Set Theory: The Basics

Notes for the Second Indian Winter School on Logic
Indian Institute of Technology, Kanpur

A. C. Walczak-Typke
Kurt Gödel Research Center for Mathematical Logic
University of Vienna
Vienna, Austria

January 14-26, 2008

Contents

About These Notes

Origins and Sources

The pages before you are a much abbreviated digest of the first 4 chapters of the scriptum that I use for my Axiomatic Set Theory lectures at the University of Vienna. The notes for those lectures are heavily based on the lectures of Prof. Bogdan Węglorz, given at the University of Wrocław, that I attended as an undergraduate. His lectures were in turn based upon Kenneth Kunen's *Set Theory* and Thomas Jech's *Set Theory*. Thus, the lion's share of these notes has as its ultimate source those two books. Bits here and there were also taken from Karol Hrbaček and Jech's book *Introduction to Set Theory*, Wilfrid Hodge's *Model Theory*, Keith Devlin's *Aspects of Constructibility*, Roman Murawski's *Filozofia Matematyki: Zarys Dziejów*.

Format of Lectures and Notes

Sections with asterisks will probably not be presented, but they are good to know, so you should read them on your own!

Chapter 1

Motivation, Language, and Axioms

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.

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 How to Speak: 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 $\neg((\neg(\phi)) \wedge (\neg(\psi)))$;
- $(\phi) \to (\psi)$ abbreviates $(\neg(\phi)) \vee (\psi)$;
- $(\phi) \leftrightarrow (\psi)$ stands for $((\phi) \to (\psi)) \wedge ((\psi) \to (\phi))$
- $v_j \neq v_i$ 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 $\forall x\, (x \in a \to \phi)$
- Similarly, $\exists x \in a\, \phi$ stands for $\exists x\, (x \in a \wedge \phi)$
- $\exists! x \phi$ is an abbreviation of $\exists x\, (\phi(x) \wedge (\forall y)(\phi(y) \to y = x))$. 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

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

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.*

1.3 The Axioms of Set Theory

1.3.1 Statement and discussion of the axioms of ZFC

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.

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 1.3.1. $\emptyset$ 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 1.3.2.

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

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. $\square_{1.3.2}$

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

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

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. This is an easy EXERCISE.

We can now define:

Definition 1.3.3. A *(Kuratowski) ordered pair* is defined to be

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

Clearly, $\langle x, y\rangle = \langle x', y'\rangle \rightarrow x = x' \wedge y = y'$. The "clearly" is an easy EXERCISE.

Axiom 4 (Union)**.**

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

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

Definition 1.3.4. 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 1.3.5. 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:

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

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 1.3.6. 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.

Definitions 1.3.7. 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 1.3.8. *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 1.3.9. f is called a *function* iff f is a relation and

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

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

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, $f"C = \mathrm{rng}(f \restriction C) = \{f(x) : x \in C\}$. 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 $\mathrm{rng}(f) = B$. A function that is both a surjection and an injection is called a *bijection*.

We can use functions to compare relations.

Definition 1.3.10. 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)**.**

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

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

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

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.

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.

1.3.2 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 usually note when a theorem can be proved within one of these partial systems.

Chapter 2

Orders and Ordinals

2.1 Orders

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

Definition 2.1.1. 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 2.1.2. 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 2.1.3. 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 2.1.4. *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. $\square_{2.1.4}$

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

Lemma 2.1.5. *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 $f(z) S g(z)$ or $g(z) S f(z)$. Let us assume that $f(z) S g(z)$. 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, $g(t) S g(z)$, which gives tRz because g is an isomorphism. This means that t is R-smaller than z and $t \in X$. Contradiction. $\square_{2.1.5}$

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

Theorem 2.1.6. *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. The details here are left as an EXERCISE. $\square_{2.1.6}$

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 2.1.7. *The following statements are equivalent:*

AC Axiom of choice

WOP Well-ordering Principle

We postpone the proof of Theorem 2.1.7 until the next section.

2.2 Ordinals

We begin with some definitions:

Definition 2.2.1. 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 2.2.2. 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 $\mathrm{pred}(\alpha, \beta)$ instead of $\mathrm{pred}(\alpha, \beta, \in_\alpha)$.

