Cardinal and Ordinal Numbers
Math 6300

Klaus Kaiser

April 9, 2007

Contents

Chapter 1

Introduction

In this course, we will develop set theory like any other mathematical theory – on the basis of a few given axioms and generally accepted practices of logic. When we are studying algebraic structures like groups, we have in mind structures like $\mathcal{S}_n$, $\mathbb{Z}$, $\mathbb{V}$, $\mathbb{G}$ whose elements are called permutations, integers, vectors or just group elements, and we use suggestive notations like ϕ, n, v, g to denote the objects these groups are made of. For groups, we also have a binary operation acting on the elements which we like to denote as *composition*, *addition*, *translation* or *multiplication* and we use suggestive symbols, like "$\circ$" for composition and "$+$" for addition. A group has a unique *identity* element which is usually denoted as "e". As a matter of convenience, an additional unary operation is then added, which assigns to a group element its unique *inverse*. The theory of groups is then governed by a few simple axioms $x(yz) = (xy)z$, $xx^{-1} = x^{-1}x = e$ and $xe = ex = x$. Groups are then developed solely on the basis of these axioms with the aforementioned examples serving as illustrations and motivations.

The situation for set theory is somewhat different. Unless you have already seen some axiomatic set theory or mathematical logic, you probably have not the fuzziest idea about different models of set theory. Sets are just arbitrary collections of objects and manipulations of sets , like forming *intersection*, *union*, and *complement* correspond to basic logical connectives, namely *and*, *or*, *not*. It seems that there is only one universe of sets. Our knowledge about it, however, may increase over time. Before *Borel* and *Lebesgue*, mathematicians didn't recognize measurable sets of real numbers. But they were there, just as the planet Pluto existed before it was discovered around 1930. Of course, mathematical objects are not physical. They are mental constructs owing their existence to our ability to speak and think. But is it safe to speak about all possible sets? Indeed, most mathematicians believe that it is safe to accept the idea of a universe of sets in which all of mathematics is performed. Mathematical statements then should be either true or false even when we know that they are undecidable right now. There is a belief that further insights will eventually resolve all open problems one way or the other. However, certain precautions must be exercised in order to avoid inconsistencies., like *Russel's* paradox about the set of all sets which don't contain themselves. (If $r(x)$ stands for the predicate $not(x \in x)$, then forming the Russel class $r = \{x|r(x)\}$ leads to the contradiction $(r \in r)$ iff $not(r \in r)$. In *Naive Set Theory*, methods for constructing new sets from given ones are presented and some sort of "etiquette" for doing it right is established. Such an approach, however, can be confusing. For example, most mathematicians don't feel any need for *Kuratowski's* definition of an ordered pair (a, b) as the set $\{\{a\}, \{a, b\}\}$ or to go through a lengthy justification that $s(n, 0) = n, s(n, m') = s(n, m)'$ defines the sum of two natural numbers n and m. On the other hand, certain proofs in analysis where sequences $s(n)$ are constructed argument by argument leaving at every n infinitely many options open, certainly take some time for getting used to. How in the world can we talk about the sequence $s(n), n \in \mathbb{N}$ as a finished product, when at point n we just don't know what $s(m)$ for $m \geq n$ will be? Sure, the

Axiom of choice is supposed to do the job. But what kind of axiom is this anyway? Is it part of our logic, or is it a technical property of a theory of sets? How come that generations of mathematicians didn't notice that they had been using it all the time? In *Axiomatic Set Theory* we assume that there is a mathematical structure $\mathcal{U}$ which we call the *universe* and whose elements are called sets. On $\mathcal{U}$ a binary relation $\in$ is defined which is called the *membership* relation. The basic assumption then is that $(\mathcal{U}, \in)$ satisfies the *Zermelo-Fraenkel Axioms* of set theory. We think that $\mathcal{U}$ is a set in the naive, familiar sense whose objects are called sets. Because we don't want $\mathcal{U}$ to be a member of $\mathcal{U}$, we call $\mathcal{U}$ the universe. This is mainly a precaution which we exercise in other branches of mathematics, too. A function space consists of functions, but is itself not a function. The relation $\in$ is a relation between sets in $\mathcal{U}$. We use the notation $\in$ only for this relation, in particular, instead of $x \in \mathcal{U}$ we say *x is in $\mathcal{U}$* or that *x belongs to $\mathcal{U}$*. We are now going to describe the axioms of set theory. These are statements about the universe of sets every mathematician would consider as self-evident. We are going to claim that there are sets, in particular an *empty set* and an infinite set and that we can construct from given sets certain new sets, like the union and the power set of a set. There is one difficulty in defining sets by properties. Because we only have the membership relation $\in$ at our disposal, any property about sets should be expressed in terms of $\in$ and logical procedures. For this reason we have to develop a language of axiomatic set theory first. The existence of sets sharing a common property is then governed by the axiom of comprehension. It will turn out that for example $x = x$ or not $(x \in x)$ never define sets, no matter what the universe is, resolving Russell's paradox. A certain amount of mathematical logic seems to be unavoidable in doing axiomatic set theory. But all that is necessary is to explain the syntax of the first order language for set theory. We do not have to say what a formal proof is. Similarly, the interpretation of formulas in the model $(\mathcal{U}, \in)$ is considered as self evident; delving into semantic considerations is equally unnecessary. Also, this is standard mathematical practice. In algebra you have no problems understanding the meaning of any particular equation, say the commutative law, but it needs to be made clear what a polynomial as a formal expression is. Because in axiomatic set theory we have to make statements concerning all formulas, we have to say what a formula is. Only the most important facts about set theoretic constructions, cardinals and ordinals are discussed in this course. Advanced topics of topology, for example, need more set theory. But these notes contain enough material for understanding classical algebra and analysis.

References

K. Devlin, *Fundamentals of Contemporary Set Theory.* Springer (1979).

K. Devlin, *The Joy of Sets.* Springer (1993).

K. Hrbacek, T. Jech, *Introduction to Set Theory.* Marcel Dekker, Inc. (1984).

J. L. Krivine, *Introduction to Axiomatic Set Theory.* Reidel (1971).

K. Kunen, *Set Theory.* North Holland (1980).

Y. Moschovakis, *Notes on Set Theory.* Springer (1994).

J. D. Monk, *Introduction to Set Theory.* McGraw-Hill Book Company (1969)

J. H. Shoenfield, *Mathematical Logic.* Addison Wesley (1967).

R. Vaught, *Set Theory.* Birkhäuser (1994).

These are very good text books on set theory and logic. The book by Monk is still useful for learning the basics of cardinal and ordinal arithmetic. Devlin's 93 book contains a chapter on recent

research on P. Aczel's Anti-Foundation-Axiom. The books by Kunen, Krivine and Shoenfield are advanced graduate texts, i.e., aimed at students who want to specialize in logic. The book by Kunen is a comprehensive text on set theory while Krivine is a good introduction into the classical relative consistency proofs, that is, the ones based on inner models of set theory. Shoenfield contains a final, far reaching, chapter on set theory. The following two articles are quite interesting. Shoenfield analyzes the truth of the ZF axioms, while Hilbert outlines the transition from finitary, constructive mathematics (which underlies, for example, our intuitive understanding of the natural numbers as well as the syntax of logic) towards a formalistic point of view about mathematics.

D. Hilbert, *Über das Unendliche.* Math. Annalen 25 (1925).

J. H. Shoenfield, *Axioms of Set Theory.* In: Handbook of Mathematical Logic. (North Holland, Amsterdam)

The Independence Proofs of Cohen are clearly presented in:

P. J. Cohen, *Set Theory and the Continuum Hypothesis.* Benjamin (1966).

The formal analysis of logic and set theory has important practical applications in form of non-standard methods. There is an extensive literature on this vital subject. The following books are exceptionally well written; the book by Robinson is a classic of this field.

R. Goldblatt, *Lectures on the Hyperreals. An Introduction to Nonstandard Analysis.* Springer (1998).

A. E. Hurd, P.A. Loeb, *An Introduction to Nonstandard Real Analysis.* Academic Press (1985).

E. Nelson, *Radically Elementary Probability Theory.* Princeton (1987).

A. Robinson, *Non-Standard Analysis.* North-Holland (1974).

The *Axiom of Foundation,* that is, a set cannot contain itself, should be true for the universe of sets, but it does not have any significant consequences. So we have separated it from the other axioms and develop set theory without this axiom. The following two books analyze strong negations of the Foundation Axiom and provide applications to self referential statements and Computer Science.

P. Aczel, *Non-Well-Founded-Sets.* Center for the Study of Language and Information Publications, Stanford (1988).

J. Barwise, J. Etchemendy, *The Liar.* Oxford (1987).

A good deal of the history of modern set theory is contained in

John W. Dawson, Jr. *Logical Dilemmas, The Life and Work of Kurt Gödel.* A K Peters, Wellesley, Massachusetts (1997).

Chapter 2

The Zermelo Fraenkel Axioms of Set Theory

The Axiom of Extensionality. If every element of the set a is an element of the set b and every element of the set b is an element of the set a, then $a = b$.

In other words, two sets are equal iff they contain the same elements. This should not be considered as a definition of equality of sets. Equality is an undefined, primitive relation and clearly, equal sets have the same elements. The axiom of extensionality merely states a condition on the relation $\in$. We may formalize extensionality:

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

The elements of the universe $(\mathcal{U}, \in)$ are in the first place just objects without any structure. What matters is their relationship to other elements with respect to $\in$. We may think of $\mathcal{U}$ as a directed graph where the sets in $\mathcal{U}$ are nodes and $a \in b$ corresponds to an edge $a \leftarrow b$. Part of the universe may have nodes called 0, 1, 2, $\{1\}$ and edges $0 \leftarrow 1$, $0 \leftarrow 2$, $1 \leftarrow 2$, $1 \leftarrow \{1\}$:

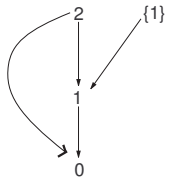

Figure 2.1: Snapshot of the Universe

An edge $0 \leftarrow \{1\}$ would violate the axiom of extensionality, because then 2 and $\{1\}$ would have the same elements.

The Null Set Axiom. There is a set with no elements:

$$\exists x \forall y \, \neg(y \in x)$$

By extensionality, there is only one such set. It is denoted by $\emptyset$ and called the *empty set*. It is a *constant* within the universe $\mathcal{U}$, i.e., a unique element defined by a formula.

The Pairing Axiom. For any sets a and b there is a set c whose only elements are a and b:

$$\forall x \forall y \exists z \forall t \big[(t \in z) \leftrightarrow \big((t = x) \vee (t = y)\big)\big]$$

By extensionality again, there is for given a and b only one such set c. We write $c = \{a, b\}$ for the set whose only elements are a and b. If a and b are the same set, then c has only one element, namely a. That is, for any set a of the universe $\mathcal{U}$ there is a set whose only element is a. This set is called the *singleton* $\{a\}$; $\{a, b\}$ is called a *pair* if a is different from b. Three applications of the pairing axiom lead to the existence of the set $\{\{a\}, \{a, b\}\}$. This is *Kuratowski's* definition of the *ordered pair* (a, b) of a and b. One easily proves the

Theorem 2.1 *One has that* $(a, b) = (a', b')$ *if and only if* $a = a'$ *and* $b = b'$.

The Union Axiom. For any set a there is a set b whose members are precisely the members of members of a:

$$\forall x \exists y \forall z \big[(z \in y) \leftrightarrow \exists t\big((t \in x) \wedge (z \in t)\big)\big]$$

The set b is called the *union* of a and denoted by $\bigcup a$ or $\bigcup\{x | x \in a\}$. We mention some consequences:

- For any sets a, b, c there is a set d whose elements are a, b and c:
  $d = \bigcup\{\{a, b\}, \{c\}\}$
- The union of $c = \{a, b\}$ is denoted by $a \cup b$. It is easy to see that
  $a \cup b = \{x | x \in a \text{ or } x \in b\}$.

Let a and b be sets. We say that a is a subset of b if every element of a is also an element of b:

$$(x \subseteq y) \equiv \forall z \big[(z \in x) \rightarrow (z \in y)\big]$$

The left-hand side is not a formula, because $\in$ is the only relation of our universe; $(x \subseteq y)$ is only an abbreviation of the formula in the variables x and y on the right hand side. In particular we have by extensionality that

$$\forall x \forall y \big[(x = y) \leftrightarrow \big((x \subseteq y) \wedge (y \subseteq x)\big)\big]$$

The Power Set Axiom. Let a be a set of the universe $\mathcal{U}$. Then there is a set b whose elements are precisely the subsets of a:

$$\forall x \exists y \forall z \big[(z \in y) \leftrightarrow (z \subseteq x)\big]$$

The set b is called the power set of a and we use the notation $b = \mathcal{P}(a)$. We have $\mathcal{P}(\emptyset) = \{\emptyset\}$, $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$, $\mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.

If a is any set of our universe, any $c \in \mathcal{P}(a)$ corresponds to an intuitive subset of a, namely $\{d | d \leftarrow c\}$ where for each such d, $d \leftarrow a$ holds. However, not every proper collection of edges $d \leftarrow a$ will lend itself to a set c of the universe. For example, if $\mathcal{U}$ happens to be countable then any infinite set a in $\mathcal{U}$ will have "subsets" which don't correspond to sets in $\mathcal{U}$. What kind of properties now lead to subsets? We have reached the point where we have to talk a bit about mathematical logic.

The Language of Axiomatic Set Theory

We are going to describe a *formal language* that has the following ingredients.

1. Symbols
   (a) An **unlimited** supply of *variables* x_0, x_1, $x_2 \ldots$.

(b) The elements of the universe $\mathcal{U}$ are the *constants* of the language.

(c) The *membership* symbol $\in$ and the *equality* symbol $=$.

(d) The symbols for the *propositional connectives*: $\wedge$ which stands for **and**, $\vee$ which stands for **or**, $\neg$ which stands for **not**, $\rightarrow$ which stands for **if, then**, $\leftrightarrow$ which stands for **if and only if**.

(e) For each variable x_n one has the *universal quantifier* $\forall x_n$ which stands for **for all** x_n and the *existential quantifier* $\exists x_n$ which stands for **there exists some** x_n.

2. Formation Rules for Formulas

(a) Let u and v stand for any variable or constant. Then $(u \in v)$ and $(u = v)$ are formulas. These are the *atomic* formulas.

(b) If P and Q are formulas then $(P \wedge Q)$, $(P \vee Q)$, $\neg P$, $(P \rightarrow Q)$, $(P \leftrightarrow Q)$ are formulas.

(c) If P is a formula then $\forall x_n P$ and $\exists x_n P$ are formulas.

Only expressions that can be constructed by finitely many applications of these rules are formulas. For better readability, different kinds of parentheses will be used, and letters, like x, y, z, ... will stand for variables. There are standard conventions concerning the priorities of the binary propositional connectives in order to avoid an excessive accumulation of parentheses.

The axioms of set theory as stated so far are all formulas, actually *sentences*, that is, all occurrences of variables are *bound*. If Q is a formula then every occurrence of x_n within P of a subformula $\forall x_n P$ or $\exists x_n P$ of Q is said to be bound. Variables x_n which are not bound, i.e., which are not within the *scope* of a quantifier $\forall x_n$ or $\exists x_n$ of Q, are said to be *free*. If we underline in a formula a variable then this variable is meant to occur only bound.

