma)\, (q \Vdash \phi(\pi, \tau_1, \ldots, \tau_n)).$$

Proof. Statements 1-3 are true of $\Vdash^*$ by definition, and so hold for $\Vdash$ by Theorem 9.4.6, Statement 1.

For Statement 4, fix a generic G with $p \in G$. By definition of $\Vdash$, there is an $a \in \sigma_G$ such that $(\phi(a, \tau_1, \ldots, \tau_n))^{M[G]}$. Furthermore, $a = \pi_G$ for some $\pi \in \mathrm{dom}(\sigma)$. By Statement 2 of Theorem 9.4.6, there is an $r \in G$ such that $r \Vdash \phi(\pi, \tau_1, \ldots, \tau_n)$. If q is a common extension of p and r, then $q \leq p$ and $q \Vdash \phi(\pi, \tau_1, \ldots, \tau_n)$. $\Box_{9.4.7}$

9.5 The generic extension is a model of ZFC

Now, we use our results thus far to show that the generic extension $M[G]$ is a model of ZFC.

We will use a form of Choice that is different from those presented thus far.

Lemma 9.5.1. *The Axiom of Choice holds iff*

$$\forall x\, \exists \alpha \in \boldsymbol{ON} \exists f\, (f \text{ is a function } \wedge \mathrm{dom}(f) = \alpha \wedge x \subset \mathrm{rng}(f)).$$

Proof. If x is a set, α is an ordinal, and we have a function f that is a mapping as described above, then we can define a well-ordering of x in the following manner. Let $g(z) = \min(f^{-1}\{z\})$ for $z \in x$. Then g maps x 1-1 into α. Then, let $yRz \iff g(y) < g(z)$. The relation R well-orders x. $\Box_{9.5.1}$

Theorem 9.5.2. *Let M be a countable transitive model for ZFC, $\langle \mathbb{P}, \leq, \mathbb{1}\rangle$ a partial order in M, and G $\mathbb{P}$-generic over M. Then $M[G]$ satisfies ZFC.*

Proof. **Extensionality, Foundation, Pairing, Union** were verified in Lemma 9.3.14.

We check **Comprehension**. We must check that whenever $\sigma, \tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$ and $\phi(x, v, y_1, \ldots, y_n)$ is any formula,

$$\{a \in \sigma_G : (\phi(a, \sigma_G, \tau_{1G}, \ldots, \tau_{nG}))^{M[G]}\} \in M[G].$$

Let

$$\rho = \{\langle \pi, p\rangle \in \mathrm{dom}(\sigma) \times \mathbb{P} : p \Vdash (\pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \ldots, \tau_n))\}.$$

By the definability of forcing as given by Statement 1 of Theorem 9.4.6, $\rho \in M^{\mathbb{P}}$. We check that $\rho_G = \{a \in \sigma_G : \phi(a)^{M[G]}\}$: (out of laziness, we leave out mention of $\tau_1, \ldots, \tau_n$.) First, any element of ρ_G is of the form π_G, where $\langle \pi, p\rangle \in \rho$ for some $p \in G$. By definition of ρ, $p \Vdash (\pi \in \sigma \wedge \phi(\pi))$. Thus, by the definition of $\Vdash$, $\pi_G \in \sigma_G$ and $\phi(\pi_G)^{M[G]}$. Thus, $\rho_G \subseteq \{a \in \sigma_G : \phi(a)^{M[G]}\}$. To establish equality, assume $a \in \sigma_G$ and $\phi(a)^{M[G]}$. Then, $a = \pi_G$ for some $\pi \in \mathrm{dom}(\sigma)$. Thus $(\pi_G \in \sigma_G \wedge \phi(\pi_G))^{M[G]}$. Since, by statement 2 of Theorem 9.4.6, any statement true in $M[G]$ is forced, there is a $p \in G$ such that $p \Vdash (\pi \in \sigma \wedge \phi(\pi))$. Then, $\langle \pi, p\rangle \in \rho$, so $\pi_G \in \rho_G$.

Next, we check **Replacement**. That is, we check that for each formula $\phi(x, v, r, z_1, \ldots, z_n)$ and each $\sigma_G, \tau_{1G}, \ldots, \tau_{nG} \in M[G]$, if

$$(\forall x \in \sigma_G\, \exists! y\, (\phi(x, y, \sigma_G, \tau_{1G}, \ldots, \tau_{nG}))^{M[G]},$$

then there is a $\rho \in M^{\mathbb{P}}$ such that

$$\forall x \in \sigma_G\, \exists y \in \rho_G\, (\phi(x, y, \sigma_G, \tau_{1G}, \ldots, \tau_{nG}))^{M[G]}.$$

(Again, out of laziness, we leave out mention of $\tau_1, \ldots, \tau_n$.) Let $S \in M$ be such that $S \subset M^{\mathbb{P}}$ and

$$\forall \pi \in \mathrm{dom}(\sigma)\, \forall p \in \mathbb{P}\, (\exists \mu \in M^{\mathbb{P}}\, (p \Vdash \phi(\pi, \mu)) \to \exists \mu \in S\, (p \Vdash \phi(\pi, \mu))).$$

The set S exists because, by Statement 1 of Theorem 9.4.6, $p \Vdash \phi(\pi, \mu)$ is defined by a formula relativized to M, so by the Reflection Theorem 6.4.3 in M we can take $S = R(\alpha)^M \cap M^{\mathbb{P}}$ for a suitable α. Let $\rho = S \times \mathbb{1}$. Then $\rho_G = \{\mu_G : \mu \in S\}$. Fix $x \in \sigma_G$. We show that $\exists y \in \rho_G\, (\phi(x, y))^{M[G]}$: Note that $x = \pi_G$ for some $\pi \in \mathrm{dom}(\sigma)$. By assumption, $(\exists y\, \phi(\pi_G, y))^{M[G]}$. So, for some $v \in M^{\mathbb{P}}$, $\phi(\pi_G, v_G)^{M[G]}$. By Statement 2 of Theorem 9.4.6, there is a $p \in G$ such that $p \Vdash \phi(\pi, v)$. Then, there is a $\mu \in S$ such that $p \Vdash \phi(\pi, \mu)$. Thus we have $\mu_G \in \rho_G$ and $(\phi(\pi_G, \mu_G))^{M[G]}$.

It may seem that we have proved a stronger form of Replacement – one which weakens the $\exists! y$ in the hypothesis to $\exists y$. However, this "stronger" axiom is actually a version of reflection and is derivable in ZF.

Infinity holds since ω $(= (\check{\omega}_G)$ is in $M[G]$. Thus, $M[G]$ satisfies ZF − P.

——————— HERE ENDED SPRING 2008 LECTURE 8 (90 min) ————

For the **Power Set Axiom**, fix $\sigma_G \in M[G]$. We aim to find a name $\rho \in M^{\mathbb{P}}$ such that $\forall x \in M[G]\, (x \subset \sigma_G \to x \in \rho_G)$. To this end, let $\rho = S \times \{\mathbb{1}\}$, where

$$S = \{\tau \in M^{\mathbb{P}} : \mathrm{dom}(\tau) \subset \mathrm{dom}(\sigma)\} = (\mathscr{P}(\mathrm{dom}(\sigma) \times \mathbb{P}))^M.$$

Fix any $\mu \in M^{\mathbb{P}}$ such that $\mu_G \subset \sigma_G$. We show that $\mu_G \in \rho_G$. Let

$$\tau = \{\langle \pi, p\rangle : \pi \in \mathrm{dom}(\sigma) \wedge p \Vdash \pi \in \mu\}.$$

Then $\tau \in S$, so $\tau_G \in \rho_G$. The argument will be complete once we show that $\mu_G = \tau_G$. To see that $\mu_G \subseteq \tau_G$, note that since $\mu_G \subseteq \sigma_G$, any element of μ_G is of the form π_G for some $\pi \in \text{dom}(\sigma)$. Since $\pi_G \in \mu_G$, there is a $p \in G$ such that $p \Vdash \pi \in \mu$. Thus $\langle \pi, p \rangle \in \tau$, so $\pi_G \in \tau_G$. For the other inclusion $\tau_G \subseteq \mu_G$, note that any element of τ_G is of the form π_G, where $\langle \pi, p \rangle \in \tau$ for some $p \in G$. Then $p \Vdash \pi \in \mu$, so $\pi_G \in \mu_G$.