Theorem 2.2.3.

1. *If α is an ordinal and $y \in \alpha$, then y is also an ordinal and $y = \mathrm{pred}(\alpha, y)$;*

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 2.1.5, 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 exist (there will be a loop). I leave the details as an EXERCISE.

(3): To prove this, use (1), (2), and Theorem 2.1.6. 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 $y \cap C = \emptyset$. $\square_{2.2.3}$

Theorem 2.2.3 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 2.2.4 (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 2.2.3 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 2.2.3, no ordinal is a member of itself. $\square_{2.2.4}$

Lemma 2.2.5. *If A is a transitive set of ordinals, then A itself is an ordinal.*

The proof of the above lemma is clear from the definitions and is left as an EXERCISE.

The following gives us a main point of ordinals.

Theorem 2.2.6. *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 2.2.3 (2).

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

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

Let $C = \operatorname{rng}(f)$. By the Replacement Axiom, C is a set. Using Lemma 2.2.5, 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 $B = \operatorname{pred}(A, b, R)$. In the latter case, this would mean that $b \in B$, which is not possible. $\square_{2.2.6}$

Remark 2.2.7.

1. *The proof of Theorem 2.2.6 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 $\langle \operatorname{pred}(A, a, R), R\rangle \cong x$. 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 2.2.6.*

3. *Theorem 2.2.6 allows us to use ordinals as representatives of well-order types.*

Definition 2.2.8. If $\langle A, R\rangle$ is a well-ordering, then $\operatorname{type}(\langle A, R\rangle)$ is the unique ordinal α such that $\langle A, R\rangle \cong \alpha$.

Definition 2.2.9. 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 $\exists x\,(x$ 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 2.2.10.

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

The proof of the above lemma is left as an EXERCISE.

2.3 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 2.1.7 (AC $\Leftrightarrow$ WOP) (which we have not yet proved) means that we can well-order every set. Theorem 2.2.6 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 2.3.1. We define the *successor* of an ordinal:

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

A simple lemma, the proof of which is left as an EXERCISE:

Lemma 2.3.2. *For any ordinal α,*

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

Definition 2.3.3. 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 2.3.4. $0 = \emptyset$, $1 = S(0)$, $2 = S(1)$, $3 = S(2)$, $4 = S(3)$,... etc.

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

Definition 2.3.5. An ordinal α is a *natural number* iff $\forall\beta \leqslant \alpha\,(\beta = 0 \vee$ β 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: $\omega = \{z \in x : z$ is a natural number.$\}$

Definition 2.3.6. ω is the set of natural numbers.

The set ω is an ordinal by Lemma 2.2.5. 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 2.3.7. *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.

$\square_{2.3.7}$

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.

2.4 Ordinal Addition and Multiplication

Now we define some basic arithmetic operations on the ordinals.

Definition 2.4.1. $\alpha + \beta = \text{type}(\langle \alpha \times \{0\} \cup \beta \times \{1\}, R\rangle)$, 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 2.4.2. *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 $\alpha \times \{0\} \cup \beta \times \{1\} \cup \gamma \times \{2\}$. $\square_{2.4.2}$

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

Definition 2.4.3. 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 (the proof is left as an EXERCISE).

Lemma 2.4.4. *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 2.4.5. (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 2.4.5 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 2.4.6. 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 2.4.7. 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 $\zeta < \beta$.

2.5 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 2.5.1.

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

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

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 2.5.2 (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 2.2.3(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 2.2.3(5), let β be the smallest element of $\alpha \cap \mathbf{C}$. Clearly, β is then the smallest element of $\mathbf{C}$. $\square_{2.5.2}$

Mathematically, Theorems 2.2.3(5) and 2.5.2. are very similar. Formally, there is an enormous difference between them. Theorem 2.2.3(5) is the abbreviation of one provable sentence. On the other hand, Theorem 2.5.2 is a theorem schema which represents infinitely many theorems.

It is possible, of course, to state Theorem 2.5.2 without classes. To do this, we would have to say: for each formula $\mathbf{C}(x, z_1, \dots, z_n)$, the following is a theorem:

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

Note that here we think of $\mathbf{C}$ as defining $\{x : \mathbf{C}(x, z_1, \dots, z_n)\}$, with $z_1, \dots, z_n$ 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 2.5.2) 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 2.2.3(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.