Formulas can be represented by certain labelled, directed trees. An atomic formula is just a node, e.g.,

$$(x \in a)$$

which is a tree. If Γ_1 is the tree for P_1 and if Γ_2 is the tree for P_2 , then the tree for $(P_1 \wedge P_2)$ is the graph:

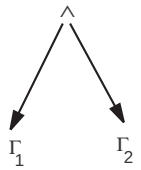

Figure 2.2: The Graph of a Conjunction

Any node of the tree Γ for the formula Q determines a subformula P of Q. For example, a node labelled $\wedge$ determines a conjunction $P \equiv (P_1 \wedge P_2)$ as a subformula of Q, where P_1 and P_2 are subformulas of P; P_1 and P_2 are the scope of the node $\wedge$. Similarly, a node $\forall x$ determines a subformula $P \equiv \forall x_n P_1$, where the subformula P_1 of P is the scope of the node $\forall x_n$ within Q.

Whenever we indicate a formula P as $P(x_0, x_1, \ldots, x_{n-1})$, it is understood that the free variables of P, if there are any, are are among $x_0, x_1, \ldots, x_n$. The constants within a formula are often called *parameters*. So we write $P(x_0, \ldots, x_{n-1}, a_0, \ldots, a_{m-1})$ to indicate the free variables and parameters of a formula. A sentence P is either true or false in the universe $\mathcal{U}$. More generally, if $P(x_0, \ldots, x_{n-1})$ is a formula with free variables $x_0, \ldots, x_{n-1}$ and if $a_0, \ldots, a_{n-1}$ belong to $\mathcal{U}$, then a simultaneous

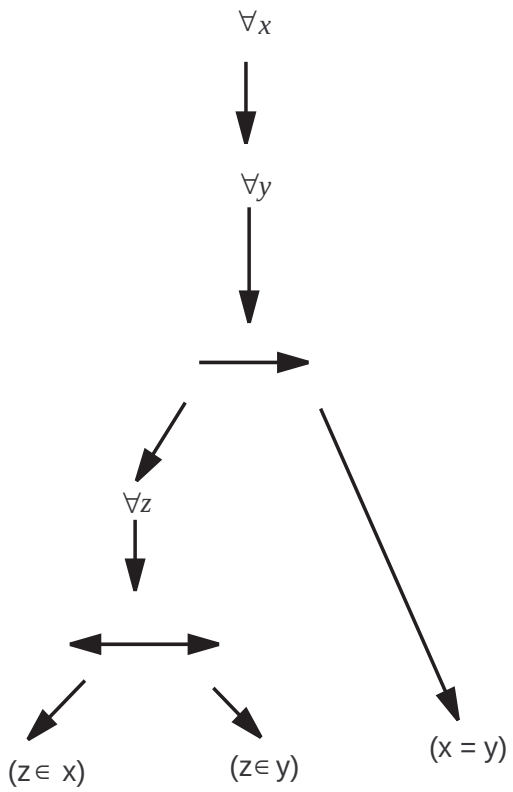

Figure 2.3: The Graph of the Extensionality Axiom

substitution of the x_i by the a_i makes $P(a_0, \ldots a_{n-1})$ either true or false. When we say that a formula $P(x_0, \ldots x_{n-1})$ holds on $\mathcal{U}$, it is meant that its *closure*, i.e.,

$$\forall x_0 \ldots \forall x_{n-1} P(x_0, \ldots, x_{n-1})$$

holds on $\mathcal{U}$. Because we have used the equality sign $=$ as a symbol within the language, equality of formulas, or more generally their equivalence, is denoted by $\equiv$, e.g., $x = y \equiv y = x$. That is, we write $P \equiv Q$ if and only if $P \leftrightarrow Q$ is a theorem of logic. Formulas without parameters are called *pure* formulas of set theory.

A formula in one free variable, or *argument*, is called a class.

$$S(x, a) \equiv (x \in a) \text{ and } R(x) \equiv \neg(x \in x)$$

are examples of classes. The first class defines a set, namely a, while the second one does not define a set: b satisfies $S(x, a)$ iff $b \in a$; there is no set r such that b satisfies $R(x)$ iff $b \in r$.

Formulas $P(x_0, \ldots, x_{n-1})$ are called *n-ary relations.* Formulas in two arguments are called *binary* relations. We also use the terms *predicates*, *properties* and *expressions* for formulas. Let $E(x)$ be a class. We say that $R(x, y)$ is a *relation on* $E(x)$ if

$$\forall x \forall y \big[R(x, y) \rightarrow \big(E(x) \wedge E(y) \big) \big]$$

holds on $\mathcal{U}$.

Let $R(x, y)$ be a binary relation. We define *domain* and *range* as the classes

$$\text{dom\_ of\_}R(x, \underline{y}) \equiv \exists y R(x, y) \text{ and } \text{ran\_of\_}R(\underline{x}, y) \equiv \exists x R(x, y)$$

Then $R(x, y)$ is a relation on

$$E(z) \equiv \text{dom\_ of\_}R(z, \underline{y}) \vee \text{ran\_of\_}R(\underline{x}, z)$$

which we call the *extent* of $R(x, y)$.

A binary relation $R(x, y)$ is called *reflexive* if

$$\forall x \forall y \big[R(x, y) \rightarrow \big((R(x, x) \wedge R(y, y) \big) \big]$$

holds on $\mathcal{U}$.

The relation $R(x, y)$ is *symmetric* if

$$\forall x \forall y \big[R(x, y) \rightarrow R(y, x) \big]$$

holds on $\mathcal{U}$.

The relation $R(x, y)$ is *transitive* if

$$\forall x \forall y \forall z \big[\big(R(x, y) \wedge R(y, z) \big) \rightarrow R(x, z) \big]$$

holds on $\mathcal{U}$.

The binary relation $E(x, y)$ is called an *equivalence* if it is reflexive, symmetric and transitive. It is easy to see that for any reflexive relation, e.g., an equivalence $E(x, y)$ one has that,

$$\text{dom\_of\_}E(x, \underline{y}) \equiv E(x, x) \text{ and } \text{ran\_of\_}E(\underline{x}, y) \equiv E(y, y)$$

and therefore $E(x, y)$ is a relation on its domain $D(x)$.

The binary relation $R(x,y)$ is called *anti-symmetric* if

$$\forall x \forall y \big[\big(R(x,y) \wedge R(y,x) \big) \rightarrow (x = y) \big]$$

holds on $\mathcal{U}$.

A binary relation $PO(x,y)$ which is reflexive, transitive and anti-symmetric is called a *partial order*. Again we have by reflexivity that domain and range define the same class and that $PO(x,y)$ is a relation on its domain. A partial order $L(x,y)$ is called *linear* or *total*, if

$$\forall x \forall y \big[\big(D(x) \wedge D(y) \big) \rightarrow \big(L(x,y) \vee L(y,x) \big) \big]$$

holds on $\mathcal{U}$. $D(x)$ denotes the domain of $L(x,y)$.

An $(n+1)$-ary relation $F(x_0, \ldots, x_{n-1}, y)$ is called *functional* if

$$\forall x \ldots \forall x_{n-1} \forall y_1 \forall y_2 \big[\big(F(x_0, \ldots, x_{n-1}, y_1) \wedge F(x_0, \ldots, x_{n-1}, y_2) \big) \rightarrow (y_1 = y_2) \big]$$

holds on $\mathcal{U}$.

We define for any relation $P(x_0, \ldots, x_{n-1}, y)$ domain and range

$$\begin{aligned} \text{dom\_of\_}P(x_0, \ldots, x_{n-1}, \underline{y}) &\equiv \exists y P(x_o, \ldots, x_{n-1}, y) \\ \text{ran\_of\_}P(\underline{x_0}, \ldots, \underline{x_{n-1}}, y) &\equiv \exists x_0 \ldots \exists x_{n-1} P(x_0, \ldots, x_{n-1}, y) \end{aligned}$$

The domain is an $n-ary$ relation $D(x_0, \ldots, x_{n-1})$ while the range is a class $R(y)$.

The binary relation

$$P(x,y) \equiv \forall z \big[(z \in y) \leftrightarrow (z \subseteq x) \big]$$

is functional in the variable x. It assigns to a set a the power set $b = \mathcal{P}(a)$. We have that

$$\text{dom\_of\_}P(x, \underline{y}) \equiv (x = x) \;\; \text{and ran\_of\_}P(x,y) \equiv \exists x \forall z \big[(z \in y) \leftrightarrow (z \subseteq x) \big]$$

Instead of $P(x,y)$ we will often use the more suggestive notation $y = \mathcal{P}(x)$. We similarly write $z = (x,y)$, $z = \{x,y\}$ and $z = x \cup y$ for the corresponding functional predicates.

We define for a formula $F(x,y,z)$ the expression

$$fun\big(F(\underline{x}, \underline{y}, z)\big) \equiv \forall x \forall y_1 \forall y_2 \big[F(x, y_1, z) \wedge F(x, y_2, z) \rightarrow y_1 = y_2 \big]$$

which holds for a set a in $\mathcal{U}$ if and only if $F(x,y,a)$ is functional.

The Schemes of Replacement and Comprehension

In the previous section we didn't stipulate the existence of sets. For example, domain and range of a binary relation were defined as classes, i.e., as formulas in one variable. Of course, given a binary relation on a given set a, domain and range should be subsets of a. The existence of sets according to standard constructions in mathematics is guaranteed by

The Axiom Scheme of Replacement. Let $F(x,y,x_0, \ldots, x_{n-1})$ be a pure formula of axiomatic set theory such that for sets $a_0, \ldots, a_{n-1}$ the binary relation $F(x,y,a_0,a_1, \ldots, a_{n-1})$ is functional. Let a be any set. Then there is a set b such that $d \in b$ holds if and only if there is some $c \in a$ such that $F(c,d,a_0, \ldots, a_{n-1})$ holds on $\mathcal{U}$:

$$\forall x_0 \ldots \forall x_{n-1} \big(fun\big(F(\underline{x}, \underline{y}, x_0, \ldots, x_{n-1})\big) \rightarrow \forall x \exists y \forall v [v \in y \leftrightarrow \exists u [u \in x \wedge F(u, v, x_0, \ldots, x_{n-1})]] \big)$$

Because this is supposed to hold for **every** pure formula $F(x, y, x_0, \ldots, x_{n-1})$, where at least x and y are free, this list of axioms is called a *scheme*. It is called replacement because it allows us to replace some of the elements c of the set a simultaneously by sets d in order to create a set b. As a first application of replacement we will deduce its weaker cousin

The Scheme of Comprehension. Let $A(x, x_0, \ldots, x_{n-1})$ be a pure formula of axiomatic set theory and let $a_0, \ldots, a_{n-1}$ be sets. Then for any set a there is a set b which consists exactly of those elements c of a for which $A(c, a_0, \ldots, a_{n-1})$ holds on $\mathcal{U}$:

$$\forall x_0 \ldots \forall x_{n-1} \forall x \exists y \forall z \big[z \in y \leftrightarrow \big(z \in x \wedge A(z, x_0, \ldots, x_{n-1}) \big) \big]$$

In order to deduce this from replacement, we only have to note that

$$F(x, y, x_0, \ldots, x_{n-1}) \equiv \big(y = x \wedge A(x, x_0, \ldots, x_{n-1}) \big)$$

is functional.

The standard notation for the subset b of a, which is defined by the property $A(x, a, .., a)$, is

$$b = \{x | x \in a \wedge A(x, a_0, \ldots, a_{n-1})\}$$

Constructions within the Universe

The existence of the union of a set a was stipulated as an axiom. We don't need a further axiom for the intersection.

The Intersection of a Set. Let a be **non-empty** set. Then there is a set b whose members are precisely the members of all members of a.

$$\forall x \big[\neg (x = \emptyset) \rightarrow \exists y \forall z \big[(z \in y) \leftrightarrow \forall t \big(t \in x \rightarrow z \in t \big) \big] \big]$$

This follows at once from comprehension. Note that the intersection of the set a is contained in any of its members c. The standard notation for the intersection of a set a is $\bigcap a$ or $\bigcap \{x | x \in a\}$. Why is it important to assume that the set a is non-empty?

The Cartesian product of Two Sets. Let a and b be sets. Then there is a set c such that $e \in c$ if, and only if, $e = (f, g)$ where $f \in a$ and $g \in b$:

$$\forall x \forall y \exists z \forall t \big[(t \in z) \leftrightarrow \exists u \exists v \big(t = (u, v) \wedge (u \in x) \wedge (v \in y) \big) \big]$$

The equation $z = (x, y)$ is shorthand for the functional relation $Q(x, y, z)$ which says that z is the ordered pair (x, y), which according to Kuratowski's definition is the set $\{\{x\}, \{x, y\}\}$. Thus:

$$Q(x, y, z) \equiv \forall t \Big[(t \in z) \leftrightarrow \exists u \exists v \Big\{ \big(t = u \vee t = v \big) \wedge \forall s \big[s \in u \leftrightarrow s = x \big] \wedge \forall s' \big[s' \in v \leftrightarrow \big(s' = x \vee s' = y \big) \big] \Big\} \Big]$$

If $= (f, g) = \{\{f\}, \{f, g\}\}$ where $f \in a$ and $g \in b$ then $\{f\} \in \mathcal{P}(a)$ and $\{f, g\} \in \mathcal{P}(a \cup b)$. Hence $(f, g) \in \mathcal{P}(\mathcal{P}(a \cup b))$. We now apply comprehension to

$$P(z, a, b) \equiv \exists u \exists v \big[Q(u, v, z) \wedge u \in a \wedge v \in b \big]$$

which says that "z is an ordered pair whose two components belong to a and b", respectively and get the desired result as

$$c = \{e | e \in \mathcal{P}(\mathcal{P}(a \cup b)) \wedge P(e, a, b)\}$$

The set c is called the *cartesian product* $a \times b$ of a and b. The cartesian product of finitely many sets is similarly defined. The formula

$$C(x, y, z) \equiv \forall t \big[t \in z \leftrightarrow \exists u \exists v \big[Q(u, v, t) \wedge (u \in x) \wedge (v \in y) \big] \big]$$

is functional and says that z is the cartesian product of x and y.

We remark that a binary relation $R(x, y)$ may be perceived as a unary relation $R^*(z)$:

$$R^*(z) \equiv \exists x \exists y \big(Q(x, y, z) \wedge R(x, y) \big)$$

That is, $R^*(e)$ holds if and only if $e = (c, d)$ and $R(c, d)$ holds.

Relations as Sets. Let $R(x, y)$ be a binary relation. Assume that domain and range of $R(x, y)$ are sets a and b, respectively. Then define the set

$$r = \{e | e = (c, d) \in a \times b, R(c, d)\}$$

We now have $e = (c, d) \in r$ if and only if $R^*(e)$ holds. In this sense we may identify a binary relation, for which the extent is a set, by a set of ordered pairs.

Graphs of Functions. If the binary relation $F(x, y)$ is functional and the domain of $F(x, y)$ is a set a, then, according to replacement, the range is also a set. Let b be any set containing the range of $F(x, y)$. The set

$$\{e | e = (c, d) \in a \times b, F(c, d)\}$$

is called the graph of the function $f : a \to b$.