The key to the proof of the Power Set Axiom in $M[G]$ lies in the fact that in M there is a set of names which contains representatives for any possible subset of σ_G, even though the collection of all μ such that $\mu_G \subseteq \sigma_G$ (or even $\sigma_G = \emptyset$) is usually not contained in a set of M.

Finally, we check the **Axiom of Choice**. Fix $x = \sigma_G \in M[G]$. By the assumption that the Axiom of Choice holds in M, let $\text{dom}(\sigma) = \{\pi_\gamma : \gamma < \alpha\}$, where the function which takes γ to π_γ in in M. Let

$$\tau = \{\text{op}(\check{\gamma}, \pi_\gamma) : \gamma < \alpha\} \times \{\mathbb{1}\}.$$

Then $\tau \in M$ and $\tau_G = \{\langle \gamma, \pi_{\gamma G} \rangle : \gamma < \alpha\}$. So, τ_G is a function with $\text{dom}(\tau_G) = \alpha$ and $\sigma_G \subseteq \text{rng}(\tau_G)$. $\square_{9.5.2}$

———————— HERE ENDED SPRING 2007 LECTURE 8 (135 min) ————-

Corollary 9.5.3. *Let M be a countable transitive model for ZFC. Then there is a countable transitive model $N \supset M$ such that N satisfies ZFC $+ \mathbf{V} \neq \mathbf{L}$.*

Proof. We use the notation of the previous theorem. Choose $\mathbb{P}$ such that $G \notin M$. This will be true whenever $\mathbb{P}$ satisfies the condition of Lemma 9.2.5 (i.e. when $\forall p \in \mathbb{P}\, \exists q, r \in \mathbb{P}\, (q \leq p \wedge r \leq p \wedge q \perp r)$). An example is $\mathbb{P}$ consisting of finite partial functions from ω to $\{0, 1\}$. Then, let $N = M[G]$. Since, by Lemma 9.3.11, $\text{o}(N) = \text{o}(M)$, and $\mathbf{L}^N = \mathbf{L}^M \subset M$, thus N satisfies $\mathbf{V} \neq \mathbf{L}$. $\square_{9.5.3}$

As we continue our development of forcing, we will often discuss the relation $p \Vdash \phi$ where ϕ will be a statement of some mathematical complexity. We will not write ϕ explicitly as a formula in the first-order language of set theory. Rather we will standard mathematical notation which we will consider to be an abbreviation for the first-order formula. We do not have to worry about the exact way we write the unabbreviated formula, since two formulas which are equivalent in ZFC are forced by the same conditions. More precisely, the following holds:

Lemma 9.5.4.

1. *Let $\phi(x_1, \ldots, x_n)$ and $\psi(x_1, \ldots, x_n)$ be formulas, and assume*

$$ZFC \vdash \forall x_1, \ldots, x_n\, (\phi(x_1, \ldots, x_n) \rightarrow \psi(x_1, \ldots, x_n)).$$

Then, for any countable transitive model M for ZFC, partial order $\mathbb{P} \in M$, $p \in \mathbb{P}$, and $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$,

$$(p \Vdash \phi(\tau_1, \ldots, \tau_n)) \rightarrow (p \Vdash \psi(\tau_1, \ldots, \tau_n)).$$

2. *If we assume also that*

$$ZFC \vdash \forall x_1, \ldots, x_n\, (\phi(x_1, \ldots, x_n) \iff \psi(x_1, \ldots, x_n)),$$

then

$$(p \Vdash \phi(\tau_1, \ldots, \tau_n)) \iff (p \Vdash \psi(\tau_1, \ldots, \tau_n)).$$

Proof. For statement 1, for any G which is $\mathbb{P}$-generic over M, $M[G]$ satisfies ZFC, so

$$\phi(\tau_{1G},\ldots,\tau_{nG})^{M[G]} \to \psi(\tau_{1G},\ldots,\tau_{nG})^{M[G]}.$$

Thus, statement 1 follows from the definition of $\Vdash$.

Statement 2 follows from statement 1. $\square_{9.5.4}$

In the following section(s), we discuss how to find a partial order $\mathbb{P}$ so that $M[G]$ will satisfy particular further set theoretic axioms.

9.6 The consistency of the failure of the Continuum Hypothesis with ZFC

In this section, we will give the most famous relative consistency proof produced by forcing: that of $\mathrm{Con}(ZFC{+}\neg CH)$. The method of this section entails forcing over a countable transitive model M with finite partial functions from one set I to another set J. This will allow us to construct models in which $2^{\aleph_0}$ is $\aleph_2$, $\aleph_5$, $\aleph_{\omega_1}$, or anything else not obviously contradictory.

Definition 9.6.1. We define the set of *finite partial functions* from a set I to another set J as

$$\mathrm{PrtFn}^{<\omega}(I,J) = \{p : |p| < \omega \wedge p \text{ is a function } \wedge \mathrm{dom}(p) \subset I \wedge \mathrm{rng}(p) \subset J\}.$$

The set $\mathrm{PrtFn}^{<\omega}(I,J)$ is ordered by $p \le q \iff p \supset q$.

The set $\mathrm{PrtFn}^{<\omega}(I,J)$ is a partial order with largest element $\mathbb{1} = \emptyset$. Since the notion of finiteness is absolute, so in the definition of $\mathrm{PrtFn}^{<\omega}(I,J)$. Thus, if $I, J \in M$, then $\mathrm{PrtFn}^{<\omega}(I,J) = \mathrm{PrtFn}^{<\omega}(I,J)^M \in M$.

Earlier, we mentioned $\mathrm{PrtFn}^{<\omega}(\omega,2)$, and here the idea is similar. We can now generalize an example given earlier for the specific partial order $\mathrm{PrtFn}^{<\omega}(\omega,2)$.

Lemma 9.6.2. *If $I, J \in M$, I is infinite, $J \neq \emptyset$, and G is a $\mathrm{PrtFn}^{<\omega}(I,J)$-generic filter over M, then $\bigcup G$ is a function from I onto J.*

Proof. If G is a filter in $\mathrm{PrtFn}^{<\omega}(I,J)$, then $\bigcup G$ is a function with $\mathrm{dom}(\bigcup G) \subset I$ and $\mathrm{rng}(\bigcup G) \subset J$. If $J \neq \emptyset$, then $D_i = \{p \in \mathrm{PrtFn}^{<\omega}(I,J) : i \in \mathrm{dom}(p)\}$ is dense for all $i \in I$. By absoluteness, $D_i \in M$ if $I, J \in M$. Thus, if G is generic over M, $G \cap D_i \neq \emptyset$ for each $i \in I$. Thus $\mathrm{dom}(\bigcup G) = I$. Likewise, if I is infinite, $\{p \in \mathrm{PrtFn}^{<\omega}(I,J) : j \in \mathrm{rng}(p)\}$ is dense and in M, so $\mathrm{rng}(\bigcup G) = J$. $\square_{9.6.2}$

One thing that can be shown using this kind of partial order is that the notion of a cardinal need not be absolute for M and $M[G]$. Thus, let κ be an uncountable cardinal of M – that is, $\kappa \in M$ and (κ is an uncountable cardinal)^M. Let $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\omega,\kappa)$, and let G be $\mathbb{P}$-generic over M. Then $\bigcup G \in M[G]$ (by absoluteness of $\bigcup$), and G is a function from ω onto κ. So, in $M[G]$, κ is a countable ordinal. In such a case, we say that $\mathbb{P}$ *collapses* κ.

Using a different pair of sets I and J, we can use $\mathrm{PrtFn}^{<\omega}(I,J)$ to get a model in which the Continuum Hypothesis is false.