Theorem 2.5.3 (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*

$$\forall \alpha \, (\mathbf{G}(\alpha) = \mathbf{F}(\mathbf{G} \restriction \alpha)). \tag{2.1}$$

Proof. To show **uniqueness**, assume that there are functions $\mathbf{G_1}$ and $\mathbf{G_2}$ that both satisfy 2.1. Then, it is possible to prove that $\forall \alpha \, (\mathbf{G_1}(\alpha) = \mathbf{G_2}(\alpha))$ by transfinite induction on α. We leave the details of this 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) $\delta > \alpha$. $\square_{2.5.3}$

To state Theorem 2.5.3 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 2.5.3), 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.

2.6 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.

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

Definition 2.6.1. 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 2.6.2. *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 $\alpha \cdot \beta \leqslant \gamma$. The details are left as an EXERCISE. $\square_{2.6.2}$

2.7 Proof of AC $\Leftrightarrow$ WOP

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

Proof. **AC$\Rightarrow$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$\Rightarrow$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. $\square_{2.1.7}$

Chapter 3

Cardinal Numbers

3.1 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 3.1.1.

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

Theorem 3.1.2 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 $y \in \mathrm{rng}(h)$. 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 $h(g(y)) = g^{-1}(g(y)) = y$. $\square_{3.1.2}$

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 3.1.3. 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 3.1.4. α is a *cardinal* if $\alpha = |\alpha|$.

Lemma 3.1.5. *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 3.1.2, we get the result. $\square_{3.1.5}$

Lemma 3.1.6. *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 3.1.5. $\square_{3.1.6}$

Corollary 3.1.7. *ω is a cardinal, and each $n \in \omega$ is a cardinal.*

Definition 3.1.8. 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 need the Powerset Axiom for an uncountable set to exist.

3.2 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 3.2.1.

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

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

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 3.2.2. *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 3.1.6 (2). $\square_{3.2.2}$

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

Lemma 3.2.3. *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. $\square_{3.2.3}$

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 3.2.4. *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 3.2.2.

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

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

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 3.1.) So, $\text{type}(\kappa \times \kappa, \ll) \leqslant \kappa$, so $|\kappa \times \kappa| \leqslant \kappa$. Since clearly $|\kappa \times \kappa| \geqslant \kappa$, we have equality. $\square_{3.2.4}$

Corollary 3.2.5. *Let κ and λ be infinite cardinals. Then,*

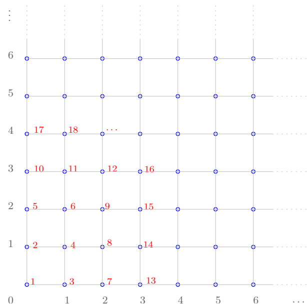

Figure 3.1: *A tiny initial portion of the well-ordering $\ll$. This shows that, in the worst case, the predecessors of a given pair are contained in the square defined by that pair.*

1. $\kappa \oplus \lambda = \kappa \otimes \lambda = \max(\kappa, \lambda)$;
2. $|\kappa^{<\omega}| \leqslant \omega \otimes \kappa = \kappa$. *(This was defined in Definition 2.4.5.)*

Proof. We prove only **(2)**: We use the proof of Theorem 3.2.4 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 $|\kappa^{<\omega}| \leqslant \omega \otimes \kappa = \kappa$. $\square_{3.2.5}$

3.3 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 3.3.1. 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 3.3.2 ((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. $\square_{3.3.2}$

With the help of the Axiom of Choice, one can deduce from Theorem 3.3.2 that there exists a cardinal $> \omega$, in particular, $|\mathscr{P}(\omega)|$.

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

Theorem 3.3.3 ((ZF$^-$) Hartogs, 1906)**.**

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

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 $\sup(S)$. $\square_{3.3.3}$

Definition 3.3.4 ((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 3.3.5. $\aleph_\alpha = \omega_\alpha$ 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 3.3.6.