The *projections* are important examples of functional relations:

$$\begin{aligned} F_1(x, y, z, p) &\equiv Q(x, y, z) \wedge p = x \\ F_2(x, y, z, q) &\equiv Q(x, y, z) \wedge q = y \end{aligned}$$

We have that $F_1(a, b, c, d)$ holds on $\mathcal{U}$ iff $c = (a, b)$ and $d = a$, i.e., d is the first component of the ordered pair c. The predicate

$$P_1(t, x) \equiv \exists u \exists v \big[F_1(u, v, t, x) \big]$$

then holds if "x is the first component of the ordered pair t". That f is a function from a to b is expressed by $F(a, b, f)$ where

$$\begin{aligned} F(x, y, z) \equiv\ & \exists p \big[C(x, y, p) \wedge z \subseteq p \big] \wedge \forall u \big[u \in x \to \exists t \big(t \in z \wedge P_1(t, u) \big) \big] \wedge \\ & \forall t \forall t' \forall s \forall u \forall u' \big[t \in z \wedge t' \in z \wedge P_1(t, s) \wedge P_1(t', s) \wedge P_2(t, u) \wedge P_2(t', u') \to u = u' \big] \end{aligned}$$

The Exponentiation of Sets. Let a and b be sets. Then there is a set c whose elements are given by the functions $f : a \to b$.

The function $f : a \to b$ is an element of $\mathcal{P}(a \times b)$. Hence c is defined by comprehension:

$$c = \{f | f \in \mathcal{P}(a \times b), F(a, b, f)\}$$

For the set c one uses *exponential* notation $c = b^a$.

Union and Intersection of a Family of Sets. A function s with domain i is sometimes called a *family of sets* $a_j, j \in i$, where, of course $a_j = s(j)$. The union of the family s is the union of the range r of s, which is, according to the replacement axiom, a set u. We write $u = \bigcup\{a_j | j \in i\} = \bigcup s$. The intersection of a non-empty family s is defined similarly.

The Cartesian Product of a Family of Sets. Let s be a family of sets, indexed by the set i. A function $f : i \to u$ from i into the union u of the range r of s is called a *choice function* if for every $j \in i$ one has that $f(j) \in a_j$. Then there is a set c whose members are all the choice functions for s. This set is called the *cartesian product of the family* s and is denoted by $c = \prod\{a_j | j \in i\}$.

This follows from comprehension: We will use the expression $(x, y) \in z$ as shorthand for $\exists p(Q(x, y, p) \wedge (p \in z))$. Then $c = \{f | f \in u^i \wedge \forall x \forall y \forall z\big(\big((x, y) \in f \wedge (x, z) \in s\big) \to (y \in z)\big)\}$

The Remaining Axioms of ZF

Within the universe $\mathcal{U}$ we certainly can find the sets $\underline{0} = \emptyset$, $\underline{1} = \{\underline{0}\}$, $\underline{2} = \{\underline{0}, \underline{1}\}$,..., $\underline{n} = \{\underline{0}, \underline{1}, \ldots, \underline{n-1}\}$. Note that $\underline{n+1} = \underline{n} \cup \{\underline{n}\}$ where $\underline{n}$ is not a member of $\underline{n}$. Hence $\underline{n}$ has exactly n elements and $n \mapsto \underline{n}$ is an injective map from the "set" $\mathbb{N}$ of natural numbers into the universe $\mathcal{U}$. The sets $\underline{n}$ are called the *natural number objects* of $\mathcal{U}$. Notice that we have $n < m$ if and only if $\underline{n} \in \underline{m}$, and $\underline{n} \in \underline{m}$ is the same as $n \subset m$, $\subset$ standing for strict inclusion. On the basis of the axioms stated so far we have no way of telling whether there is a set whose elements are exactly the sets $\underline{n}$.

The Axiom of Infinity. There is a set ω whose elements are exactly the natural number sets $\underline{n}$.

This concludes the list ZF of axioms for axiomatic set theory. Our definition of ω as the set of all natural number sets $\underline{n}$ is only preliminary; it is not even given by a first order sentence of our language of set theory. After we have studied ordinals in general, ω will be defined as the set of all *finite* ordinal numbers[1]. Of course, there is no danger to think that the finite ordinals are just the ordinary finite numbers $\underline{n}$. And the vast majority of mathematicians feel that way. On the other hand, any axiomatic definition of ω allows for elements ν which are nonstandard, i.e., different from any ordinary number $\underline{n}$. However, whether one realizes this possibility or not seems to be irrelevant for the formal development of mathematics.

There are two more axioms most mathematicians consider as"true", the *Axiom of Choice* and the *Axiom of Foundation*. These axioms are listed separately, mainly because because a great deal of set theory can be developed without them.

The Axiom of Choice (AC). The Cartesian product of a family of non-empty sets is itself non-empty.

The Axiom of Foundation (AF). Every non-empty set a contains a set b which is disjoint to a.

Both axioms are independent of ZF. The axiom of choice is necessary for proving many essential theorems concerning infinite sets, e.g., that every vector space has a basis. The axiom of foundation provides the universe $\mathcal{U}$ with more structure in the sense that every set will have a rank as measure of its complexity. However, even strong negations of AF are consistent with ZF and such models of set theory have become an important research tool in computing science for the analysis of self-referential statements.

[1]More elementary is **Dedekind's** approach: For any set x, $x \cup \{x\}$ is called the *successor* x^+ of x. A set i is called *inductive* if we ave that $\emptyset \in i$ and $x \in i$ implies that x^+ is in i. Dedekind's version of the axiom of infinity then says that there is an inductive set i_0. It is obvious that i_0 must contain all $\underline{n}$. Then he defines the set ω of natural numbers as the intersection of all inductive sets. This can be done because it is enough to intersect all inductive subsets of i_0.

Chapter 3

Ordinals

3.1 Well-Orderings

A binary relation r on a set w is called a *well-ordering* if it is a partial order such that every non-empty subset has a smallest element. Because any pair has a minimum, r is actually a total order. Instead of $(c, d) \in r$ one writes $c \leq d$. Actually, for well-orderings one prefers the *irreflexive* or *strict version* of r. If r is a partial order then one has

(i') $r \setminus \Delta = r'$ is irreflexive, i.e., $(c, c) \notin r$.

(ii) r' is transitive.

The conditions (i') and (ii) imply that

(iii) r' is anti-symmetric, i.e., $c < d$ implies that $d < c$ cannot hold.

On the other hand, if r' is an irreflexive, transitive relation then $r = r' \cup \Delta$ is a partial order.

If we assume the axiom of infinity in it's naive version, i.e., ω is the set of the standard natural numbers $\underline{n}$, then a prime example of a well-ordered set is provided by ω, together with the relation $\in$ restricted to the set ω. We have:

$$n < m \text{ iff } \underline{n} \in \underline{m} \text{ iff } \underline{n} \subset \underline{m}$$

Because $\mathbb{N} = (N, <)$ is well-ordered, the same holds true for $(\omega, \in)$. Notice, that $\mathbb{N}$ lives outside the universe $\mathcal{U}$ and $n \mapsto \underline{n}$ is an isomorphism which is not an element of $\mathcal{U}$. That $\mathbb{N}$ is well-ordered by $<$ is equivalent to the induction axiom, which we may take for granted.

On the other hand, we can prove, with the help of the axiom of foundation, that $(\omega, \in)$ is well-ordered. Let b be a non-empty subset of ω. According to AF there is some $\underline{n}$ in b which is disjoint to b. That is, $\underline{m} \in \underline{n}$ yields $\underline{m} \notin b$. In other words, $\underline{n}$ is the smallest element of b.

Definition 3.1 A binary relation r on a set w is *well-founded* if there is no strictly decreasing map $f : \omega \to w$, $\underline{n} \mapsto a_{\underline{n}}$, i.e., $a_{\underline{0}} > a_{\underline{1}} > a_{\underline{2}} > \ldots$ which belongs to $\mathcal{U}$.

Proposition 3.1 *A well-ordering is well-founded.*

Proof. Otherwise the range $\{a_{\underline{0}}, a_{\underline{1}}, \ldots\}$ would be a set b without a smallest element. □

Definition 3.2 A binary relation r on a set a satisfies the *minimal condition* if any non-empty subset b of a contains a minimal element, i.e., there is some $c \in b$ such that for no $d \in b$ one has that $(d, c) \in r$.

Definition 3.3 Let $\leq$ be a partial order on the set a. Then any $c \in a$ determines the *(initial) segment* $s(c) = \{b | b \in a, b < c\}$.

Proposition 3.2 (AC) *A partial order $\leq$ on a set a which is well-founded satisfies the minimal condition.*

PROOF. Assume that a non-empty subset b of a does not have a minimal element. Then the set $d = \{s(c) | c \in b\}$ does not contain the empty set. Let f be a choice function on d. Pick any c_0 from b. Then $\underline{0} \mapsto c_{\underline{0}}, 1 \mapsto f(s(c_{\underline{0}})) = c_{\underline{1}}, 2 \mapsto f(s(c_{\underline{2}})) = c_{\underline{2}}, \ldots$ defines a strictly decreasing sequence which belongs to $\mathcal{U}$. (The proof that we can define in such a way a sequence is a good excercise.) This contradicts the well-foundedness of our relation. □

In particular:

Proposition 3.3 (AC) *A total, well-founded order is a well ordering.* □

As a generalization of complete induction for natural numbers we have:

Proposition 3.4 (The Principle of Proof by Induction) *Let $(w, <)$ be well-ordered system and let a be a subset of w. Assume that*

(i) *The set a contains the smallest element o of w.*

(ii) *Assume that $c \in a$ in case that $d \in a$ for all $d < c$.*

Then one has that $a = w$.

PROOF. Assume that $b = w \setminus a$ is non-empty. Let c be the smallest element of b. We have $c > o$ by (i). Now, any $d < c$ belongs to a and therefore c belongs to a, by (ii). But this contradicts the choice of c. □

Notice that (ii) actually implies (i). There is no $d < o$, and therefore $o \in a$, by default.

Definition 3.4 Let $R(x, y)$ be a partial ordering. The class $S(x)$ is called a *segment* of $R(x, y)$ if $\forall x \forall y\ R(x, y) \wedge S(y) \rightarrow S(x)$ holds on $\mathcal{U}$.

Proposition 3.5 *The proper segments of a well-ordered set w are exactly the segments $s(a)$.*

PROOF. Assume that the subset s of w is a proper segment. Then $w \setminus s$ is non-empty and has a smallest element a. But then $s = s(a)$. □

Corollary 3.6 *The map $a \mapsto s(a)$ is strictly increasing and establishes a bijective order preserving map between the well-ordered set $(w, <)$, and the ordered set system $\mathcal{S} = (\{s(a) | a \in w\}, \subset)$ of segments of w.* □

Definition 3.5 An injective, order preserving map map between partially ordered sets is called an *order embedding*. A bijective order embedding for which the inverse is also order preserving is called an *(order) isomorphism*. An order isomorphism of a partially ordered set to itself is called an *(order) automorphism*.

Proposition 3.7 *If f is a bijective order embedding from the totally ordered set $(a, \leq_1)$ to the partially ordered set $(b, \leq_2)$ then $\leq_2$ is a total order and f is an order isomorphism.* □

Corollary 3.8 *A well-ordered set is order isomorphic to its system of proper segments.* □

Proposition 3.9 *Let $(w, <)$ be a well ordered system and $f : w \to w$ be strictly increasing. Then one has $f(c) \geq c$ for all $c \in w$.*

PROOF. For the proof assume that the set $a = \{c | f(c) < c\}$ is non-empty. Then a has a smallest element b. Then $f(b) < b$ because $b \in a$. Hence $f(f(b) < f(b)$ because f is strictly increasing. Therefore, $f(b) \in a$. But then one has that $b \leq f(b)$, which is a contradiction. □

Proposition 3.10 *Let f be an automorphism of the well-ordered system $(w, <)$. Then f is the identity map.*

PROOF. Assume otherwise. Then $a = \{b | f(b) > b\}$ is non-empty and has a smallest element b_0. Then $f(b_0) > b_0$ because $b_0 \in a$. Let $b_0 = f(c_0)$. From $f(b_0) > f(c_0)$ one infers $b_0 > c_0$ and therefore, by the choice of b_0 , $f(c_0) = c_0$. Hence $b_0 = f(c_0) = c_0$, which is a contradiction. □

Corollary 3.11 *An isomorphism between isomorphic well-ordered systems $(w, <)$ and $(w', <')$ is unique.*

PROOF. If f and g are two such isomorphisms then $g^{-1} \circ f$ is the identity on w. That is, $f = g$. □

Corollary 3.12 *A well-ordered system $(w, <)$ is not isomorphic to any of its segments $s(a)$.*

PROOF. For any such isomorphism we would have $f(a) \geq a$, because of Proposition 3.9, and $f(a) \in s(a)$, i.e., $f(a) < a$. □

Corollary 3.13 *Assume that $s(a) \cong s(a')$ holds in the well-ordered system $(w, <)$. Then $a = a'$ and the isomorphism is the identity.*

PROOF. Assume $a < a'$. Then $s(a) \subset s(a')$ and the well-ordered system $w' = s(a')$ would be isomorphic to its segment $s(a)$. This contradicts Corollary 3.12. □

Theorem 3.14 *Any two well-ordered systems $(w_1, <_1)$ and $(w_2, <_2)$ are either isomorphic or one is isomorphic to a segment of the other one. Any such isomorphism is unique.*

PROOF. We cannot have $w_1 \cong (s(b), <_2)$ and $w_2 \cong (s(a), <_1)$ because this would yield $w_1 \cong s(b) \subset w_2 \cong s(a)$, hence $w_1 \cong s(a')$ for some $a' < a$. This contradicts the statement of Corollary 3.12. For the proof of the theorem we define a binary relation

$$F = \{(a, b) | s(a) \cong s(b)\} \subseteq w_1 \times w_2.$$

It is easy to see that the domain and the range of F are segments of w_1 and of w_2, respectively. Moreover, F is functional and strictly increasing. If we had $dom(F) = s(a')$ and $ran(F) = s(b')$ then F would establish an isomorphism $s(a') \cong s(b')$. Hence $(a', b') \in F$, and therefore $a' \in s(a')$, which is a contradiction. Hence, either $dom(F)$ or $ran(F)$, or both, are all of w_1 or w_2, respectively. So either $F : w_1 \to s(b')$ or $F : w_1 \to w_2$ or $F^{-1} : w_2 \to s(a')$. □

Definition 3.6 Let $R(x, y)$ be a total order relation on its domain $D(x)$. Then R(x,y) is called a well-ordering if for any element a in the domain one has that $S(x, a) \equiv R(x, a) \wedge x \neq a$ is a set and well-ordered by $R(x, y)$.

This terminology is justified by the following

Lemma 3.15 *Let $R(x, y)$ be a well-ordering with domain $D(x)$ and let $T(x)$ be any non-empty subclass of $D(x)$. Then $T(x)$ has a smallest element.*

PROOF. Let a be a set which satisfies $T(x)$ and $D(x)$. Then $T(x) \wedge S(x,a)$ is a set. If this set is empty, then a is the smallest element of $T(x)$. Otherwise it has a smallest element with respect to $R(x,y)$ which is then the smallest element of $T(x)$. □

Corollary 3.16 *Let $R_1(x,y)$ and $R_2(x,y)$ be two well-orderings with proper classes $D_1(x)$ and $D_2(y)$ as their domains. Then there is a unique functional relation $F(x,y)$ which defines an isomorphism between $D_1(x)$ and $D_2(x)$.*

PROOF. This is just a small modification of the proof for Theorem 3.14. As before, define

$$F(x,y) \equiv D_1(x) \wedge D_2(y) \wedge \{u|S_1(u,x)\} \cong \{v|S_2(v,y)\}.$$

Because $D_1(x)$ is a proper class, it cannot be isomorphic to any $S_2(y,b)$ because these segments are by definition sets. □

Corollary 3.17 *Let $R_1(x,y)$ and $R_2(x,y)$ be two well-orderings with domains $D_1(x)$ and $D_2(x)$, respectively. Assume that $D_1(x)$ is a set a and $D_2(x)$ is a proper class. Then a is order isomorphic to a unique initial segment $S(x,b)$ of $R_2(x,y)$.* □

The question now is whether there are any well-ordered classes. This brings us to ordinals.