Lemma 9.6.3. *Let κ be an uncountable cardinal of M, and let $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\kappa\times\omega,2)$. If G is $\mathrm{PrtFn}^{<\omega}(\kappa\times\omega,2)$-generic over M, then $(2^\omega \geqslant |\kappa|)^{M[G]}$.*

Proof. If G is $\mathbb{P}$-generic over M, for $\mathbb{P}$ as above, then $\bigcup G : \kappa \times \omega \longrightarrow 2$. We can think of G as coding a κ sequence of functions from ω into 2. Namely, let $f_\alpha(n) = (\bigcup G)(\alpha, n)$ for $\alpha < \kappa$, $n < \omega$. By absoluteness, the sequence $\langle f_\alpha : \alpha < \kappa \rangle$ (i.e. the function that assigns f_α to each α) is in $M[G]$. Furthermore, the f_α are all distinct: If $\alpha \neq \beta$, let

$$D_{\alpha\beta} = \{p \in \mathbb{P} : \exists n \in \omega\, (\langle \alpha, n\rangle \in \mathrm{dom}(p) \wedge \langle \beta, n\rangle \in \mathrm{dom}(p) \wedge p(\alpha, n) \neq p(\beta, n))\}.$$

Then $D_{\alpha\beta}$ is dense and in M, so $G \cap D_{\alpha\beta} \neq \emptyset$, which implies $f_\alpha \neq f_\beta$. Thus, $M[G]$ contains a κ-sequence of distinct functions from ω into $\{0, 1\}$. $\square_{9.6.3}$

Taking $\kappa = (\aleph_2)^M$, the above lemma would seem to imply that $2^{\aleph_0} \geqslant \aleph_2$ in $M[G]$, i.e. that CH fails in $M[G]$. But we cannot immediately jump to this conclusion so. We sill have to check that $\kappa = (\aleph_2)^{M[G]}$ also holds. This is not immediately obvious since, as we saw in Lemma 6.4.8, there are partial orders that force an uncountable cardinal in M to become a countable ordinal in $M[G]$.

9.6.1 The countable (anti-)chain condition

That $\mathrm{PrtFn}^{<\omega}(\kappa \times \omega, 2)$ does not collapse κ to a countable ordinal involves the fact that this partial order has certain combinatorial properties in M. In particular, we have in mind the following (very badly named) property:

Definition 9.6.4. A partial order $\langle \mathbb{P}, \leq \rangle$ has the *countable chain condition* (abbreviated *c.c.c.*) iff every antichain in $\mathbb{P}$ is countable.

I emphasize that we are interested in the combinatorial properties of $\mathbb{P}$ *in* M. The partial order $\mathrm{PrtFn}^{<\omega}(\kappa \times \omega, 2)$ has c.c.c. in $\mathbf{V}$ since M is countable, but this is irrelevant! That $(\mathrm{PrtFn}^{<\omega}(\kappa \times \omega, 2)$ has c.c.c.$)^M$ follows from the following more general result, relativized to M:

Lemma 9.6.5. *If I is arbitrary and J is countable, then* $\mathrm{PrtFn}^{<\omega}(I, J)$ *has c.c.c.*

We need a definition and a lemma to prove Lemma 9.6.5.

Definition 9.6.6. A family $\mathscr{A}$ of sets is called a Δ-*system*, or a *quasi-disjoint* family iff there is a fixed set r, called the *root* of the quasi-disjoint family, such that $a \cap b = r$ when ever a and b are distinct members of $\mathscr{A}$. Draw a picture of a witch's broom, where the handle is the root.

Lemma 9.6.7 (Δ-system Lemma)**.** *Let κ be any infinite cardinal. Let $\lambda > \kappa$ be regular and satisfy*

$$\forall \alpha < \lambda\, (|\alpha^{<\kappa}| < \lambda).$$

Assume $|\mathscr{A}| \geqslant \lambda$ and $\forall x \in \mathscr{A}\, (|x| < \kappa)$. Then, there is a family $\mathscr{B} \subseteq \mathscr{A}$ such that $|\mathscr{B}| = \lambda$ and $\mathscr{B}$ is a quasi-disjoint family.

Proof. We can assume $|\mathscr{A}| = \lambda$, shrinking $\mathscr{A}$ if necessary. Then, $|\bigcup \mathscr{A}| \leqslant \lambda$. Since it is irrelevant what the elements of $\mathscr{A}$ exactly are, we can assume that $\bigcup \mathscr{A} \subset \lambda$. Then, each $x \in \mathscr{A}$ has some order type $< \kappa$ as a subset of λ. Since λ is regular and $\lambda > \kappa$, there is some $\varrho < \kappa$ such that $\mathscr{A}_1 = \{x \in \mathscr{A} : x \text{ has type } \varrho\}$ has cardinality λ. here we are applying the pigeonhole principle. We fix such a ϱ, and look only at $\mathscr{A}_1$.

For each $\alpha < \lambda$, $|\alpha^{<\kappa}| < \lambda$ implies that less than λ elements of $\mathscr{A}_1$ are subsets of α. Thus, $\bigcup \mathscr{A}_1$ is unbounded in λ. If $x \in \mathscr{A}_1$ and $\xi < \varrho$, let $x(\xi)$ be the ξ-th element of x. Since λ is regular, there is some ξ such that $\{x(\xi) : x \in \mathscr{A}_1\}$ is unbounded in λ. Fix ξ_0 to be the least such ξ (note that ξ_0 might be 0!). Let

$$\alpha_0 = \sup\{x(\eta) + 1 : x \in \mathscr{A}_1 \wedge \eta < \xi_0\}.$$

Then $\alpha_0 < \lambda$ and $x(\eta) < \alpha_0$ for all $x \in \mathscr{A}_1$ and all $\eta < \xi_0$.

By transifinte recursion on $\mu < \lambda$, we pick $x_\mu \in \mathscr{A}$ so that $x_\mu(\xi_0) > \alpha_0$ and $x_\mu(\xi_0)$ is above all elements of earlier x_ν. That is,

$$x_\mu(\xi_0) > \max(\alpha_0, \sup\{x_\nu(\eta) : \eta < \varrho \wedge \nu < \mu\}).$$

Let $\mathscr{A}_2 = \{x_\mu : \mu < \lambda\}$. Then $|\mathscr{A}_2| = \lambda$ and $x \cap y \subset \alpha_0$ whenever x and y are distinct elements of $\mathscr{A}_2$. Since $|\alpha_0^{<\kappa}| < \lambda$, there is an $r \subset \alpha_0$ and a $\mathscr{B} \subset \mathscr{A}_2$ with $|\mathscr{B}| = \lambda$ and $\forall x \in \mathscr{B}\,(x \cap \alpha_0 = r)$. Thus, $\mathscr{B}$ forms a Δ-system with root r. $\square_{9.6.7}$

Proof. Let $p_\alpha \in \mathrm{PrtFn}^{<\omega}(I, J)$ for $\alpha < \omega_1$ and let $a_\alpha = \mathrm{dom}(p_\alpha)$. By the Δ-system Lemma for $\kappa = \omega$ and $\lambda = \omega_1$, there is an uncountable $X \subset \omega_1$ such that $\{a_\alpha : \alpha \in X\}$ forms a Δ-system with some root r. Since J is countable, ${}^rJ$ is as well, so there are only countably many possibilities for $p_\alpha \restriction r$. Thus, there is an uncountable $Y \subset X$ such that the $p_\alpha \restriction r$ for $\alpha \in Y$ are all the same. But then the p_α for $\alpha \in Y$ are all compatible. Thus, there can never be a family $\{p_\alpha : \alpha < \omega\}$ of incompatible conditions. $\square_{9.6.5}$

There are lots more examples of c.c.c. partial orders. The importance of c.c.c. in forcing is the following lemma. This lemma gives a way of approximating, within M, any function which appears in $M[G]$.

Lemma 9.6.8. *Assume $\mathbb{P} \in M$, ($\mathbb{P}$ is c.c.c.)^M, and $A, B \in M$. Let G be $\mathbb{P}$-generic over M, and let $f \in M[G]$, with $f : A \longrightarrow B$. Then there is a map $F : A \longrightarrow \mathscr{P}(B)$ with $F \in M$, $\forall a \in A\,(f(a) \in F(a))$ and $\forall a \in A\,((|F(a)| \leqslant \omega)^M)$.*

Proof. Fix $\tau \in M^{\mathbb{P}}$ with $f = \tau_G$. Since any statement true in $M[G]$ is forced, there is a $p \in G$ such that

$$p \Vdash \tau \text{ is a function from } \check{A} \text{ into } \check{B}.$$

Formally, we are applying Statement 2 of Theorem 9.4.6 here to a formula $\phi(x, y, z)$ which asserts that x is a function from y into z. By Lemma 9.5.4, exactly which formula ϕ we use does not matter.