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. $\square_{3.3.6}$

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

Lemma 3.3.7 ((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 $Y \preccurlyeq X$. $\square_{3.3.7}$

Note: As in the Cantor's Theorem 3.3.2, one can prove, even without the Axiom of Choice, that there exists a mapping from $\mathscr{P}(\omega)$ onto ω_1, but one cannot prove the existence of a 1-1 function from ω_1 into $\mathscr{P}(\omega)$.

Lemma 3.3.8 ((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. $\square_{3.3.8}$

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.

3.4 Cardinal Exponentiation

Definition 3.4.1 (ZF$^-$)**.** $A^B = {}^B A = \{f : f \text{ is a function } \wedge \operatorname{dom}(f) = B \wedge \operatorname{rng}(f) \subseteq A\}$.

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

Definition 3.4.2 (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 3.4.3. *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 ${}^\lambda 2 \preccurlyeq {}^\lambda\kappa \approx {}^\lambda\lambda \preccurlyeq \mathscr{P}(\lambda \times \lambda) \approx \mathscr{P}(\lambda) \approx 2^\lambda$. $\Box_{3.4.3}$

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 3.4.4 (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 ${}^C({}^BA) \approx {}^{C \times B}A$. $\Box_{3.4.4}$

Definition 3.4.5 (AC).

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

Cantor showed that $2^{\aleph_\alpha} \geqslant \aleph_{\alpha+1}$ (Theorem 3.3.2), 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. But! Cohen showed in 1963 that if ZFC is consistent, then so is ZFC + ¬CH. So, the continuum hypothesis is independent of ZFC.

3.5 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 3.5.1. If $f : \alpha \longrightarrow \beta$, f maps α *cofinally* iff $\operatorname{rng}(f)$ is unbounded in β. The *cofinality* of β, written $\operatorname{cf}(\beta)$, is the least α such that there is a map from α cofinally into β.

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

Lemma 3.5.2. *There is a cofinal map $f : \operatorname{cf}(\beta) \longrightarrow \beta$ which is strictly increasing (i.e. $\zeta < \nu \rightarrow f(\zeta) < f(\nu)$).*

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

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

$\Box_{3.5.2}$

Lemma 3.5.3. *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. $\square_{3.5.3}$

Corollary 3.5.4. $\mathrm{cf}(\mathrm{cf}(\beta)) = \mathrm{cf}(\beta)$.

Proof. We use Lemma 3.5.3 on a strictly increasing function $f : \mathrm{cf}(\beta) \longrightarrow \beta$, whose existence is guaranteed by Lemma 3.5.2. $\square_{3.5.4}$

Definition 3.5.5. An ordinal β is *regular* iff β is a limit ordinal and $\mathrm{cf}(\beta) = \beta$.

So, by Corollary 3.5.4, $\mathrm{cf}(\beta)$ is regular for every limit β.

Lemma 3.5.6. *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. $\square_{3.5.6}$

Definition 3.5.7. An infinite cardinal κ is *regular* if $\mathrm{cf}(\kappa) = \kappa$. It is *singular* if $\mathrm{cf}(\kappa) < \kappa$.

Lemma 3.5.8. *ω and all infinite* $\mathrm{cf}(\beta)$ *are regular.*

Lemma 3.5.9 (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 3.3.8 also of cardinality $\leqslant \kappa$ (and in particular $\neq \kappa^+$). Contradiction. $\square_{3.5.9}$

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

Proof. This results from Lemma 3.5.3. $\square_{3.5.10}$

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 3.5.11.

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 3.5.12 (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. $\Box_{3.5.12}$

Corollary 3.5.13 (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 3.5.12 with $\kappa = 2^\lambda$. $\Box_{3.5.13}$

Lemma 3.5.14 (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 3.4.3.
2. By Lemma 3.5.12 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 $|{}^\lambda\alpha| \leqslant \max(\alpha, \lambda)^+ \leqslant \kappa$.

$\Box_{3.5.14}$

Finally, we give a definition which may be useful later.

Definition 3.5.15 (AC). $A^{<\beta} = \{{}^\alpha A : \alpha < \beta\}$.

Note: If $\kappa \geqslant \omega$, then $|\kappa^{<\omega}| = \kappa$ and
$$|\kappa^{<\lambda}| = \sup\{\kappa^\theta : \theta < \lambda \wedge \theta \text{ is a cardinal.}\}.$$