3.2 Ordinals

A set a is *transitive* if every element b of a is a subset of a. That is, the set a satisfies the predicate

$$Trans(x) \equiv \forall z\big[(z \in x) \rightarrow (z \subseteq x)\big].$$

For example, every natural number set $\underline{n}$ is transitive. The set a is transitive if $c \in b$ and $b \in a$ implies that $c \in a$. This explains the terminology. The set α is called an *ordinal* if

(i) α is transitive.

(ii) $\in$ if restricted to α is a strict well-ordering of α.

The second condition is formalized by the predicate:

$$\begin{aligned} Well(z) \quad \equiv \quad & \forall x \forall y\big[\big(x \in z \wedge y \in z\big) \rightarrow \big(\neg(x \in y) \vee \neg(y \in x)\big)\big] \wedge \\ & \forall u \forall v \forall w\big[\big(u \in z \wedge v \in z \wedge w \in z \wedge u \in v \wedge v \in w\big) \rightarrow u \in w\big] \wedge \\ & \forall x\Big[\big(x \subseteq z \wedge x \neq \emptyset\big) \rightarrow \exists u\Big\{u \in x \wedge \forall y\big[y \in x \rightarrow \big(u \in y \vee u = y\big)\big]\Big\}\Big] \end{aligned}$$

Well(α) holds if $\in$ restricted to α is an irreflexive , transitive relation and where every subset s of α contains some β such that for any $\gamma \in s$ one has that $\gamma \notin \beta$. It is customary to denote ordinals by small Greek letters. The class of ordinals is defined by the predicate

$$Ord(x) \equiv Trans(x) \wedge Well(x)$$

Proposition 3.18 *Let α be an ordinal. Then α is not an element of itself.*

PROOF. That α is an ordinal means that $(\alpha, <)$ is a well-ordered system where $\beta < \gamma$ holds for elements $\beta, \gamma \in \alpha$ iff $\beta \in \gamma$. If we had $\alpha \in \alpha$, then α as an element of itself would violate the irreflexivity of $<$ on α. □

Proposition 3.19 *Every element β of an ordinal α is itself an ordinal.*

PROOF. We first have to show that β is transitive. Let $\gamma \in \beta$ and $\delta \in \gamma$. We need to show that $\delta \in \beta$. Because of the transitivity of α we have that the element β of α is also a subset of α, i.e., $\beta \subseteq \alpha$. Hence $\gamma \in \beta$ yields $\gamma \in \alpha$. But then, by transitivity of α again, $\gamma \in \alpha$ yields $\gamma \subseteq \alpha$. So $\delta \in \gamma$ gives $\delta \in \alpha$. Hence the elements δ , γ , β are all in α and $\delta < \gamma < \beta$. Hence $\delta < \beta$ and this is $\delta \in \beta$.
Because β is a subset of α, it is well-ordered by $\in$, restricted to β. □

Proposition 3.20 *The segments of an ordinal α are α and the elements of α.*

PROOF. Because α is well-ordered we have that the segments of α are α and the sets $s(\beta)$ for $\beta \in \alpha$. But $s(\beta) = \{\gamma | \gamma \in \alpha, \gamma < \beta\} = \{\gamma | \gamma \in \alpha, \gamma \in \beta\} = \{\gamma | \gamma \in \beta\} = \beta$. □

Corollary 3.21 *Let α and β be ordinals where $\beta \subset \alpha$. Then $\beta \in \alpha$.*

PROOF. We show that β is a proper segment of α. Indeed, let $\gamma \in \beta$ and $\delta < \gamma$ where $\delta \in \alpha$. But then $\delta \in \gamma$ by definition of $<$ on α. Because β is an ordinal we have that $\gamma \subset \beta$. So $\delta \in \beta$. Hence β is a segment and as a proper subset of α, it is an element of α. □

Corollary 3.22 *Let β and α be ordinals. Then $\beta \subset \alpha$ if and only if $\beta \in \alpha$.* □

Proposition 3.23 *Let α be any ordinal. Then $\alpha^+ = \alpha \cup \{\alpha\}$ is also an ordinal.*

PROOF. Let $\gamma \in \alpha^+$ and $\beta \in \gamma$. If $\gamma \in \alpha$ then $\gamma \subset \alpha$ and therefore $\beta \in \alpha$, and $\beta \in \alpha^+$ because of $\alpha \subseteq \alpha^+$. If $\gamma = \alpha$ then $\beta \in \alpha$ and $\beta \in \alpha^+$. Hence α^+ is transitive.
The elements of α^+ are the elements of α and α. It is easy to see that $\beta < \gamma$ iff $b \in \gamma$ defines a total ordering on α^+ with α as largest element. If s is any non-empty subset of α^+, then α is the minimum of s in case that $s = \{\alpha\}$, otherwise it is $min(s \cap \alpha)$. Hence α^+ is well-ordered by $\in$. □

Lemma 3.24 *Let α and β be ordinals. Then $\delta = \alpha \cap \beta$ is an ordinal which is equal to α or equal to β.*

PROOF. We show that δ is a segment of α. Clearly, $\delta \subseteq \alpha$. Let $\gamma \in \delta$ and $\rho < \gamma$ where $\rho \in \alpha$. According to the definition of $<$ on α we have that $\rho \in \gamma$. Now, $\gamma \in \delta \subseteq \beta$ and therefore $\gamma \in \beta$, i.e., $\gamma \subseteq \beta$ because β is an ordinal. Therefore, $\rho \in \beta$. Hence, $\rho \in \alpha \cap \beta = \delta$. Thus δ is a segment of α and therefore $\delta = \alpha$ or $\delta \in \alpha$. We need to show that $\delta \in \alpha$ implies that $\delta = \beta$. By symmetry we also have that $\delta \in \beta$ or $\delta = \beta$. But $\delta \in \alpha$ and $\delta \in \beta$ leads to $\delta \in \alpha \cap \beta = \delta$, i.e., $\delta \in \delta$ which is impossible for ordinals. □

Theorem 3.25 *Let α and β be ordinals. Then $\alpha = \beta$ or $\alpha \in \beta$ or $\beta \in \alpha$ and these cases are mutually exclusive*

PROOF. $\alpha \cap \beta = \alpha$ is the same as $\alpha \subseteq \beta$ and $\alpha \subset \beta$ is the same as $\alpha \in \beta$. The claim now follows immediately from the lemma. □

Let α and β be any ordinals. We define that $\alpha < \beta$ iff $\alpha \in \beta$. The relation

$$R(x, y) \equiv Ord(x) \wedge Ord(y) \wedge (x \in y)$$

is a strict order on its domain, which is the class $Ord(x)$. For any ordinal β, we have that $R(x, \beta)$ is the set β which is well-ordered by $R(x, y)$.

Theorem 3.26 *The class $Ord(x)$ of ordinals is well-ordered by the membership relation $\in$, i.e., by $\alpha < \beta$ iff $\alpha \in \beta$.* □

Proposition 3.27 *Let $\beta = \bigcup a$ be the union of a set a of ordinals. Then β is an ordinal and one has $\beta \geq \alpha$ for every $\alpha \in a$. Furthermore, if γ is an ordinal such that $\gamma \geq \alpha$ holds for each $\alpha \in a$, then $\gamma \geq \beta$. That is, every set of ordinals has a least upper bound within the class of ordinals.*

PROOF. By the very definition of the union of a set a, we have that the union β of a contains every member α of a: $\beta \supseteq \alpha$ for each $\alpha \in a$; and obviously, any set c with that property must contain β. For ordinals, $\beta \supseteq \alpha$ is the same as $\alpha < \beta$ or $\alpha = \beta$ and therefore we only have to show that β is an ordinal. We first show that β is transitive. So let $\gamma \in \beta$ and $\delta \in \gamma$. Then $\gamma \in \alpha$ for some $\alpha \in a$. Now α is an ordinal, therefore $\gamma \subset \alpha$. But then $\delta \in \alpha$ and $\delta \in \beta$ because of $\beta \supseteq \alpha$. Now let γ and ρ be any elements of β. Because they are ordinals we have that either $\gamma = \rho$ or $\gamma \in \rho$ or $\rho \in \gamma$. That is, β is totally ordered by $\in$. If c is any non-empty subset of β then $c \cap \alpha$ must be non-empty for at least one $\alpha \in a$. Let $\delta = min(c \cap \alpha)$ and let $\gamma \in c$. We cannot have $\gamma \in \delta \subset \alpha$ because δ was the smallest element of c in α. Hence $\delta \leq \gamma$. Therefore, β is well-ordered by $\in$. □

Theorem 3.28 *The class $Ord(x)$ of ordinals is a proper class, i.e., there is no set which contains all ordinals.*

PROOF. Assume that a set b contains all ordinals. Then, by comprehension, there would be a set a that consists exactly of all ordinals. By the previous proposition, $\beta = \bigcup a$ would be an ordinal $\beta$and therefore $\beta \in \beta^+ \in a$. But then $\beta \in \beta$ by the definition of union. But this is impossible for ordinals.□

Corollary 3.29 *Let $R(x, y)$ be any well-ordering where the domain $D(x)$ of $R(x, y)$ is a proper class. Then there is exactly one functional relation between $D(x)$ and $Ord(x)$ which defines an order isomorphism between these classes.* □

Corollary 3.30 *Let $(w, <)$ be any well-ordered system. Then there is exactly one ordinal α such that $(w, <)$ and $(\alpha, \in)$ are order isomorphic.* □

For any ordinal α one has that $\alpha^+ = \alpha \cup \{\alpha\}$ is the smallest ordinal greater than α. It is called the *successor* of α and commonly denoted as $\alpha + \underline{1}$. The successor of $\alpha + 1$ is denoted as $\alpha + \underline{2}$, etc. The empty set $\emptyset$ is the smallest ordinal $\underline{0}$, and $\underline{1} = \{\underline{0}\}$ is the successor of $\underline{0}$.

$$\underline{0} < \underline{1} < \underline{2} < \dots$$

are the finite ordinals $\underline{n}$ and

$$\omega = \bigcup \{\underline{n} | \underline{n} \in \omega\}$$

is the smallest infinite ordinal. The existence of this ordinal has been explicitly stated as an axiom. This slow march through the ordinals continues with

$$\omega + \underline{1} < \omega + \underline{2} < \dots$$

and then comes as the second infinite limit ordinal

$$\omega + \omega = \bigcup \{\omega + \underline{n} | \underline{n} \in \omega\} = \omega \times \underline{2}$$

that is the next ordinal after ω which is not a successor, i.e., a *limit ordinal*; $\underline{0}$ is the only *finite* limit ordinal.

Theorem 3.31 (Proof by Induction on Ordinals) *Let $P(x)$ be any property. Assume that:*

(i) *$P(\underline{0})$ holds.*

(ii) *$P(\alpha)$ holds provided $P(\beta)$ holds for all $\beta < \alpha$.*

Then P(x) holds for all ordinals.

PROOF. Indeed, if we had an ordinal γ for which $P(x)$ would not hold, then we could find a smallest such ordinal α. But then $P(\beta)$ for all $\beta < \alpha$ and therefore $P(\alpha)$ by condition (ii). Note that (ii) actually implies (i). □

Induction on Ordinals generalizes Complete Induction on $\mathbb{N}$. We are now going to describe how functions can be defined recursively on the ordinals. Let $F(x, y)$ be any functional relation on the class $D(x)$ and let $T(x)$ be a subclass of $D(x)$, i.e., $T(x) \to D(x)$ holds on $\mathcal{U}$. Then $F(x, y) \restriction T(x) \equiv T(x) \wedge F(x, y)$ is called the *restriction* of $F(x, y)$ to $T(x)$.
Whenever we perceive a functional relation as a set, then we are actually identifying the function with its graph. So, for example, $F(x, y) \restriction a$ stands for $graph\big(F(x, y) \restriction a\big) = \{c | c \in (a \times b), c = (d, e), F(d, e)\}$, where b is the range of $F(x, y) \restriction a$.

Theorem 3.32 (Definition by Recursion on Ordinals) *Let $H(x, y, z)$ be a functional relation where the domain is $D(x, y) \equiv Ord(x) \wedge (y = y)$. That is, for any ordinal α and any set a there is a unique set b such that $H(\alpha, a, b)$ holds on $\mathcal{U}$. For this we write: $z = H(x, y), Ord(x)$. Then there is a unique functional relation $F(x, y)$ on $Ord(x)$ such that for any ordinal α one has that*

$$F(\alpha, b) \text{ iff } H(\alpha, F \restriction \alpha, b)$$

i.e.,

$$F(\alpha) = H(\alpha, F \restriction \alpha)$$

PROOF. Let α be any fixed ordinal. We first show that there is a unique function f_α on α such that

$$(*)_\alpha \quad f_\alpha(\beta) = H(\beta, f_\alpha \restriction \beta),\ \beta < \alpha$$

Note, if one has that $H(\underline{0}, \emptyset, a_{\underline{0}})$, then necessarily

$$f_\alpha(\underline{0}) = a_{\underline{0}}, f_\alpha(\underline{1}) = H(\underline{1}, \{(\underline{0}, a_{\underline{0}})\}) = a_{\underline{1}}, f_\alpha(\underline{2}) = H(\underline{2}, \{(\underline{0}, a_{\underline{0}}), (\underline{1}, a_{\underline{1}})\}), \ldots$$

In order to have a formal proof for uniqueness, assume that we have functions f_α and g_α both satisfying the condition $(*)_\alpha$. Then $f_\alpha(\underline{0}) = g_\alpha(\underline{0}) = a_{\underline{0}}$. We wish to show that if f_α and g_α agree for all $\gamma < \beta, \beta \in \alpha$, then $f_\alpha(\beta) = g_\alpha(\beta)$. Now: $f_\alpha(\beta) = H(\beta, f_\alpha \restriction \beta) = H(\beta, g_\alpha \restriction \beta) = g_\alpha(\beta)$. Hence, by the induction principle, $f_\alpha(\beta) = g_\alpha(\beta)$ holds for all $\beta < \alpha$, i.e., $f_\alpha = g_\alpha$.

With respect to the existence of f_α, note that we have already $f_{\underline{0}}, f_{\underline{1}}, \ldots$ and we may form the set of all ordinals β for which f_β exists:

$$\tau = \{\beta | \beta < \alpha, \exists f_\beta \exists c_\beta\ f_\beta : \beta \to c_\beta, f_\beta(\gamma) = H(\gamma, f_\beta \restriction \gamma), \gamma < \beta\} = \{\beta < \alpha | \exists f_\beta\ (*)_\beta\}$$

We are going to show that τ is a segment of α. Let $\beta \in \tau$ and $\gamma < \beta$. Put $f_\gamma = f_\beta \restriction \gamma$. Then for $\delta < \gamma$,

$$f_\gamma(\delta) = (f_\beta \restriction \gamma)(\delta) = f_\beta(\delta) = H(\delta, f_\beta \restriction \delta) = H(\delta, (f_\beta \restriction \gamma) \restriction \delta) = H(\delta, f_\gamma \restriction \delta)$$

That is, f_γ satisfies $(*)_\gamma$. Hence, $\gamma \in \tau$, and moreover, we have shown that if $\beta \in \tau$ and $\gamma < \beta$, one has that $f_\gamma = f_\beta \restriction \gamma$. As a segment of the ordinal α, τ is also an ordinal $\leq \alpha$. In order to show that $\tau = \alpha$, we define a function f_τ by

$$f_\tau(\beta) = H(\beta, f_\beta),\ \beta \in \tau$$

We wish to show that f_τ satisfies $(*)_\tau$. To this end let $\gamma < \beta < \tau$. Then one has:

$$f_\tau(\gamma) = H(\gamma, f_\gamma) = H(\gamma, f_\beta \restriction \gamma) = f_\beta(\gamma)$$

Hence: $f_\tau \restriction \beta = f_\beta$ and therefore, $f_\tau(\beta) = H(\beta, f_\tau \restriction \beta)$ holds for all $\beta < \tau$. Thus the function f_τ satisfies $(*)_\tau$. If we had $\tau < \alpha$, then $\tau \in \tau$ by the definition of τ. But $\tau \in \tau$ is impossible for ordinals as we have shown before.