Define

$$F(a) = \{b \in B : \exists q \leq p\,(q \Vdash \tau(\check{a}) = \check{b})\}.$$

By the definability of $\Vdash$, $F \in M$.

Fix $a \in A$. To see that $f(a) \in F(a)$, let $b = f(a)$. Then, there is an $r \in G$ such that $r \Vdash \tau(\check{a}) = \check{b}$, and r and p have a common extension, q. Then, $q \Vdash \tau(\check{a}) = \check{b}$. Thus $b \in F(a)$.

To see that $(|F(a)| \leqslant \omega)^M$, we apply the Axiom of Choice in M to find a function $Q \in M$ such that $Q : F(a) \longrightarrow \mathbb{P}$ and, for $b \in F(a)$, $Q(b) \leq p$ and $Q(b) \Vdash \tau(\check{a}) = \check{b}$. If $b, b' \in F(a)$, and $b \neq b'$, then $Q(b) \perp Q(b')$, since they force inconsistent statements. To be more precise, if $Q(b)$ and $Q(b')$ were compatible, there would be a generic H containing both of them. Then, in $M[H]$, $\tau_H : A \longrightarrow B$, $\tau_H(a) = b$, and $\tau_H(a) = b'$. Thus, $\{Q(b) : b \in F(a)\}$ is an antichain in $\mathbb{P}$. So, since $Q \in M$ and $(\mathbb{P}$ is c.c.c.$)^M$, $(|F(a)| \leqslant \omega)^M$. $\square_{9.6.8}$

The countable (anti-)chain condition has relevance to the absoluteness of cardinals.

Definition 9.6.9. If $\mathbb{P} \in M$, $\mathbb{P}$ *preserves cardinals* iff whenever G is a $\mathbb{P}$-generic over M,

$$\forall \beta \in \mathrm{o}(M)\, ((\beta \text{ is a cardinal})^M \iff (\beta \text{ is a cardinal})^{M[G]}).$$

Note that since ω is absolute, preservation of cardinals is only a problem for $\beta > \omega$. Also, if β is a cardinal of $M[G]$, it is immediately a cardinal of M since any function in M from a smaller ordinal onto β would be in $M[G]$ also. Thus, $\mathbb{P}$ preserves cardinals iff

$$\forall \beta \in \mathrm{o}(M)\, ((\beta > \omega \wedge (\beta \text{ is a cardinal})^M) \rightarrow (\beta \text{ is a cardinal})^{M[G]}).$$

—————— HERE ENDED SPRING 2007 LECTURE 9 (135 min) ———-

One can now easily see, using Lemma 9.6.8, that if $(\mathbb{P}$ is c.c.c.$)^M$, then $\mathbb{P}$ preserves cardinals – just take $B = \beta$ and A an ordinal $< \beta$. In fact $\mathbb{P}$ preserves cofinalities as well, which is a slightly stronger assertion.

Definition 9.6.10. If $\mathbb{P} \in M$, $\mathbb{P}$ *preserves cofinalities* iff whenever G is $\mathbb{P}$-generic over M and γ is a limit ordinal in M,

$$\mathrm{cf}(\gamma)^M = \mathrm{cf}(\gamma)^{M[G]}.$$

Lemma 9.6.11. *If $\mathbb{P}$ preserves cofinalities, then $\mathbb{P}$ preserves cardinals.*

Proof. Assume $\mathbb{P}$ preserves cofinalities. If $\alpha \geqslant \omega$ is a regular cardinal of M, then $\mathrm{cf}(\alpha)^{M[G]} = \mathrm{cf}(\alpha)^M = \alpha$, so α is a regular cardinal of $M[G]$. If $\beta > \omega$ is a limit cardinal of M, then the regular (in fact, successor) cardinals of M are unbounded in β. Since these remain regular in $M[G]$, β is a limit cardinal in $M[G]$ as well. Since every infinite cardinal is either regular or a limit cardinal (or both), every infinite cardinal of M is a cardinal of $M[G]$. $\square_{9.6.11}$

There are examples of forcings $\mathbb{P}$ which preserve cardinals without preserving cofinalities – so-called Prikry forcing.

The following lemma gives a simpler condition that needs to be checked for preservation of cofinalities.

Lemma 9.6.12. *Assume $\mathbb{P} \in M$ and whenever G is $\mathbb{P}$-generic over M and κ is a regular uncountable cardinal of M, $(\kappa$ is regular$)^{M[G]}$. Then $\mathbb{P}$ preserves cofinalities.*

Proof. Let γ be a limit ordinal in M, and let $(\kappa = \mathrm{cf}(\gamma))^M$. Then there is a function $f \in M$ such that f maps κ into γ cofinally and f is strictly increasing (here we are using Lemma 4.5.2 within M). Since $(\kappa$ is regular$)^M$, $(\kappa$ is regular$)^{M[G]}$ (here we are applying absoluteness of ω in the case that $\kappa = \omega$). Since $f \in M[G]$, $(\kappa = \mathrm{cf}(\gamma))^{M[G]}$ (here we are applying Lemma 4.5.3 within $M[G]$). $\square_{9.6.12}$

Theorem 9.6.13. *If $\mathbb{P} \in M$ and $(\mathbb{P}$ has c.c.c.$)^M$, then $\mathbb{P}$ preserves cofinalities, and hence cardinals.*

Proof. Assume to the contrary. Then, by the previous Lemma 9.6.12, there is a $\kappa \in M$ with $\kappa > \omega$, $(\kappa \text{ is regular})^M$, and $(\kappa \text{ is not regular})^{M[G]}$. Thus, there is an $\alpha < \kappa$ and a function $f \in M[G]$ such that f maps α cofinally into κ. By Lemma 9.6.8, let F be in M, with $F : \alpha \longrightarrow \mathscr{P}(\kappa)$, $\forall \xi < \alpha\,(f(\xi) \in F(\xi))$, and $\forall \xi < \alpha\,(|F(\xi)| \leqslant \omega)^M$. Let $S = \bigcup_{\xi<\alpha} F(\xi)$. Then $S \in M$ and S is an unbounded subset of κ. Applying in M the fact that the union of $|\alpha|$ countable sets has cardinality $|\alpha|$, $(|S| = |\alpha| < \kappa)^M$, contradicting that $(\kappa \text{ is regular})^M$. $\square_{9.6.13}$

Now we finally have everything we need to produce a model of the negation of the Continuum Hypothesis. The forcing that we are about to describe is often called *Cohen forcing*.

Let $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\omega_2^M \times \omega, 2)$. Then this $\mathbb{P}$ has c.c.c. in M and thus preserves cardinals. Thus $\omega_2^M = \omega_2^{M[G]}$. Lemma 9.6.3 shows that $(2^\omega \geqslant \omega_2)^{M[G]}$.

This leads us to the next question: can 2^ω be exactly ω_2? If we were to start with a model M in which (G)CH does not hold and, say, $(2^\omega \geqslant \omega_3)^M$, then the same holds in any cardinal-preserving extension of M. However, if the ground model M is a model for GCH, forcing with $\mathrm{PrtFn}^{<\omega}(\omega_2^M \times \omega, 2)$ makes 2^ω exactly ω_2 in $M[G]$. Generally, we will use the values of cardinal exponents in M to put an upper bound on cardinal exponents in $M[G]$.

We will get these upper bounds by going through the proof of the Power Set Axiom in $M[G]$ more carefully. To just show the Power Set Axiom, it was enough, given $\sigma \in M^{\mathbb{P}}$, to find in M some set S of names which represented all possible subsets of σ. Now, we will try to find such an S *of small cardinality.*

Definition 9.6.14. If $\sigma \in \mathbf{V}^{\mathbb{P}}$, a *nice name* for a subset of σ is $\tau \in \mathbf{V}^{\mathbb{P}}$ of the form $\bigcup\{\{\pi\} \times A_\pi : \pi \in \mathrm{dom}(\sigma)\}$, where each A_π is an antichain in $\mathbb{P}$.

We will be using the notion of a nice name in M, but the property of being a nice name is absolute.