The functions f_α, α an ordinal, form a class $G(x)$. The predicate $G(x)$ is the formalization of: "f is a function, $dom(f)$ is some ordinal α, $\forall_{\beta<\alpha}\big[f_\alpha(\beta) = H(\beta, f_\alpha \restriction \beta)\big]$". As we already have shown, the functions for $G(x)$ are pairwise compatible, i.e., restrictions of each other. Hence they can be glued together to a functional relation $F(x, y)$ where we have:

$$F(\alpha, b) \text{ iff } \exists\tau\big[f_\tau(\alpha) = b\big] \text{ iff } f_{\alpha^+}(\alpha) = b \text{ iff } H(\alpha, f_{\alpha^+} \restriction \alpha) = b \text{ iff } H(\alpha, F \restriction \alpha) = b \text{ , i.e.}$$

$$F(\alpha) = H(\alpha, F \restriction \alpha)$$

Finally, we must show uniqueness of our $F(x, y)$. So assume that $F'(x, y)$ also satisfies the recursion formula: $F'(\alpha) = H(\alpha, F' \restriction \alpha)$. Put $F' \restriction \alpha = f'_\alpha$ and let $\beta < \alpha$. Then:

$$f'_\alpha(\beta) = F'(\beta) = H(\beta, F' \restriction \beta) = H(\beta, (F' \restriction \alpha) \restriction \beta) = H(\beta, f'_\alpha \restriction \beta).$$

According to the very first part of the proof we conclude that $f_\alpha = f'_\alpha$ holds for every α. Hence $F(x, y) \equiv F'(x, y)$. □

Here is a first application of the Recursion Principle: Define a functional relation

$$H(x, y) = \begin{cases} \bigcup\{\mathcal{P}(f(\beta))|\beta < \alpha\} & \text{if } x \text{ is an ordinal } \alpha \text{ and } y \text{ a function } f \text{ on } \alpha \\ \\ \emptyset & \text{otherwise} \end{cases}$$

This leads to a recursively defined functional relation $V(x)$ on the ordinals: $V(\alpha) = H(\alpha, V \restriction \alpha) = \bigcup\big\{\mathcal{P}(V(\beta))|\beta < \alpha\big\}$. We have $V(\underline{0}) = \bigcup\emptyset = \emptyset$ and if $\beta < \alpha$, $V(\beta) = \bigcup\big\{\mathcal{P}(V(\gamma))|\gamma < \beta\big\} \subseteq \bigcup\big\{\mathcal{P}(V(\gamma))|\gamma < \alpha\big\} = V(\alpha)$. Because taking the power set is a monotone operation, we have $V(\alpha^+) \supseteq \mathcal{P}(V(\alpha)) \supseteq \mathcal{P}((V(\gamma)), \gamma \leq \alpha$. Thus, $V(\alpha^+) = \bigcup\big\{\mathcal{P}(V(\gamma))|\gamma \leq \alpha\big\} = \mathcal{P}(V(\alpha))$. If α is a limit ordinal then $\beta < \alpha$ yields $\beta^+ < \alpha$. Hence, $V(\alpha) = \bigcup\big\{\mathcal{P}(V(\beta))|\beta < \alpha\big\} = \bigcup\big\{V(\beta^+)|\beta < \alpha\big\} = \bigcup\{V(\beta)|\beta < \alpha\big\}$.

The Zermelo-Fraenkel Hierarchy of Sets. The cumulative hierarchy of sets is defined by

$$V_{\underline{0}} = \emptyset\,; V_{\alpha^+} = \mathcal{P}(V_\alpha)\,; V_\alpha = \bigcup\big\{V_\beta|\beta < \alpha\big\}, \text{ if } \alpha \text{ is a limit ordinal.}$$

$V(x) = \exists z\big[Ord(z) \wedge x \in V_z\big]$ is the class of sets belonging to that hierarchy.

In order to define addition of ordinals, we take the functional relation:

$H(x, y) = \beta$ if $x = 0$; $H(x, y) = f(\alpha)^+$ if $x = \alpha^+$ and y is a function f on α^+ ; $H(x, y) = \bigcup\big\{f(\gamma)|\gamma < \alpha\big\}$ if x is is a limit ordinal α, y a function on α; $H(x, y) = \emptyset$, otherwise.

We then have a unique function $S(x)$ on $Ord(x)$ such that $S(\alpha) = H(\alpha, S \restriction \alpha)$. This is $S(\underline{0}) = \beta$; $S(\alpha^+) = S(\alpha)^+$; $S(\alpha) = \bigcup\big\{S(\gamma)|\gamma < \alpha\big\}$. Instead of $S(\alpha)$ we write $(\beta + \alpha)$.

The Addition of Ordinals $\beta + \underline{0} = \beta$; $(\beta + \alpha^+) = (\beta + \alpha)^+$; $(\beta + \alpha) = \bigcup\big\{\beta + \gamma|\gamma < \alpha\big\}$ if $\alpha \neq \underline{0}$ is a limit ordinal, defines a unique operation on $Ord(x)$. It generalizes the ordinary addition of natural numbers.

The Multiplication of Ordinals $\beta \cdot \underline{0} = \underline{0}$; $\beta \cdot \alpha^+ = \beta \cdot \alpha + \beta$; $\beta \cdot \alpha = \bigcup\{(\beta \cdot \gamma)|\gamma < \alpha\}$ if $\alpha \neq \underline{0}$ is a limit ordinal, defines a unique operation on $Ord(x)$. It generalizes the ordinary multiplication of natural numbers.

The Exponentiation of Ordinal Numbers $\beta^{\underline{0}} = \underline{1}$; $\beta^{\alpha^+} = \beta^\alpha \cdot \beta$; $\beta^\alpha = \bigcup\{\beta^\gamma|\gamma < \alpha\}$ if $\alpha \neq \underline{0}$ is a limit ordinal, defines a unique operation on $Ord(x)$. It generalizes the ordinary exponentiation of natural numbers.

For example, $\alpha + \underline{1} = (\alpha + \underline{0})^+ = \alpha^+$, which is in agreement with our earlier convention. Also, $\underline{0}^+ = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \underline{1}$. We get:

$$\begin{array}{lll}
\underline{0}+\underline{1} & = (\underline{0}+\underline{0})^+ = \underline{0}^+ = \underline{1} & \text{i.e., } \underline{0}^+ = \underline{1} \\
\underline{0}+\underline{2} & = (\underline{0}+\underline{1})^+ = (\underline{0}^+)^+ = \underline{1}^+ = \underline{2}, & \text{i.e., } (\underline{0}^+)^+ = \underline{2} \\
\underline{0}+\underline{3} & = (\underline{0}+\underline{2})^+ = ((\underline{0}^+)^+)^+ = \underline{2}^+ = \underline{3}, & \text{i.e., } ((\underline{0}^+)^+)^+ = \underline{3} \\
\underline{0}+\underline{4} & = (\underline{0}+\underline{3})^+ = (((\underline{0}^+)^+)^+)^+ = \underline{3}^+ = \underline{4}, & \text{i.e., } (((\underline{0}^+)^+)^+)^+ = \underline{4}
\end{array}$$

Hence:

$$\underline{2}+\underline{2} = (\underline{0}^+)^+ + (\underline{0}^+)^+ = ((\underline{0}^+)^+ + (\underline{0}^+))^+ = (((\underline{0}^+)^+ + \underline{0})^+)^+ = (((\underline{0}^+)^+)^+)^+ = \underline{4}$$

Kant claimed that a statement like $2 + 2 = 4$ is a *synthetic* judgment, i.e., a fact which must be considered as being *a priori* true. Our argument showed that it can be proven on the basis of our ability to recognize symbols like $(((\underline{0}^+)^+)^+)^+$ as $\underline{3}^+$, i.e., as $\underline{4}$. In this sense, $2+2=4$ admits a proof, i.e., it is the result of an *analytical* judgment.

At no point in our development of ordinals did we use our axiom of infinity, i.e., that there is a set ω of natural numbers $\underline{n}$, $n \in \mathbb{N}$. We are now in a position to adopt a more general approach towards this axiom.

The Finite Ordinals. An ordinal $\alpha > 0$ is called *finite*, if α as well as every ordinal β smaller than α, except $\underline{0}$, is a successor, i.e., it has a *predecessor*. The class of finite ordinals is given by the formula:

$$FOrd(x) = Ord(x) \wedge \forall y\big[Ord(y) \wedge (y \leq x) \wedge (y \neq \emptyset) \rightarrow \exists z\big\{y = z \cup \{z\}\big\}\big]$$

Ordinals which are not finite are called *infinite*. If ν is a finite ordinal then all $\mu \leq \nu$ are finite. If ν is finite then ν^+ is also finite. Clearly, all *standard* finite ordinals, i.e., the ordinals $\underline{n}$ are finite. However, there is no way of showing that all finite ordinals are standard. A definition of finite as being some $\underline{n}$ is a *descriptive* definition in contrast to the *formal* one above. According to a fundamental result of mathematical logic, the so called *Löwenheim-Skolem Theorem*, there is no scheme of formulas $F(x, a), a \in b$, which is satisfied exactly by the standard numbers $\underline{n}$. Finite ordinals ν which are *nonstandard* are of infinite magnitude in the sense that $\nu \geq \underline{n}$ holds for all $n \in \mathbb{N}$. This is very easy to see.

The Axiom of Infinity (revised). There is a set ω whose elements are exactly the finite ordinals.

We see that $\omega = \bigcup\{\nu|\nu \in \omega\}$, hence ω is an ordinal, and ω is not finite because of $\omega \notin \omega$. Hence ω cannot have a predecessor, i.e., ω is a limit ordinal. Because all ordinals smaller than ω are finite, i.e., they are zero or have predecessors, ω *is the smallest infinite limit ordinal.* The existence of an infinite limit ordinal is equivalent to the axiom of infinity.

Theorem 3.33 (Proof by Induction for Finite Ordinals) *Let $P(x)$ be a property. Assume:*

(i) *$P(\underline{0})$ holds.*

(ii) *$P(\nu^+)$ holds, in case that $P(\nu)$ holds.*

Then $P(\nu)$ holds for all $\nu \in \omega$.

PROOF. Indeed, if there is some ordinal λ for which $P(\lambda)$ is not true, then $\lambda > \underline{0}$ for the smallest such ordinal, because of (i), and $\lambda \geq \omega$ because of (ii). □

Addition and multiplication of finite ordinals provide the foundation of a rigorous development of real numbers and the calculus. By the very nature of any such axiomatic or formal approach, *infinitesimals* are lurking in the shadows e.g. as reciprocals of nonstandard finite numbers. However, they cannot be discovered by any formal means. An approach of the Infinitesimal Calculus within any model of ZF by adding an additional predicate which expresses the extraterrestrial quality of being standard has been advocated by **Ed. Nelson** in his book on *Radically Elementary Probability Theory*. It is quite interesting to note that Fraenkel discarded a theory of infinitesimals as useless for serious mathematics. However axiomatic set theory together with formal logic finally reestablished the intuitive methods of Euler and Cauchy, and made them again available not only as an attractive alternative to the $\epsilon-\delta$ approach of the calculus but also as an intuitive and powerful tool for even the most advanced parts of applied mathematics. And most of this *NonStandard Analysis* was developed by **Abraham Robinson** who studied as a freshman set theory and mathematical logic under A. A. Fraenkel.

3.3 The Orderstructure of the Ordinal Sum, the Ordinal Product and Ordinal Exponentiation. Finite Arithmetic

We are going to analyze the well ordered systems $(\alpha + \beta, \in)$, $(\alpha \cdot \beta)$ and β^α in terms of $(\alpha, \in)$ and $(\beta, \in)$. We first need to prove a few facts about the addition and multiplication of ordinals.

$$\alpha + \underline{0} = \alpha \,; \underline{0} + \alpha = \alpha \tag{3.1}$$

The first equation is part of the recursive definition for addition. In order to prove the second equation, we use induction on α. Assume that we have $\underline{0} + \delta = \delta$ for every $\delta < \alpha$. If $\alpha = \sigma^+$ then

$$\underline{0} + \alpha = \underline{0} + \sigma^+ = (\underline{0} + \sigma)^+ = \sigma^+ = \alpha$$

If α is a limit ordinal then

$$\underline{0} + \alpha = \bigcup \{\underline{0} + \delta | \delta < \alpha\} = \bigcup \{\delta | \delta < \alpha\} = \alpha.$$

$$\gamma = \alpha + \beta \text{ is a limit ordinal in case that } \beta \text{ is a limit ordinal} \tag{3.2}$$

Assume $\sigma < \gamma = \bigcup\{\alpha + \delta | \delta < \beta\}$. Then $\sigma \in \alpha + \delta$ for some $\delta < \beta$. But then one has that $\sigma^+ \in (\alpha + \delta)^+ = \alpha + \delta^+ \in \gamma$, because $\delta^+ \in \beta$.
In particular, $1 + \omega$ must be a limit ordinal while $\omega + 1$ is the successor of ω. Actually $1 + \omega = \bigcup \{1 + \nu | \nu \in \omega\} = \omega$. Hence, addition of ordinals is not commutative.

$$\alpha + \beta < \alpha + \gamma \text{ if and only if } \beta < \gamma \tag{3.3}$$

We proceed by induction on γ. That is, we assume that for every ordinal $\delta < \gamma$ one has that $\alpha + \beta < \alpha + \delta$ in case that $\beta < \delta$. If $\gamma = \sigma^+$, then $\beta \leq \sigma$ and

$$\alpha + \beta \leq (\alpha + \sigma) < (\alpha + \sigma)^+ = \alpha + \sigma^+ = \alpha + \gamma.$$

If γ is a limit ordinal, then $\beta^+ < \gamma$ and

$$\alpha + \beta < (\alpha + \beta)^+ = \alpha + \beta^+ \leq \bigcup \{\alpha + \delta | \delta < \gamma\} = \alpha + \gamma$$

On the other hand, if $\alpha + \beta < \alpha + \gamma$ then $\beta \neq \gamma$ and $\beta < \gamma$, by what we have shown.

$$\text{If } \ \alpha + \beta = \alpha + \gamma \ \text{ then } \ \beta = \gamma \tag{3.4}$$

This is trivial, because $\beta \neq \gamma$ leads according to 3.3 to an inequality.

$$\text{If } \ \alpha < \beta \text{ then } \ \alpha + \gamma \leq \beta + \gamma \tag{3.5}$$