Lemma 9.6.15. *If $\mathbb{P} \in M$ and $\sigma, \mu \in M^{\mathbb{P}}$, then there is a nice name $\tau \in M^{\mathbb{P}}$ for a subset of σ such that*

$$\mathbb{1} \Vdash (\mu \subset \sigma \rightarrow \mu = \tau).$$

Proof. For each $\pi \in \mathrm{dom}(\sigma)$, let $A_\pi \subset \mathbb{P}$ be such that:

1. $\forall p \in A_\pi\,(p \Vdash \pi \in \mu)$,
2. A_π is an antichain in $\mathbb{P}$, and
3. A_π is maximal with respect to conditions 1 and 2.

We can assume $\langle A_\pi : \pi \in \mathrm{dom}(\sigma)\rangle \in M$ by definability of $\Vdash$ and Zorn's Lemma applied within M. Let

$$\tau = \bigcup\{\{\pi\} \times A_\pi : \pi \in \mathrm{dom}(\sigma)\}.$$

To show that $\mathbb{1} \Vdash (\mu \subset \sigma \rightarrow \mu = \tau)$, we show that whenever G is $\mathbb{P}$-generic over M, $\mu_G \subset \sigma_G \rightarrow \mu_G = \tau_G$. Assume $\mu_G \subset \sigma_G$.

First, we show that $\mu_G \subset \tau_G$: Fix $a \in \mu_G$. Since $\mu_G \subset \sigma_G$, $a = \pi_G$ for some $\pi \in \mathrm{dom}(\sigma)$. If $A_\pi \cap G \neq \emptyset$, fix $p \in A_\pi \cap G$. Then $\langle \pi, p \rangle \in \tau$, and $p \in G$, so $a = \pi_G \in \tau_G$. But, if $A_\pi \cap G = \emptyset$, let $q \in G$ be such that $\forall p \in A\,(p \perp q)$ (take a look at Lemma 9.3.16). Let $q' \in G$ be such that $q' \Vdash \pi \in \mu$, and let r be a common extension of q and q'. Then $A_\pi \cup \{r\}$ satisfies conditions 1 and 2 above, contradicting the maximality of A_π.

To show that $\tau_G \subset \mu_G$, fix $a \in \tau_G$. Then $a = \pi_G$, where $\langle \pi, p \rangle \in \tau$ for some $p \in G$. By definition of τ, $p \Vdash \pi \in \mu$, so $a = \pi_G \in \mu_G$. $\square_{9.6.15}$

If τ is a nice name for a subset of σ, it need not in general be true that $\tau_G \subset \sigma_G$, but that does not matte. The important thing is that every subset of σ does get represented by a nice name.

Lemma 9.6.16. *Assume hat $\mathbb{P} \in M$ and that in M, $\mathbb{P}$ is c.c.c., $|\mathbb{P}| = \kappa \geqslant \omega$, λ is an infinite cardinal, and $\theta = \kappa^\lambda$. Let G be $\mathbb{P}$-generic over M. Then in $M[G]$, $2^\lambda \leqslant \theta$.*

Proof. In M, every antichain in $\mathbb{P}$ is countable, so there are at most κ^ω such antichains. Since $\mathrm{dom}(\check{\lambda}) = \{\xi : \xi < \lambda\}$ has cardinality λ there are at most $(\kappa^\omega)^\lambda = \kappa^\lambda = \theta$ nice names for subsets of $\check{\lambda}$. Let τ_α, $(\alpha < \theta)$, enumerate, in M, all nice names for subsets of $\check{\lambda}$.

In $M[G]$, there is a function f with domain θ such that $f(\alpha) = \tau_{\alpha G}$ for each $\alpha < \theta$. Namely, $f = \pi_G$, where $\pi = \{\langle \mathrm{op}(\check{\alpha}, \tau_\alpha), \mathbb{1} \rangle : \alpha < \theta\}$. But by Lemma 9.6.15, $\mathscr{P}(\lambda)^{M[G]} \subset \mathrm{rng}(f)$, so $(2^\lambda \leqslant \theta)^{M[G]}$. $\square_{9.6.16}$

We can apply the previous lemma to show that the size of the continuum can be almost anything.

Lemma 9.6.17. *Let κ be an infinite cardinal of M such that $(\kappa^\omega = \kappa)^M$, and let $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\kappa \times \omega, 2)$. Let G be $\mathbb{P}$-generic over M. Then $(2^\omega = \kappa)^{M[G]}$.*

Proof. Applying Lemma 9.6.16 with $\lambda = \omega$ gives us $2^\omega \leqslant \kappa$ in M. However, by Lemma 9.6.3, $2^\omega \geqslant \kappa$ in M. Since $\mathbb{P}$ has c.c.c. in M, κ is still a cardinal in $M[G]$. $\square_{9.6.17}$

So, in particular, if M satisfies GCH, then in M $\kappa^\omega = \kappa$ whenever $\mathrm{cf}(\kappa) > \omega$ (Lemma 4.5.14). It follows that it is consistent for the continuum to be anything not cofinal with ω (since by König's Lemma 4.5.12, $\mathrm{cf}(2^\omega) > \omega$). Thus, we have the following:

Corollary 9.6.18.

1. $\mathrm{Con}(ZFC) \to \mathrm{Con}(ZFC + 2^\omega = \omega_2)$,
2. $\mathrm{Con}(ZFC) \to \mathrm{Con}(ZFC + 2^\omega = \omega_{\omega_1})$, *etc.*

Proof. We have already discussed why the method of generic extensions gives relative consistency results.

We can start with M satisfying $ZFC + GCH$ since in ZFC we can prove the existence of a countable transitive model for any finite number of axioms of $ZFC + (\mathbf{V} = \mathbf{L})$, and $\mathbf{V} = \mathbf{L}$ implies GCH.

To get statement 2, we start with M satisfying GCH and apply Lemma 9.6.17 with $(\kappa = \omega_{\omega_1})^M$. Then, in $M[G]$, $2^\omega = \kappa$. Since $\mathbb{P}$ preserves cardinals, $\kappa = \omega_{\omega_1}$ in $M[G]$. $\square_{9.6.18}$

The continuum can also be weakly inaccessible.

Corollary 9.6.19. *The following four theories are equiconsistent. That is*

$$\mathrm{Con}(T_1) \iff \mathrm{Con}(T_2) \iff \mathrm{Con}(T_3) \iff \mathrm{Con}(T_4),$$

where

T_1	*is*	*ZFC* $+$ *GCH* $+ \exists\kappa\,(\kappa$ *is strongly inaccessible*$)$.
T_2	*is*	*ZFC* $+ \exists\kappa\,(\kappa$ *is weakly inaccessible*$)$.
T_3	*is*	*ZFC* $+ 2^\omega$ *is weakly inaccessible.*
T_4	*is*	*ZFC* $+ \exists\kappa < 2^\omega\,(\kappa$ *is weakly inaccessible*$)$.

Proof. That $\mathrm{Con}(\mathbf{T_3}) \to \mathrm{Con}(\mathbf{T_2})$ and $\mathrm{Con}(\mathbf{T_4}) \to \mathrm{Con}(\mathbf{T_2})$ is clear.

$\mathrm{Con}(\mathbf{T_2}) \to \mathrm{Con}(\mathbf{T_1})$: Notice that as a theorem of ZFC, if κ is weakly inaccessible, then κ is weakly inaccessible in $\mathbf{L}$ and hence, by GCH in $\mathbf{L}$, strongly inaccessible in $\mathbf{L}$. This, within T_2 we can prove that $\mathbf{L}$ is an inner model for T_1.

$\mathrm{Con}(\mathbf{T_1}) \to \mathrm{Con}(\mathbf{T_3})$ and $\mathrm{Con}(\mathbf{T_1}) \to \mathrm{Con}(\mathbf{T_4})$: Let M be a countable transitive model for T_1. If $\mathbb{P}$ is c.c.c. in M and κ is weakly inaccessible in M, then, by preservation of cofinalities, κ will be both regular and a limit cardinal in $M[G]$., and hence κ will remain weakly inaccessible in $M[G]$. Thus, if $\lambda > \kappa$ and λ is a cardinal in M, then forcing with $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\lambda \times \omega, 2)$ makes $M[G]$ a model for T_4. If κ is strongly inaccessible in M, then $(\kappa^\omega = \kappa)^M$, so forcing with $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\kappa \times \omega, 2)$ makes $(2^\omega = \kappa)^{M[G]}$, and so $M[G]$ satisfies T_3.