We prove this by induction on γ. That is, we assume the claim for all $\delta < \gamma$. If $\gamma = \sigma^+$, then

$$\alpha + \gamma = \alpha + \sigma^+ = (\alpha + \sigma)^+ \leq (\beta + \sigma)^+ = \beta + \sigma^+ = \beta + \gamma$$

If γ is a limit ordinal, then

$$\alpha + \gamma = \bigcup \{\alpha + \delta | \delta < \gamma\} \leq \bigcup \{\beta + \delta | \delta < \gamma\} = \beta + \gamma$$

Note that $\underline{0} < \underline{1}$ but $\underline{0} + \omega = \underline{1} + \omega$. Hence, we have right cancellation but not left cancellation.

$$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) \tag{3.6}$$

We prove this by induction on γ. That is, we assume that associativity holds for every $\delta < \gamma$. If $\gamma = \sigma^+$ then ,

$$(\alpha + \beta) + \sigma^+ = ((\alpha + \beta) + \sigma))^+ = (\alpha + (\beta + \sigma))^+ = \alpha + (\beta + \sigma)^+ = \alpha + (\beta + \sigma^+).$$

If γ is a limit ordinal, then

$$(\alpha + \beta) + \gamma = \bigcup \{(\alpha + \beta) + \delta | \delta < \gamma\} = \bigcup \{\alpha + (\beta + \delta) | \delta < \gamma\}$$

For $\rho = \beta + \delta$, where $\delta < \gamma$, we have that $\rho = \beta + \delta < \beta + \gamma$. Hence:

$$\bigcup \{(\alpha + (\beta + \delta) | \delta < \gamma\} \subseteq \bigcup \{\alpha + \rho | \rho < \beta + \gamma\}$$

Now assume that $\rho < \beta + \gamma$. Then $\rho \in \bigcup \{\beta + \delta | \delta < \gamma\}$, i.e., $\rho \in \beta + \delta$ for some $\delta < \gamma$. Hence, $\alpha + \rho \leq \alpha + (\beta + \delta)$. This is:

$$\bigcup \{\alpha + \rho | \rho < \beta + \gamma\} \subseteq \bigcup \{\alpha + (\beta + \delta) | \delta < \gamma\}$$

Thus:

$$\bigcup \{\alpha + (\beta + \delta) | \delta < \gamma\} = \bigcup \{\alpha + \rho | \rho < \beta + \gamma\}$$

We already know that $\beta + \gamma$ is a limit ordinal. Hence:

$$\bigcup \{\alpha + \rho | \rho < \beta + \gamma\} = \alpha + (\beta + \gamma)$$

This proves associativity.

$$\alpha < \beta \ \text{ if and only if } \ \alpha + \delta = \beta \ \text{ for a unique } \ \delta > \underline{0}. \tag{3.7}$$

If $\delta > \underline{0}$, then $= \alpha + \underline{0} < \alpha + \delta = \beta$. We also know that there can be only one δ such that $\alpha + \delta = \beta$. Let $\alpha < \beta$. We induct on β, i.e., we assume that for every γ, $\alpha \leq \gamma < \beta$, one can find some δ such that $\alpha + \delta = \gamma$.
If $\beta = \sigma^+$, then $\alpha \leq \sigma < \beta$ and $\alpha + \delta = \sigma$ for some δ. Hence,

$$\alpha + \delta^+ = (\alpha + \delta)^+ = \sigma^+ = \beta$$

If β is a limit ordinal, then let

$$s = \{\delta | \alpha + \delta = \gamma < \beta\}$$

The set s is because of 3.3 a proper segment of ordinals, and therefore an ordinal, i.e.,$s = \delta_0$. If $\delta \in \delta_0$ then, because β is a limit ordinal, also $\delta^+ \in \delta_0$. That is, δ_0 is a limit ordinal. But then:

$$\alpha + \delta_0 = \bigcup \{\alpha + \delta | \delta \in \delta_0\} = \bigcup \{\gamma | \gamma < \beta\} = \beta$$

Theorem 3.34 (The Order Structure of the Ordinal Sum) *If $\gamma = \alpha + \beta$ then $\gamma = \alpha \cup (\gamma \setminus \alpha)$ decomposes the well-ordered system $w = (\gamma, \in)$ into two disjoint subsets $w_1 = (\alpha, \in)$ and $w_2 = (\gamma \setminus \alpha, \in)$ where each element of w_1 precedes each element of w_2 and where w_1 is order-isomorphic to α and w_2 is order-isomorphic to β. That is, γ is the **order sum** of α and β.*

PROOF. w_1 is actually equal to α, and therefore isomorphic to α; by what we have shown, the map $\delta \mapsto \alpha + \delta$ is an isomorphism between β and w_2. □

In order to analyze the structure of the ordinal product, we need a few facts.

$$\alpha \cdot \underline{0} = \underline{0} \cdot \alpha = \underline{0}\ ;\ \alpha \cdot \underline{1} = \underline{1} \cdot \alpha = \alpha \tag{3.8}$$

We have $\alpha \cdot \underline{0} = \underline{0}$ by definition and $\alpha \cdot \underline{1} = \alpha \cdot \underline{0}^+ = \alpha \cdot \underline{0} + \alpha = \alpha$.
We prove $\underline{0} \cdot \alpha = \underline{0}$ by induction on α. Hence we assume that for every $\delta < \alpha$ one has that $\underline{0} \cdot \alpha = \underline{0}$ holds. If $\alpha = \sigma^+$, then

$$\underline{0} \cdot \alpha = \underline{0} \cdot \sigma^+ = \underline{0} \cdot \sigma + \underline{0} = \underline{0} + \underline{0} = \underline{0}$$

If α is a limit ordinal, then

$$\underline{0} \cdot \alpha = \bigcup \{\underline{0} \cdot \delta | \delta < \alpha\} = \bigcup \underline{0} = \underline{0}$$

We prove $\underline{1} \cdot \alpha = \alpha$ by induction on α. Hence we assume that for every $\delta < \alpha$ one has that $\underline{1} \cdot \alpha = \delta$ holds. If $\alpha = \sigma^+$, then

$$\underline{1} \cdot \alpha = \underline{1} \cdot \sigma^+ = \underline{1} \cdot \sigma + \underline{1} = \sigma + \underline{1} = \sigma^+ = \alpha$$

If α is a limit ordinal, then

$$\underline{1} \cdot \alpha = \bigcup \{\underline{1} \cdot \delta | \delta < \alpha\} = \bigcup \{\delta | \delta < \alpha\} = \alpha$$

$$\gamma = \alpha \cdot \beta \quad \text{is a limit ordinal in case that } \alpha \text{ or } \beta \text{ are limit ordinals.} \tag{3.9}$$

If β is a limit ordinal, then

$$\gamma = \bigcup \{\alpha \cdot \delta | \delta < \beta\}$$

Let $\sigma < \gamma$, i.e., $\sigma \in \alpha \cdot \delta$ for some $\delta < \beta$. We must have $\alpha > \underline{0}$, by 3.8, and therefore, by 3.3,

$$\sigma < \alpha \cdot \delta < \alpha \cdot \delta + \alpha = \alpha \cdot \delta^+ < \alpha \cdot \delta^+ + \alpha = \alpha \cdot \delta^{++};\ \text{ where } \delta^{++} \text{ ¡ } \beta$$

Now, $\sigma < \alpha \cdot \delta$ is the same as $\sigma^+ \leq \alpha \cdot \delta$ and we conclude that

$$\sigma^+ < \alpha \cdot \delta^{++} \leq \gamma, \text{ i.e., } \sigma^+ < \gamma$$

Hence, γ is a limit ordinal in case that β is a limit ordinal. Now assume that α is a limit ordinal. But then it is enough to assume that β is a successor, i.e., $\beta = \sigma^+$. But then:

$$\alpha \cdot \beta = \alpha \cdot \sigma^+ = \alpha \cdot \sigma + \alpha$$

which is a limit ordinal by 3.2.

Hence, $\underline{2} \cdot \omega$ as well as $\omega \cdot \underline{2}$ are limit ordinals. But $\underline{2} \cdot \omega = \bigcup \{\underline{2} \cdot \nu | \nu \in \omega\} = \omega$ while $\omega \cdot \underline{2} = \omega \cdot \underline{1}^+ = \omega \cdot \underline{1} + \omega = \omega + \omega$. We conclude that multiplication of ordinals is not commutative. Also, $\underline{2} \cdot \omega = (\underline{1} + \underline{1}) \cdot \omega \neq \underline{1} \cdot \omega + \underline{1} \cdot \omega = \omega + \omega$, i.e., right distributivity fails.

$$\text{If } \alpha > \underline{0}, \text{ then } \beta \leq \alpha \cdot \beta \tag{3.10}$$

We prove this by induction on β. That is, we assume $\delta \leq \alpha \cdot \delta$ for every $\delta < \beta$. If β is a successor, i.e., $\beta = \sigma^+$, then:

$$\sigma \leq \alpha \cdot \sigma < \alpha \cdot \sigma + \alpha = \alpha \cdot \sigma^+ = \alpha \cdot \beta, \text{ hence } \beta = \sigma^+ \leq \alpha \cdot \sigma^+ = \alpha \cdot \beta$$

If β is a limit ordinal, then

$$\alpha \cdot \beta = \bigcup \{\alpha \cdot \delta | \delta < \beta\} \supseteq \bigcup \{\delta | \delta < \beta\} = \beta, \text{ i.e., } \alpha \cdot \beta \geq \beta.$$

$$\text{If } \alpha > \underline{0} \text{ and if } \beta > \underline{0} \text{ then } \alpha \cdot \beta > \underline{0} \tag{3.11}$$

This is an immediate consequence of 3.10.

$$\text{If } \alpha > \underline{0} \text{ and } \beta < \gamma, \text{ then } \alpha \cdot \beta < \alpha \cdot \gamma \tag{3.12}$$

We induct on γ. That is, we assume for all $\beta < \delta < \gamma$ that $\alpha \cdot \beta < \alpha \cdot \delta$. If $\gamma = \sigma^+$, then $\beta \leq \sigma$ and $\alpha \cdot \beta \leq \alpha \cdot \sigma$. This is obvious for $\beta = \sigma$, and for $\beta < \sigma$ it is our induction hypothesis. Hence,

$$\alpha \cdot \beta < \alpha \cdot \sigma + \alpha = \alpha \cdot \sigma^+ = \alpha \cdot \gamma$$

If γ is a limit ordinal, then

$$\alpha \cdot \gamma = \bigcup \{\alpha \cdot \delta | \delta < \gamma\} \supseteq \alpha \cdot \beta^+ = \alpha \cdot \beta + \alpha > \alpha \cdot \beta, \text{ i.e., } \alpha \cdot \gamma > \alpha \cdot \beta$$

In particular,

$$\text{If } \alpha > \underline{0} \text{ then } \alpha \cdot \beta = \alpha \cdot \gamma \text{ only if } \beta = \gamma$$

$$\text{If } \alpha > \underline{0} \text{ and } \beta > \underline{1}, \text{ then } \alpha < \alpha \cdot \beta \tag{3.13}$$

Note, that $\alpha = \alpha \cdot \underline{1} < \alpha \cdot \beta$ by 3.8 and 3.12.

$$\text{If } \alpha < \beta, \text{ then } \alpha \cdot \gamma \leq \beta \cdot \gamma \tag{3.14}$$

We induct on γ. That is, we assume for every $\delta < \gamma$ that $\alpha \cdot \delta \leq \beta \cdot \delta$. If $\gamma = \sigma^+$ then, by 3.3:

$$\alpha \cdot \gamma = \alpha \cdot \sigma^+ = \alpha \cdot \sigma + \alpha \leq \beta \cdot \sigma + \alpha < \beta \cdot \sigma + \beta = \beta \cdot \sigma^+ = \beta \cdot \gamma$$

If γ is a limit ordinal, then

$$\alpha \cdot \gamma = \bigcup \{\alpha \cdot \delta | \delta < \gamma\} \subseteq \bigcup \{\beta \cdot \delta | \delta < \gamma\} = \beta \cdot \gamma, \text{ i.e., } \alpha \cdot \gamma \leq \beta \cdot \gamma$$

We have already noticed, that $\underline{1} < \underline{2}$ but $\underline{1} \cdot \omega = \omega = \underline{2} \cdot \omega$. Hence, the $\leq$ sign in 3.14 cannot be replaced by $<$.

$$\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma \tag{3.15}$$

We induct on γ. That is, we assume $\alpha \cdot (\beta + \delta) = \alpha \cdot \beta + \alpha \cdot \delta$ for every $\delta < \gamma$. If $\gamma = \sigma^+$ then:

$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= \alpha \cdot (\beta + \sigma^+) = \alpha \cdot (\beta + (\sigma + \underline{1})) = \alpha \cdot ((\beta + \sigma) + \underline{1}) = \alpha \cdot (\beta + \sigma) + \alpha \\ &= (\alpha \cdot \beta + \alpha \cdot \sigma) + \alpha = \alpha \cdot \beta + (\alpha \cdot \sigma + \alpha) = \alpha \cdot \beta + \alpha \cdot \sigma^+ \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \end{aligned}$$

If γ is a limit ordinal, then $\beta + \gamma$ is a limit ordinal. Hence:

$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= \bigcup \{\alpha \cdot \xi | \xi < \beta + \gamma\} = \bigcup \{\alpha \cdot \xi | \beta \leq \xi < \beta + \gamma\} = \bigcup \{\alpha \cdot \xi | \xi = \beta + \delta, \delta < \gamma\} \\ &= \bigcup \{\alpha \cdot \beta + \alpha \cdot \delta | \delta < \gamma\} \end{aligned}$$

We may assume that $\alpha > \underline{0}$, otherwise our claim 3.15 becomes obvious. Hence, for $\rho = \alpha \cdot \delta$ where $\delta < \gamma$ we have by 3.12 that $\rho < \alpha \cdot \gamma$. This yields:

$$\bigcup \{\alpha \cdot \beta + \alpha \cdot \delta | \delta < \gamma\} \subseteq \bigcup \{\alpha \cdot \beta + \rho | \rho < \alpha \cdot \gamma\}$$

Now let $\rho < \alpha \cdot \gamma$. The set $\delta_0 = \{\delta | \alpha \cdot \delta \leq \rho\}$ is obviously a segment and therefore an ordinal. Because $\rho < \alpha \cdot \gamma$ one has that $\delta < \gamma$, for each $\delta \in \delta_0$. We claim that δ_0 is a successor. Otherwise, $\alpha \cdot \delta_0 = \bigcup \{\alpha \cdot \delta | \delta < \delta_0\}$ and, because $\alpha \cdot \delta \leq \rho$ holds for each $\delta \in \delta_0$, one has that $\alpha \cdot \delta_0 \leq \rho$. But this is $\delta_0 \in \delta_0$, which is a contradiction. Hence, $\delta_0 = \epsilon^+$ and $\alpha \cdot \epsilon \leq \rho < \alpha \cdot \gamma$ shows that $\epsilon < \gamma$. But then also $\epsilon^+ < \gamma$, because γ is a limit ordinal. This is, $\alpha \cdot \delta_0 > \rho$ where $\delta_0 < \gamma$. Hence:

$$\bigcup \{\alpha \cdot \beta + \rho | \rho < \alpha \cdot \gamma\} \subseteq \bigcup \{\alpha \cdot \beta + \alpha \cdot \delta | \delta < \gamma\}$$

This is:

$$\bigcup \{\alpha \cdot \beta + \alpha \cdot \delta | \delta < \gamma\} = \bigcup \{\alpha \cdot \beta + \rho | \rho < \alpha \cdot \gamma\}$$