Formally, to see that the above considerations yield a finitistic relative consistency proof of $\mathrm{Con}(T_1) \to \mathrm{Con}(T_3)$ or of $\mathrm{Con}(T_1) \to \mathrm{Con}(T_4)$, we can apply our discussion about finitistic relative consistency proofs that use the forcing method with T_1 as the basic theory instead of ZFC. Thus, in T_1 we can prove the existence of a countable transitive model M for any desired finite list of axioms of T_1, and then by forcing produce a countable transitive model $M[G]$ for any finite list of axioms of T_3 or T_4. $\square_{9.6.19}$

The Gödel Incompleteness Theorem implies that we cannot expect to produce relative consistency proofs of the form $\mathrm{Con}(ZFC) \to \mathrm{Con}(T_1)$.

It is also possible to calculate powers of uncountable cardinals in extensions by $\mathrm{PrtFn}^{<\omega}(\kappa \times \omega, 2)$. The particular case where $\kappa = 1$ is an oft quoted relative consistency result.

Corollary 9.6.20.

$$\mathrm{Con}(ZFC) \to \mathrm{Con}(ZFC + GCH + \mathbf{V} \neq \mathbf{L}).$$

Proof. We start with a ground model M satisfying GCH. Let $\mathbb{P} = \mathrm{PrtFn}^{<\omega}(\omega, 2)$. The proof of Corollary 9.5.3 points out that $M[G]$ satisfies $\mathbf{V} \neq \mathbf{L}$. If λ is an infinite cardinal of M, let $\theta = (\lambda^+)^M = (\omega^\lambda)^M$. By Lemma 9.6.16, $(2^\lambda \leqslant \theta)^{M[G]}$. Thus, $\forall\lambda \geqslant \omega\,(2^\lambda \leqslant \lambda^+)^{M[G]}$, so GCH holds in $M[G]$. $\square_{9.6.20}$

9.7 Models of $CH + \neg GCH$

We now look at partial orders that will allow us to build models that do not satisfy GCH, but CH holds.

Definition 9.7.1. For any infinite cardinal λ,

$$\mathrm{PrtFn}^{<\lambda}(I,J) = \{p : |p| < \lambda \wedge p \text{ is a function } \wedge \mathrm{dom}(p) \subset I \wedge \mathrm{rng}(p) \subset J\}.$$

We order $\mathrm{PrtFn}^{<\lambda}$ as usual: $p \le q \iff q \subset p$. Clearly, this is a partial order with largest element $\mathbb{1} = \emptyset$.

When $\lambda > \omega$, $\mathrm{PrtFn}^{<\lambda}(I,J)$ is NOT absolute for M! In our forcing considerations, we will always use $\mathrm{PrtFn}^{<\lambda}(I,J)^M$, where ($\lambda$ is a cardinal)^M. Useful and interesting results are only obtained if λ is a regular cardinal in M, but we will not need this latter restriction in some of the below.

Lemma 9.7.2. *If $I, J, \lambda \in M$, (λ is a cardinal)^M, $J \neq \emptyset$, $(|I| \geqslant \lambda)^M$, and G is* $\mathrm{PrtFn}^{<\lambda}(I,J)$*-generic over M, then $\bigcup G$ is a function from I to J.*

The proof of the above is analogous to that of Lemma 9.6.2. Similarly, the proof of the following is analogous to that of Lemma 9.6.3.

Lemma 9.7.3. *If (λ is a cardinal)^M, $\kappa \in M$, and G is* $\mathrm{PrtFn}^{<\lambda}(\kappa \times \lambda, 2)^M$*-generic over M, then $(2^{|\lambda|} \geqslant |\kappa|)^{M[G]}$.*

As in the previous section, the hard part here is showing that cardinals are preserved. We have to do a bit of work, since if $\lambda > \omega$, then $\mathrm{PrtFn}^{<\lambda}(I,J)$ has c.c.c. only in the trivial cases when $|I| < \omega$ or $|J| \leqslant 1$.

The work here will be split into two parts. First, we modify the c.c.c. argument to check that cardinals $> \lambda$ are preserved. Second, we introduce a new idea to check that cardinals $\leqslant \lambda$ in M remain cardinals in $M[G]$, which was trivial when $\lambda = \omega$. For all of this to work, we will eventually need that λ is regular and $2^{<\lambda} = \lambda$ in M.

As before, we check that cardinals are preserved by checking that cofinalities are preserved.

Definition 9.7.4. Assume that $\mathbb{P} \in M$ and θ is an infinite cardinal of M.

1. We say that $\mathbb{P}$ *preserves cardinals* $\geqslant \theta$ (or $\leqslant \theta$) iff whenever G is $\mathbb{P}$-generic over M, $\beta \in \mathrm{o}(M)$, and $\beta \geqslant \theta$ (respectively, $\beta \leqslant \theta$),
$$(\beta \text{ is a cardinal})^M \iff (\beta \text{ is a cardinal})^{M[G]}.$$

2. We say that $\mathbb{P}$ *preserves cofinalities* $\geqslant \theta$ (or $\leqslant \theta$) iff whenever G is $\mathbb{P}$-generic over M, γ is a limit cardinal in M, and $\mathrm{cf}(\gamma)^M \geqslant \theta$ (respectively,$\mathrm{cf}(\gamma)^M \leqslant \theta$), then
$$\mathrm{cf}(\gamma)^M = \mathrm{cf}(\gamma)^{M[G]}.$$

Lemma 9.7.5. *Under the assumptions of the previous Definition 9.7.4, if $\mathbb{P}$ preserves cofinalities $\leqslant \theta$, then $\mathbb{P}$ preserves cardinals $\leqslant \theta$. If $\mathbb{P}$ preserves cofinalities $\geqslant \theta$, and (θ is regular)^M, then $\mathbb{P}$ preserves cardinals $\geqslant \theta$.*

Lemma 9.7.6. *With the assumptions of Definition 9.7.4, assume further that whenever κ is a regular cardinal of M, $\kappa \geqslant \theta$, and G is $\mathbb{P}$-generic over M, and (κ is regular)$^{M[G]}$. Then $\mathbb{P}$ preserves cofinalities $\geqslant \theta$. Likewise for $\leqslant$ instead of $\geqslant$.*

The proofs of the above are as in the analogous lemmas from the previous section.

If we weaking "countable" in the definition of c.c.c. to "$< \theta$", then we preserve cofinalities $\geqslant \theta$.

Definition 9.7.7. A partial order $\mathbb{P}$ has the *θ-chain condition* (abbreviated *θ-c.c.*) iff every antichain in $\mathbb{P}$ has cardinality $< \theta$.

This is a more sensibly named property. Using this naming convention, c.c.c. is ω_1-c.c.

Again, analogously to the previous section, we have the following.

Lemma 9.7.8. *Assume $\mathbb{P} \in M$, $A, B \in M$, $(\theta$ is a cardinal$)^M$, and $(\mathbb{P}$ is θ-c.c.$)^M$. Let G be $\mathbb{P}$-generic over M, and let $f \in M[G]$, with $f : A \longrightarrow B$. Then, there is a map $F : A \longrightarrow \mathscr{P}(B)$ with $F \in M$, $\forall a \in A\,(f(a) \in F(a))$, and $\forall a \in A\,(|F(a)| < \theta)^M$.*

Lemma 9.7.9. *Assume $\mathbb{P} \in M$, θ is a cardinal of M, and $(\mathbb{P}$ is θ-c.c.$)^M$. Then $\mathbb{P}$ preserves cofinalities $\geqslant \theta$. Hence, if it is also true that $(\theta$ is regular$)^M$, then $\mathbb{P}$ preserves cardinals $\geqslant \theta$.*

So, which chain conditions occur in practice? Let $cc(\mathbb{P})$ be the smallest θ such that $\mathbb{P}$ has θ-c.c. A theorem of Tarski states that c.c.$(\mathbb{P})$ is finite or regular. Thus, we can remove the assumption that $(\theta$ is regular$)^M$ from the previous lemma. Furthermore, $cc(\mathbb{P})$ cannot be ω. For each $n < \omega$, there do exist $\mathbb{P}$ which are n-c.c. (namely $\mathrm{PrtFn}^{<\omega}(1, n-1)$). However, if $cc(\mathbb{P}) < \omega$, then the resulting forcing is not interesting, since any G which is $\mathbb{P}$-generic over M will be in M. If θ is weakly inaccessible, then there is an important example of a partial order with $cc(\mathbb{P}) = \theta$ – the so called *Levy collapsing order*, which is used to force a weakly inaccessible cardinal to be $\aleph_1$. sadly, I won't get to the proof of this during this lecture.