Because $\alpha \cdot \gamma$ is a limit ordinal, one has that

$$\bigcup \{\alpha \cdot \beta + \rho | \rho < \alpha \cdot \gamma\} = \alpha \cdot \beta + \alpha \cdot \gamma.$$

This proves left distributivity. As an application we note that $\alpha \cdot \underline{2} = \alpha \cdot (\underline{1} + \underline{1}) = \alpha \cdot \underline{1} + \alpha \cdot \underline{1} = \alpha + \alpha$. We extract from the proof of 3.15 the following

Lemma 3.35 *Let $\rho < \alpha \cdot \gamma$. Then there is a unique $\epsilon < \gamma$ such that $\alpha \cdot \epsilon \leq \rho < \alpha \cdot \epsilon^+$. If γ is a limit ordinal then one has that $\rho < \alpha \cdot \delta_0$ holds for some $\delta_0 < \gamma$.*

PROOF. The only thing left to show is uniqueness of ϵ. Assume that we have some $\xi > \epsilon$. Then $\xi \geq \epsilon^+$ and therefore $\alpha \cdot \xi \geq \alpha \cdot \epsilon^+ > \rho$. If $\xi < \epsilon$ then $\xi^+ \leq \epsilon$ and $\alpha \cdot \xi^+ \leq \alpha \cdot \epsilon \leq \rho$. That is, only ϵ satisfies $\alpha \cdot \epsilon \leq \rho < \alpha \cdot \epsilon^+$. □

$$\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma \tag{3.16}$$

We induct on γ, i.e., we assume that $\alpha \cdot (\beta \cdot \delta)$ holds for each $\delta < \gamma$. If $\gamma = \sigma^+$, then

$$\alpha \cdot (\beta \cdot \sigma^+) = \alpha \cdot (\beta \cdot (\sigma + \underline{1})) = \alpha \cdot (\beta \cdot \sigma + \beta) = \alpha \cdot (\beta \cdot \sigma) + \alpha \cdot \beta = (\alpha \cdot \beta) \cdot \sigma + \alpha \cdot \beta = (\alpha \cdot \beta) \cdot \sigma^+$$

This is, $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$.
If γ is a limit ordinal, then $\beta \cdot \gamma$ is a limit ordinal as well as $\alpha \cdot (\beta \cdot \gamma)$. Hence:

$$\alpha \cdot (\beta \cdot \gamma) = \bigcup \{\alpha \cdot \xi | \xi < \beta \cdot \gamma\}$$

For proving our claim 3.16 we may assume $\beta > \underline{0}$. Now, if $\xi = \beta \cdot \delta$ where $\delta < \gamma$, then, according to 3.12, one has that $\xi < \beta \cdot \gamma$. On the other hand, if $\xi < \beta \cdot \gamma$, then according to the last Lemma, and again by 3.12, one has some $\delta < \gamma$ such that $\xi < \beta \cdot \delta < \beta \cdot \gamma$. Therefore,

$$\bigcup \{\alpha \cdot \xi | \xi < \beta \cdot \gamma\} = \bigcup \{\alpha \cdot (\beta \cdot \delta) | \delta < \gamma\}$$

But, $\bigcup \{\alpha \cdot (\beta \cdot \delta) | \delta < \gamma\} = \bigcup \{(\alpha \cdot \beta) \cdot \delta) | \delta < \gamma\} = (\alpha \cdot \beta) \cdot \gamma$. This proves associativity.

Lemma 3.36 *Let $\gamma < \alpha \cdot \beta$. Then $\gamma = \alpha \cdot \rho + \sigma$ for unique $\rho < \beta$ and $\sigma < \alpha$.*

PROOF. According to Lemma 3.35, we can find some $\rho < \beta$ such that $\alpha \cdot \rho \leq \gamma < \alpha \cdot \rho^+$. This is, $\alpha \cdot \rho \leq \gamma < \alpha \cdot \rho + \alpha$. Using 3.7, $\alpha \cdot \rho + \sigma = \gamma$, and $\sigma < \alpha$ by 3.3.
Now note that α and β have to be different from $\underline{0}$ in order for γ to be less than $\alpha \cdot \beta$ Assume $\gamma = \alpha \cdot \rho_1 + \sigma_1 = \alpha \cdot \rho_2 + \sigma_2$ where $\gamma < \alpha \cdot \beta$, $\rho_1, \rho_2 < \beta$ and $\sigma_1, \sigma_2 < \alpha$. We wish to show that $\rho_1 = \rho_2$ and $\sigma_1 = \sigma_2$.
If $\rho_1 = \rho_2$, then $\alpha \cdot \rho_1 = \alpha \cdot \rho_2$ and, a fortiori, $\sigma_1 = \sigma_2$ by 3.4. If $\rho_1 > \rho_2$, then $\rho_1 \geq (\rho_2 + \underline{1})$. Hence, $\alpha \cdot \rho_1 \geq \alpha \cdot (\rho_2 + \underline{1}) = \alpha \cdot \rho_2 + \alpha$. By 3.7 we can find some $\delta \geq \underline{0}$ such that $\alpha \cdot \rho_1 = (\alpha \cdot \rho_2 + \alpha) + \delta$. We conclude, $\alpha \cdot \rho_2 + \sigma_2 = \alpha \cdot \rho_1 + \sigma_1 = \alpha \cdot \rho_2 + \alpha + \delta + \sigma_1$, i.e., $\sigma_2 = \alpha + \delta + \sigma_1 \geq \alpha$, which is a contradiction. □

Theorem 3.37 *The map $(\sigma, \rho) \mapsto \alpha \cdot \rho + \sigma$, $\sigma \in \alpha$ and $\rho \in \beta$, defines a bijection between $\alpha \times \beta$ and $\alpha \cdot \beta$. One has $\gamma_1 = \alpha \cdot \rho_1 + \sigma_1 > \gamma_2 = \alpha \cdot \rho_2 + \sigma_2$ if, and only if, either $\rho_1 > \rho_2$ or in case that $\rho_1 = \rho_2$ that $\sigma_1 > \sigma_2$. Hence, $(\alpha \cdot \beta, \in)$ is order isomorphic to $(\alpha \times \beta, <_{al})$ where $<_{al}$ denotes the* ***anti-lexicographic*** *order on $\alpha \times \beta$.*

PROOF. Note that if $\rho < \beta$ then $\rho^+ \leq \beta$. Thus, $\alpha \cdot \rho^+ = \alpha \cdot \rho + \alpha \leq \alpha \cdot \beta$. For any $\sigma < \alpha$, , we get $\alpha \cdot \rho + \sigma < \alpha \cdot \rho + \alpha \leq \alpha \cdot \beta$, i.e., $\alpha \cdot \rho + \sigma < \alpha \cdot \beta$. Together with Lemma 3.36 we have established the said bijection between the Cartesian product $\alpha \times \beta$ and the ordinal $\alpha \cdot \beta$.
We pull back the order on $\alpha \cdot \beta$ in order to make $\alpha \times \beta$ to a well-ordered system. That is,

$$(\sigma_1, \rho_1) > (\sigma_2, \rho_2) \text{ iff } \gamma_1 = \alpha \cdot \rho_1 + \sigma_1 > \gamma_2 = \alpha \cdot \rho_2 + \sigma_2$$

If $\rho_1 > \rho_2$, then $\rho_1 \geq \rho_2^+$ and $\alpha \cdot \rho_1 \geq \alpha \cdot \rho_2^+ = \alpha \cdot \rho_2 + \alpha > \alpha \cdot \rho_2 + \sigma_2$. Thus, $\alpha \cdot \rho_1 + \sigma_1 \geq \alpha \cdot \rho_1 > \alpha \cdot \rho_2 + \sigma_2$. Hence,

$$(\sigma_1, \rho_1) > (\sigma_2, \rho_2) \text{ in case that } \rho_1 > \rho_2$$

If $\rho_1 = \rho_2$ then $\gamma_1 = \alpha \cdot \rho_1 + \sigma_1 > \gamma_2 = \alpha \cdot \rho_2 + \sigma_2$ yields $\sigma_1 > \sigma_2$. Hence,

$$(\sigma_1, \rho_1) > (\sigma_2, \rho_2) \text{ in case that } \rho_1 = \rho_2 \text{ and } \sigma_1 > \sigma_2$$

This proves our theorem. □

$(\alpha \times \beta, <_{al})$ looks like a matrix with $\beta-$many rows of copies of α:

$$\begin{array}{ccccccccccc}
(\underline{0},\underline{0}) & < & (\underline{1},\underline{0}) & < & \ldots & < & (\sigma,\underline{0}) & < & \ldots & & \sigma \in \alpha \\
(\underline{0},\underline{1}) & < & (\underline{1},\underline{1}) & < & \ldots & < & (\sigma,\underline{1}) & < & \ldots & & \sigma \in \alpha \\
\ldots\ldots & . & \ldots\ldots & . & \ldots & . & \ldots\ldots & . & \ldots & & \\
\ldots\ldots & . & \ldots\ldots & . & \ldots & . & \ldots\ldots & . & \ldots & & \\
(\underline{0},\rho) & < & (\underline{1},\rho) & < & \ldots & < & (\sigma,\rho) & < & \ldots & & \sigma \in \alpha, \rho \in \beta \\
\ldots\ldots & . & \ldots\ldots & . & \ldots & . & \ldots\ldots & . & \ldots & & \\
\ldots\ldots & . & \ldots\ldots & . & \ldots & . & \ldots\ldots & . & \ldots & &
\end{array}$$

That is, $\alpha \cdot \beta$ can be realized as the order sum of $\beta-$many copies of α. There are $\beta-$many layers of copies of α.

Corollary 3.38 (The Division Algorithm) *Let $\alpha > \underline{0}$. Then every β admits a unique representation $\beta = \alpha \cdot \rho + \sigma$ where $\rho \leq \beta$ and $\sigma < \alpha$.*

PROOF. If $\alpha \geq \underline{1}$ then, using the monotinicity relation of 3.14, we infer from $\underline{1} \leq \alpha$ that $\underline{1} \cdot \beta \leq \alpha \cdot \beta < \alpha \cdot \beta^+$. Thus, $\beta < \alpha \cdot \beta^+$ and the claim now follows from Theorem 3.37. □

Interesting special cases are obtained by taking $\alpha = 2$ and $\alpha = \omega$:

- For any ordinal β one has that $\beta = \underline{2} \cdot \rho + \sigma$ where σ is either $\underline{0}$ or $\underline{1}$. That is, β is either *even* or *odd*.
- For any ordinal β one has that $\beta = \omega \cdot \rho + \nu$ where $\nu \in \omega$. That is, any ordinal β is the sum of a limit ordinal γ which is equal to $\rho-$many copies of ω pieced together, plus a finite ordinal ν.

Of course, for finite ordinals we can say more about the sum and product. But notice, our proofs are completely formal and don't rely on an intuitive understanding of what a number is and how we learned to add and multiply.

Theorem 3.39 *The set ω of finite ordinals is closed under addition and multiplication. Moreover, addition and multiplication of finite ordinals is commutative. That is, if $\nu \in \omega$ and $\mu \in \omega$ then*

(i) $\nu + \mu \in \omega \;\; ; \;\; \nu \cdot \mu \in \omega$

(ii) $\nu + \mu = \mu + \nu \;\; ; \;\; \nu \cdot \mu = \mu \cdot \nu$

PROOF. In order to prove this, we induct on μ. If $\mu = \underline{0}$, then $\nu + \underline{0} = \nu \in \omega$ by 3.1 . If $\mu = \sigma^+$ then $\nu + \mu = (\nu + \sigma)^+ = \rho^+ \in \omega$, by the induction hypothesis. Also, $\nu \cdot \underline{0} = \underline{0} \in \omega$ by 3.8. If $\mu = \sigma^+$ then $\nu \cdot \mu = \nu \cdot \sigma^+ = \nu \cdot \sigma + \nu \in \sigma$ by the first part of (i).

We first establish commutativity of addition for $\mu = \underline{0}$ and $\mu = \underline{1}$.
We have that $\nu + \underline{0} = \underline{0} + \nu = \nu$ by 3.1.
If $\mu = \underline{1}$, then $\nu + \underline{1} = \nu^+ = \underline{1} + \nu$. We prove this by induction on ν. For $\nu = \underline{0}$ this is $\underline{0} + \underline{1} = \underline{0}^+ = \underline{1} + \underline{0}$ which is true by 3.1 and because of $\underline{0}^+ = \underline{1}$. If $\nu = \sigma^+$ we have

$$\nu + \underline{1} = \sigma^+ + \underline{1} = (\sigma + \underline{1}) + \underline{1} = (\underline{1} + \sigma) + \underline{1} = \underline{1} + (\sigma + 1) = \underline{1} + \sigma^+ = \underline{1} + \nu$$

by induction and 3.6.
If $\mu = \sigma^+$ we conclude, by induction, 3.6, and $\underline{1} + \nu = \nu + \underline{1}$:

$$\nu + \sigma^+ = (\nu + \sigma)^+ = (\sigma + \nu)^+ = (\sigma + \nu) + \underline{1} = \sigma + (\nu + \underline{1}) = \sigma + (\underline{1} + \nu) = (\sigma + \underline{1}) + \underline{1} = \sigma^+ + \nu$$

Hence, $\nu + \mu = \mu + \nu$.

In order to prove commutativity for multiplication, we need

$$(*)\ (\nu + \underline{1}) \cdot \mu = \nu \cdot \mu + \mu$$

We prove this by induction on μ. For $\mu = \underline{0}$ this is $(\nu + \underline{1}) \cdot \underline{0} = \nu \cdot \underline{0} + \underline{0}$, which is true because both sides are zero by 3.8 and 3.1. If $\mu = \sigma^+$ then

$$(\nu + \underline{1}) \cdot \sigma^+ = (\nu + \underline{1}) \cdot \sigma + (\nu + \underline{1}) = (\nu \cdot \sigma + \sigma) + (\nu + \underline{1}) = (\nu \cdot \sigma + \nu) + (\sigma + \underline{1}) = \nu \cdot \sigma^+ + \sigma^+$$

which is (*) for μ.