Finally, assume the last case, that $\theta = \lambda^+$. A trivial example of a partial order $\mathbb{P}$ with $cc(\mathbb{P}) = \theta$ is $\mathrm{PrtFn}^{<\omega}(1, \lambda)$. A more interesting partial order to examine is $\mathbb{P} = \mathrm{PrtFn}^{<\lambda}(I, 2)$. Under the assumption of GCH, $cc(\mathbb{P}) = \lambda^+$ if $|I| \geqslant \lambda$. If GCH does not hold, then $cc(\mathbb{P}) = (2^{<\lambda})^+$. That $cc(\mathbb{P}) \geqslant (2^{<\lambda})^+$, I leave as an exercise. But we need the other inequality for our considerations.

———————— HERE ENDED SPRING 2007 LECTURE 10 (135 min) ———-

Lemma 9.7.10. *The partial order $\mathrm{PrtFn}^{<\lambda}(I, J)$ has $(|J|^{<\lambda})^+$-c.c.*

Proof. Let $\theta = (|J|^{<\lambda})^+$, and suppose that $\{p_\xi : \xi < \theta\}$ forms an anti-chain. We look at two cases for λ. First, assume λ is regular. Then $(|J|^{<\lambda})^{<\lambda} = |J|^{<\lambda}$, so $\forall \alpha < \theta\,(|\alpha^{<\lambda}| < \theta)$, so by the Δ-system Lemma, there is a set $X \subset \theta$ with $|X| = \theta$ such that $\{\mathrm{dom}(p_\xi) : \xi \in X\}$ forms a Δ-system with some root r. Since there are less than θ possibilities for $p_\xi \restriction r$, we have a contradiction just as in the analogous proof for $\lambda = \omega$.

Secondly, assume that λ is singular. Then, since θ is regular and $> \lambda$, we can find a regular $\lambda' < \lambda$ such that $Y = \{\xi : |p_\xi| < \lambda'\}$ has cardinality θ. Then $\{p_\xi : \xi \in Y\}$ contradicts the $(|J|^{<\lambda'})^+$-c.c. that we proved for regular λ'. $\square_{9.7.10}$

Corollary 9.7.11. *Assume $I, J \in M$. Assume further that, in M, λ is regular, $|J| \leqslant 2^{<\lambda}$, and $\theta = (2^{<\lambda})^+$. Then $\mathrm{PrtFn}^{<\lambda}(I, J)^M$ preserves cofinalities and cardinals $\geqslant \theta$.*

Proof. Lemma 9.7.10, applied within M, implies that $\mathrm{PrtFn}^{<\lambda}(I, J)$ has θ-c.c. in M since $(|J|^{<\lambda} = 2^{<\lambda})^M$. The rest follows from Lemma 9.7.9. □

Now we will use a completely different argument to show that if λ is regular in M, then $\mathrm{PrtFn}^{<\lambda}(I, \kappa)^M$ preserves cofinalities and cardinals $\leqslant \lambda$. Under GCH, $2^{<\lambda} = \lambda$, so Corollary 9.7.11 implies that all cofinalities and cardinals will be preserved. However, if in M there are cardinals κ such that $\lambda^+ \leqslant \kappa \leqslant 2^{<\lambda}$, then except in trivial cases such κ will have cardinality λ, and so will no longer be cardinals, in $M[G]$.

Definition 9.7.12. A partial order $\mathbb{P}$ is λ-*closed* iff whenever $\gamma < \lambda$ and $\{p_\xi : \xi < \gamma\}$ is a decreasing sequence of elements of $\mathbb{P}$ (i.e. $\xi < \eta \to p_\xi \geq p_\eta$), then

$$\exists q \in \mathbb{P}\, \forall \xi < \gamma\, (q \leq p_\xi).$$

Lemma 9.7.13. *If λ is regular, then $\mathrm{PrtFn}^{<\lambda}(I, J)$ is λ-closed.*

Proof. Let $q = \bigcup\{p_\xi : \xi < \gamma\}$. Then $|q| < \lambda$ since each $|p_\xi| < \lambda$ and λ is regular. $\square_{9.7.13}$

Note that if λ is singular, then $\mathrm{PrtFn}^{<\lambda}(\lambda, 2)$ is not λ-closed. Also, if $(\lambda$ is singular$)^M$, then $\mathrm{PrtFn}^{<\lambda}(\lambda, 2)^M$ collapses λ.

On the other hand, if λ is regular, then the fact that $\mathrm{PrtFn}^{<\lambda}(I, J)$ is λ-closed will be used to show that cardinals $\leqslant \lambda$ are preserved.

The proof of the next theorem should be compared to the proof of Lemma 9.7.8. That lemma used a chain condition to approximate, in M, functions from A to B in $M[G]$. The next theorem shows that functions from A to B are in fact in M if A is small enough.

Theorem 9.7.14. *Assume $\mathbb{P} \in M$, $A, B \in M$. Assume further that $(\lambda$ is a cardinal$)^M$, $(\mathbb{P}$ is λ-closed$)^M$, and $(|A| < \lambda)^M$. Let G be $\mathbb{P}$-generic over M and let $f \in M[G]$ with $f : A \longrightarrow B$. Then $f \in M$.*

Proof. Note first that it is enough to prove the statement of the theorem with A being an ordinal, $A = \alpha < \gamma$. To prove the general result, we can then let $j \in M$ be a 1-1 map from $\alpha = |A|^M < \lambda$ onto A, and apply the special case with $f \circ j : \alpha \longrightarrow B$ to show that $f \circ j$, and hence f, is in M.

Let $K = ({}^\alpha B)^M = {}^\alpha B \cap M$, and $f \in {}^\alpha B \cap M[G]$. We want to show that $f \in K$. Assume otherwise. Then we can fix $\tau \in M^{\mathbb{P}}$ such that $f = \tau_G$, and then fix $p \in G$ such that

$$p \Vdash (\tau \text{ is a function from } \check{\alpha} \text{ into } \check{B}\ \wedge \tau \notin \check{K}).$$

We will now argue based on the above forcing statement.

Working in M, we use transfinite recursion together with the Axiom of Choice to choose sequences $\{p_\eta : \eta \leqslant \alpha\}$ from $\mathbb{P}$ and $\{z_\eta : \eta < \alpha\}$ from B so that

1. $p_0 = p$,
2. $p_\eta \leq p_\xi$ for all $\xi \leqslant \eta$, and
3. $p_{\eta+1} \Vdash \tau(\check{\eta} = \check{z}_\eta$.

For successor steps in this recursion, we are given p_η, and we find $p_{\eta+1}$ and z_η in the following manner: $p_\eta \le p$, so

$$p_\eta \Vdash (\tau \text{ is a function from } \check{\alpha} \text{ into } \check{B}).$$

Since a consequence of a forced statement is forced,

$$p_\eta \Vdash \exists x \in \check{B}\, (\tau(\check{\eta} = x).$$

Thus, by statement 4 of Corollary 9.4.7, there is a $z_\eta \in B$ and $p_{\eta+1} \le p_\eta$ such that $p_{\eta+1} \Vdash \tau(\check{\eta} = \check{z_\eta}$.

At the limit steps, let $g = \langle z_\eta : \eta < \alpha \rangle$. SO, g is the function with domain α such tath $g(\eta) = z_\eta$ for each η. Then $g \in K$.

Let H be $\mathbb{P}$-generic over M,with $p_\alpha \in H$, and so each $p_\eta \in H$. Then $\tau_H(\eta) = z_\eta$ for each $\eta < \alpha$, to $\tau_H = g \in K$. But $p_0 = p \Vdash \tau \notin \check{K}$, so $\tau_H \notin K$, which is a contradiction. $\square_{9.7.14}$

Corollary 9.7.15. *Assume $\mathbb{P} \in M$, (λ is a cardinal)^M, and ($\mathbb{P}$ is λ-closed)^M. Then, $\mathbb{P}$ preserves cofinalities $\leqslant \lambda$, and hence cardinals $\leqslant \lambda$.*

Proof. Assume to the contrary. Then, by Lemma 9.7.6, there is a $\kappa \leqslant \lambda$ such that κ is a regular cardinal in M, but κ is a singular cardinal in $M[G]$. Thus, there is $\alpha < \kappa$ and $f \in M[G]$ which maps α cofinally into κ. By the previous Theorem 9.7.14, $f \in M$, which contradicts the regularity of κ in M. $\square_{9.7.15}$

Theorem 9.7.16. *Let $\lambda, I, J \in M$. Assume that in M, λ is regular, $2^{<\lambda} = \lambda$, and $|J| \leqslant \lambda$. Then $\mathrm{PrtFn}^{<\lambda}(I,J)^M$ preserves cofinalities, and hence cardinals.*

Proof. By the regularity of λ, $\mathrm{PrtFn}^{<\lambda}(I,J)^M$ is λ-closed in M, and so preserves cofinalities $\leqslant \lambda$. By the assumption that $2^{<\lambda} = \lambda$, $\mathrm{PrtFn}^{<\lambda}(I,J)^M$ has λ^+-c.c. in M, and so preserves cofinalities $\geqslant (\lambda^+)^M$. $\square_{9.7.16}$

So, we can force with orders of the form $\mathrm{PrtFn}^{<\lambda}(\kappa \times \lambda, 2)^M$ to violate GCH as badly as we wish at λ. We can use nice names to get a precise computation of 2^λ in $M[G]$. So, analogous to before, we have:

Theorem 9.7.17. *In M, assume that $\lambda < \kappa$, λ is regular, $2^{<\lambda} = \lambda$, and $\kappa^\lambda = \kappa$. Let $\mathbb{P} = \mathrm{PrtFn}^{<\lambda}(\kappa \times \lambda, 2)^M$. Then $\mathbb{P}$ preserves cardinals, and if G is $\mathbb{P}$-generic over M, then $(2^\lambda = \kappa)^{M[G]}$.*

Proof. Since we have shown preservation of cardinals, and Lemma 9.7.3 makes showing that $(2^\lambda \geqslant \kappa)^{M[G]}$ easy, all we have left to show is that $(2^\lambda \leqslant \kappa)^{M[G]}$.

In M, $\mathbb{P}$ has cardinality $\kappa^{<\lambda} = \kappa$. The partial order $\mathbb{P}$ also has λ^+-c.c., so there are at most $\kappa^\lambda = \kappa$ many antichains in $\mathbb{P}$. Thus, there are at most $\kappa^\lambda = \kappa$ nice names for subsets of λ. Let $\langle \tau_\alpha : \alpha < \kappa \rangle$ be an enumeration of the nice names, and let

$$\pi = \{\langle \mathrm{op}(\check{\alpha}, \tau_\alpha), \mathbb{1} \rangle : \alpha < \kappa\}.$$

Then, as in the analogous proof from the previous section, in $M[G]$, π_G is a function, $\mathrm{dom}(\pi_G) = \kappa$, and $\mathscr{P}(\lambda) \subset \mathrm{rng}(\pi_G)$, so $2^\lambda \leqslant \kappa$. $\square_{9.7.17}$

One can use the method of Theorem 9.7.17 to compute the powers of all cardinals in $M[G]$ (not only those of λ) in terms of the cardinal arithmetic in M. We can also use this method to violate GCH as we wish at any regular cardinal, or even at any finite number of regular cardinals. For example:

Theorem 9.7.18. *If ZFC is consistent, then the following are as well:*

1. $ZFC + CH + (2^{\omega_1} = \omega_2) + (2^{\omega_2} = \omega_{\omega_8})$.
2. $ZFC + CH + (2^{\omega_1} = \omega_5) + (2^{\omega_2} = \omega_7)$.
3. $ZFC + (2^{\omega} = \omega_3) + (2^{\omega_1} = \omega_4) + (2^{\omega_2} = \omega_6)$.

Proof. Assume our ground model M satisfies $ZFC + GCH$.

1. Let $\mathbb{P} = \mathrm{PrtFn}^{<\omega_2}(\omega_{\omega_8} \times \omega_2, 2)^M$. By Theorem 9.7.17, $\mathbb{P}$ preserves cardinals, and if G is $\mathbb{P}$-generic over M, $(2^{\omega_2} = \omega_{\omega_8})^{M[G]}$. That $2^{\omega_1} = \omega_2$ holds in $M[G]$ follows from the fact that $({}^{\omega_1}2)^M = ({}^{\omega_1}2)^{M[G]}$ by Theorem 9.7.14. So, if $F \in M$, and F maps ω_2 onto ${}^{\omega_1}2$ in M, then F maps ω_2 onto ${}^{\omega_1}2$ in $M[G]$ as well. Similarly, $(2^{\omega_1} = \omega_2)^{M[G]}$.

2. We will force twice. Let $\mathbb{P}_1 = \mathrm{PrtFn}^{<\omega_2}(\omega_7 \times \omega_2, 2)^M$, G be$\mathbb{P}_1$-generic over M, nd let $N = M[G]$. Then, by the arguments of 1,
$$((2^{\omega} = \omega_1) \wedge (2^{\omega_1} = \omega_2) \wedge (2^{\omega_2} = \omega_7))^N.$$
Furthermore, $\kappa^{\omega_1} = \kappa)^N$ whenever $((\kappa \geqslant \omega_2 <) \wedge (\kappa \text{ is regular}))^N$, since this is true in M by the assumption that GCH holds in M, and $({}^{\omega_1}\kappa)^M = ({}^{\omega_1}\kappa)^{M[G]}$.
Now, we treat N as a ground model, and force again. Let
$$\mathbb{P}_2 = \mathrm{PrtFn}^{<\omega_1}(\omega_5 \times \omega_1, 2)^N.$$
Since $(2^{<\omega_1} = \omega_1)^N$, $\mathbb{P}_2$ preserves cardinals. Let H be $\mathbb{P}_2$-generic over N. Using the same arguments as in 1, CH holds in $N[H]$. That $(2^{\omega_2} \geqslant \omega_7)$ holds in $N[H]$ follows from the fact that this inequality holds in N. To show equality, that is, that $(2^{\omega_2} = \omega_7)$ holds in $N[H]$, we use the method of Theorem 9.7.17. In particular, in N, $\mathbb{P}_2$ has ω_2-c.c. and $|\mathbb{P}_2| = \omega_5^{\omega_1} = \omega_3$, so there are only $((\omega_5)^{\omega_1})^{\omega_2} = \omega_7$ many nice names for subsets of ω_2. To see that $(2^{\omega_1} = \omega_5)^{N[H]}$, we apply Theorem 9.7.17 directly, using the fact that $(\omega_5^{\omega_1} = \omega_5)^N$.

3. Here we force three times, so that the first forcing extension satisfies
$$(2^{\omega} = \omega_1) \wedge (2^{\omega_1} = \omega_2) \wedge (2^{\omega_2} = \omega_6),$$
the second extension satisfies
$$(2^{\omega} = \omega_1) \wedge (2^{\omega_1} = \omega_4) \wedge (2^{\omega_2} = \omega_6),$$
and the third and final extension satisfies
$$(2^{\omega_1} = \omega_4) + (2^{\omega_2} = \omega_6).$$

$\square_{9.7.18}$

Note that in 2 and 3, it was important that we dealt with the largest cardinal first. For example, if in 2 we were to start with the smaller cardinal, by first forcing with a partial order $\mathbb{Q}_1 = \mathrm{PrtFn}^{<\omega_1}(\omega_5 \times \omega_1)^M$, then $M[G]$ would satisfy $2^{\omega_1} = \omega_5$. Then, $(2^{<\omega_2} \neq \omega_2)^{M[G]}$, so if we let $\mathbb{Q}_2 = \mathrm{PrtFn}^{<\omega_2}(\omega_7 \times \omega_2, 2)^{M[G]}$, then $\mathbb{Q}_2$ would not preserve cardinals. In fact, if we were to force with $\mathbb{Q}_2$ anyway, with H begin $\mathbb{Q}_2$-generic over $M[G]$, then $(\omega_5)^{M[G]}$ would have cardinality ω_2 in $M[G][H]$, $(2^{\omega_1} = \omega_2)^{M[G][H]}$.

Index

[1]