If $\mu = \underline{0}$, then $\nu \cdot \underline{0} = \underline{0} \cdot \nu = \underline{0}$ by 3.8. If $\mu = \sigma^+$ then

$$\nu \cdot \sigma^+ = \nu \cdot \sigma + \nu = \sigma \cdot \nu + \nu$$

by the definition of ordinal multiplication and by the induction hypothesis. But also,

$$\sigma^+ \cdot \nu = \sigma \cdot \nu + \nu$$

by (*). This proves $\nu \cdot \mu = \mu \cdot \nu$. □

Concerning exponentiation of ordinals, we must note that the notation $\gamma = \alpha^\beta$ is ambiguous. The ordinals α and β are sets but γ is **not** the set of maps from β into α. For example:

$$\underline{2}^\omega = \bigcup \{\underline{2}^\nu | \nu \in \omega\} = \omega$$

while the set of maps from ω into $\underline{2}$ is equivalent to $\mathbb{R}$, i.e., to the set of real numbers. Monk uses $\alpha^{\bullet\beta}$ for ordinal exponentiation. We do not adopt his convention. Later on, it will become clear from the context what kind of exponentiation is meant. We already remark, that there is no natural way to make $\underline{2}^\omega$, perceived as the set of maps from ω into $\underline{2}$ to a well ordered set. One needs the AC to well-order this set, i.e., $\mathbb{R}$. However, one has the following

Proposition 3.40 *A **finite** direct product*

$$a = a_{\underline{0}} \times \ldots \times a_{\nu - \underline{1}}$$

of well-ordered sets a_μ is well-ordered by the lexicographic order:

$$b = (b_{\underline{0}}, \ldots, b_{\nu - \underline{1}}) < c = (c_{\underline{0}}, \ldots, c_{\nu - \underline{1}}) \text{ iff } b_\mu < c_\mu \text{ if } \mu \text{ is the first index where } a \text{ and } b \text{ differ.}$$

Proof. This relation $<$ is obviously irreflexive and one can easily prove that it is also transitive. We need to show that every subset s of a has a smallest element. We define successively subsets s_μ of s and elements $c_\mu \in a_\mu$:

$$s_{\underline{0}} = \{b_{\underline{0}} | b_{\underline{0}} \text{ first component of some } b \in s\}$$

$$c_0 = \min(s_{\underline{0}})$$

$$s_{\underline{1}} = \{b_{\underline{1}} | b_{\underline{1}} \text{ second component of some } b \in s \text{ where the first component is } c_{\underline{0}}\}$$

$$c_{\underline{1}} = \min(s_{\underline{1}})$$

$$\ldots$$

$$s_{\nu - \underline{1}} = \{b_{\nu - \underline{1}} | b_{\nu - \underline{1}} \text{ last component of some } b \in s \text{ where first components are } c_{\underline{0}}, c_{\underline{1}}, \ldots, c_{\nu - \underline{2}}\}$$

$$c_{\nu - \underline{1}} = \min(s_{\nu - \underline{1}})$$

It is easy to see that $c = (c_{\underline{0}}, \cdots, c_{\nu - \underline{1}})$ is the minimum of s. □

Definition 3.7 We define for ordinals α and β the *weak cartesian power*:

$$\alpha^{(\beta)} = \{f | f(\gamma) > 0 \text{ for only finitely many } \gamma \in \beta\}$$

If f and g are different elements of $\alpha^{(\beta)}$ then, because f and g differ at only finitely many γ's there is a largest γ where $f(\gamma) \neq g(\gamma)$. We define $f < g$ in case that $f(\gamma) < g(\gamma)$.

For example, we have in $\underline{2}^{(\omega)}$ that:

$$(\underline{0}, \underline{0}, \ldots) < (\underline{0}, \underline{1} \ldots) < (\underline{0}, \underline{0}, \underline{1}, \ldots) < \ldots$$

Theorem 3.41 *The relation $<$ is a well-order on $\alpha^{(\beta)}$*

Proof. It is obviously irreflexive and transitive. We need to show that every nonempty subset s of $\alpha^{(\beta)}$ has a minimum. For f in $\alpha^{(\beta)}$ we define $\mu_0(f) = \max\{\lambda | f(\lambda) > \underline{0}\}$. We can do this because the functions in $\alpha^{(\beta)}$ are nonzero for only finitely many ordinals in β. Then let

$$\lambda_0 = \min\{\mu_0(f) | f \in s\},\ \xi_0 = \min\{f(\lambda_0) | f \in s \text{ where } \mu_0(f) = \lambda_0\},\ s_0 = \{f | f(\lambda_0) = \xi_0,\ \mu_0(f) = \lambda_0\}$$

We have $g > f$ for $g \notin s_0$ and $f \in s_0$. Thus, only functions $f \in s_0$ are candidates for $\min(s)$. If s_0 is a singleton, we are done. Otherwise, we continue and define, similarly as before, $\mu_1(f) = \max\{\lambda | f(\lambda) > \underline{0}, \lambda < \lambda_0\}$. Then let

$$\lambda_1 = \min\{\mu_1(f) | f \in s_0\},\ \xi_1 = \min\{f(\lambda_1) | f \in s_0 \text{ where } \mu_1(f) = \lambda_1\},\ s_1 = \{f | f(\lambda_1) = \xi_1,\ \mu_1(f) = \lambda_1\}$$

We have $\lambda_0 > \lambda_1$ and $g > f$ for $g \in s_0 \setminus s_1$ and $f \in s_1$. We can continue that way and create a strictly decreasing sequence of ordinals $\lambda_0 > \lambda_1 \ldots$ which must be finite[1]. Thus s_μ is a singleton for a finite ordinal μ. The element of s_μ then is the minimum for s. □

The following relations are quite obvious from the definitions:

$$\alpha^{\underline{0}} = \{\emptyset\} = \underline{1};\ \alpha^{(\beta+\underline{1})} \cong \alpha^{(\beta)} \times \alpha;\ \alpha^{(\gamma)} = \bigcup \left\{\alpha^{(\beta)} | \beta < \gamma\right\},\ \text{if } \gamma \text{ is a limit ordinal} \tag{3.17}$$

These identities make the following theorem quite obvious:

Theorem 3.42 *The order type of the exponentiation of ordinals is order isomorphic to the weak cartesian power, i.e.,*

$$\alpha^\beta \cong \alpha^{(\beta)}$$

as well-ordered sets. □

One can use this isomorphism to prove various exponential identities, e.g., $\alpha^{\beta+\gamma} = \alpha^\beta \cdot \alpha^\gamma$ and $\alpha^{\beta \cdot \gamma} = (\alpha^\beta)^\gamma$. The corresponding identities for the weak cartesian power, i.e., $\alpha^{(\beta+\gamma)} \cong \alpha^{(\beta)} \times \alpha^{(\gamma)}$ and $\alpha^{(\beta \cdot \gamma)} \cong \left(\alpha^{(\beta)}\right)^{(\gamma)}$ are quite easily established.

[1] A formalization of this process is based on "Definition by Recursion on Ordinals"(Theorem 3.32), and is left as an excercise. For example, λ_1 should be $\lambda_{\underline{1}}$ etc.

Chapter 4

The Axiom of Choice

The axiom of choice (AC) states that every family of non-empty sets admits a choice function. As an important consequence we note that every set p of mutually disjoint and non-empty sets c admits a *cross-section*, i.e., a set s such that $s \cap c \neq \emptyset$ for each $c \in p$. For s we may take the range of a choice function for the family which is the identity on p, i.e., $p \to p, c \mapsto c$. The sets c are called the *classes* of the *partition* p of the set $a = \bigcup c$, and s is also called a set of *representatives* for the partition p. Because the sets c are pairwise disjoint, $s \cap c$ is a singleton for each $c \in p$. The statement that any partition p of a set a into non-empty classes admits a set of representatives is actually equivalent to AC. Given any family $a, j \in i$, of non-empty sets we can make it to a set p of disjoint sets by defining $p = \{c_j = \{j\} \times a_j | j \in i\}$. Clearly, if $j \neq j'$ then $c_j \cap c_{j'} = \emptyset$. Then if s is a cross-section for p, we have $s \cap c_j = (j, s_j)$ for some $s_j \in a_j$, and the map $j \mapsto s_j$ is a choice function for the family $a_j, j \in i$. A choice function is easy to find in case that all sets a_j are well-ordered. One may pick from every a_j just the smallest element. Notice that a well-ordering of $a = \bigcup \{a_j | j \in i\}$ induces a well-ordering of each a_j and produces a choice function. The most substantial consequence of AC is

Theorem 4.1 (Zermelo's Well-Ordering Principle (WO)) *Every set can be well-ordered.*

Proof. Let f be any choice function on $\mathcal{P}(a) \setminus \{\emptyset\}$. The idea to well-order a is quite simple. Define $a_0 = f(a)$, $a_1 = f(a \setminus \{a_0\}), a_2 = f(a \setminus \{a_0, a_1\}), \ldots$ and set $a_0 < a_1 < \ldots$ The process of assigning ordinals to elements of a should terminate because at every such step we are taking away an element from a. The formal definition of this process is done by recursion. We pick some $c \notin a$ and define a functional relation by

$$H(x,y) = \begin{cases} f(a \setminus ran(g)) & \text{if } x = a \text{ and } y = g : \alpha \to a, ran(g) \subset a \\ \\ c & \text{otherwise} \end{cases}$$

Because the universe $\mathcal{U}$ is not a set it is clear that such a set c must exist for a. By the Recursion Theorem we have a functional relation $F(x, y)$ such that for each ordinal α, $F(\alpha) = H(\alpha, F \restriction \alpha)$ holds. If $F(\alpha) \neq c$ holds, then $F \restriction \alpha$ is a function whose domain is α, the range is $\{F(\beta) | \beta < \alpha\} \subset a$ and, moreover, $F(\alpha) = f(a \setminus \{F(\beta) | \beta < \alpha\})$. We conclude that $F(\alpha) \neq F(\beta)$ for all $\beta < \alpha$. If we had $F(\alpha) \neq c$ for all ordinals α, then $F(x)$ would establish an injective functional relation from $Ord(x)$ into the set a. It would follow from comprehension and replacement that $Ord(x)$ is a set. Hence, there is a smallest ordinal γ such that $F(\gamma) = c$. For each $\alpha < \gamma$ we have that $F(\alpha) \in a$. Hence $F \restriction \gamma : \gamma \to a$ and the map is injective. But $rang(F \restriction \gamma) = a$, because $F(\gamma) = c$. Thus $F \restriction \gamma$ is bijective and well-orders a. □

Let $(p, \leq)$ be a partially ordered set. An element $m \in p$ is called *maximal* if $m \leq a$ implies that $m = a$. The open interval $(0, 1)$ of real numbers is an example of a partially ordered set which does not have maximal elements. The maximal elements of the partially ordered set of all proper subspaces of $\mathbb{R}^n$ are the hyperplanes. In a totally unordered set, any element is maximal. In order to find a condition which guarantees the existence of maximal elements let us try to find one in $(p, \leq)$. Pick any element a_0 in p. If a_0 is not maximal, pick an element $a_1 > a_0 \ldots$ etc. If the so created sequence does not terminate at some finite $\underline{n} \in \omega$, we have to define some $a_\omega > a_{\underline{n}}$, i.e., we need to have at least one upper bound for the chain $\{a_{\underline{n}} | \underline{n} \in \omega\}$. This heuristic construction suggests

Theorem 4.2 (Zorn's Lemma (ZL)) *Let $(p, \leq)$ be a partially ordered set where every chain has an upper bound in p. Then p contains a maximal element.*

PROOF. The element $u \in p$ is called a strict upper bound for the subset a of p, if $u > b$ holds for each $b \in a$. Any element of p is a strict upper bound for $\emptyset$. There is no strict upper bound for p. Let d be the set of all subsets of p which have a strict upper bound and for $a \in d$ let $u(a)$ be the non-empty set of strict upper bounds. For a choice function f on $\mathcal{P}(p) \setminus \{\emptyset\}$ and $a \in d$ define $v(a) = f(u(a))$, i.e., v selects a strict upper bound for a subset a of p, if there is one. Of course, $v(a) \notin a$ for each $a \in d$. We define

$$H(x, y) = \begin{cases} v(rang(g)) & \text{if } x = \alpha \text{ and } y = g : \alpha \to p, ran(g) \in d \\ \\ c & \text{otherwise} \end{cases}$$

As before, we choose for c a set which is not a member of p. According to the Recursion Principle we have a functional relation $F(x, y)$ such that $F(\alpha) = H(\alpha, F \restriction \alpha)$. If $F(\alpha) \neq c$, then $F \restriction \alpha$ is a function from α into p, where the range $\{F(\beta) | \beta < \alpha\}$ belongs to d, and $F(\alpha) = v(\{F(\beta) | \beta < \alpha\})$. Therefore, $F(\alpha) \neq F(\beta), \beta < \alpha$. If we had $F(\alpha) \neq c$ for all ordinals then $F(x, y)$ would define an injective functional relation of the class $Ord(x)$ into the set p. Hence $F(\gamma) = c$ for a smallest ordinal γ. But then for $\beta < \alpha < \gamma$ we have, because of $F(\alpha) = v(\{F(\beta) | \beta < \alpha\})$, that $F(\alpha) > F(\beta)$. Hence $\{F(\beta) | \beta < \gamma\}$ is a chain in p and must have an upper bound m. On the other hand, this chain cannot have a strict upper bound because $F(\gamma) = c$. Hence $\gamma = \alpha^+$ and $F(\alpha) = m$ and there cannot be any element larger than m. This is, m is maximal. □

Corollary 4.3 *Let $(p, \leq)$ be a partially ordered set where every chain has an upper bound. Then for any $a \in p$ there is some maximal element m in p such that $m \geq a$.*

PROOF. We only have to apply Zorn's Lemma to the partially ordered set $\{b | b \geq a\}$. □

Corollary 4.4 *Let $(p, \leq)$ be a partially ordered set where every well-ordered chain has an upper bound. Then p has a maximal element.*

PROOF. For the proof of ZL we only needed an upper bound for the well-ordered chain $\{F(\alpha) | \alpha < \gamma\}$. □

Theorem 4.5 *(ZL) implies within ZF Set Theory (WO) and, a fortiori, also (AC).*

PROOF. Let a be a set we wish to well-order. To this end let s be the set of all well-orderings of subsets b of a, i.e., $s = \{r | r \subseteq b \times b, b \subseteq a, r \text{ well-order on } b\}$. We partially order s by saying that $r_1 \leq r_2$ if b_1 is a segment of b_2 . Let c be any chain in s. It is quite obvious that $u = \bigcup \{r | r \in c\}$ is a well-order on the set $b = \bigcup \{dom(r) | r \in c\}$, i.e., $u \in s$, and $u \geq r$ for each $r \in c$. Hence any chain in s has an upper bound. By ZL we conclude that s has a maximal element w. If we had $b = dom(w) \subset a$, then any element $d \in a \setminus b$ put on top of (b, w) would violate the maximality of w.
That WO implies AC is a trivial observation which we have stated before. □

Chapter 5

The Axiom of Foundation

The axiom of foundation says that every non-empty set a contains an element b such that $b \cap a = \emptyset$. It follows that a set can never be an element of itself, $\neg(x \in x)$, or more generally, there are no finite $\in -cycles$, i.e., maps defined on a finite ordinal ν such that $a_{\underline{0}} = a_{\nu - \underline{1}}$, $a_\mu \in a_{\mu + \underline{1}}$, $\underline{0} \leq \mu < \nu - \underline{1}$; and there is no map on ω such that $a_{\nu + \underline{1}} \in a_\nu$. The range of such maps would violate AF. An element $b \in a$ such that $b \cap a = \emptyset$, is *minimal* in $(a, \in)$, i.e., $b \in a$ and for no $c \in a$ one has that $c \in b$. From this remark we conclude

Proposition 5.1 (AF) *A set α is an ordinal if and only if*

(i) *α is transitive.*

(ii)' *$(\alpha, \in)$ is totally ordered.*

Proof. Recall that an ordinal is a set α which is transitive on which $\in$ is a strict well-ordering. Because of AF, $\in$ is irreflexive. If we have that $\in$ is a total order then a minimal element of α is actually the minimum. □

If a is transitive and $b \in a$ then $b \subseteq a$; thus $b \subset a$, in the presence of AF.

Theorem 5.2 (AF) *A set α is an ordinal if and only if*

(i)' *α and all of its elements β are transitive.*

Proof. We need to show (ii)'. To this end assume that there are $\beta, \gamma \in \alpha$ which are different and incomparable with respect to $\in$. Hence the set

$$a = \big\{\gamma | \gamma \in \alpha, \exists_{\beta \in \alpha} \big[(\beta \neq \gamma) \wedge \neg(\beta \in \gamma) \wedge \neg(\gamma \in \beta)\big]\big\}$$

is non-empty and has a minimal element γ_0. Because α is transitive we have $\gamma_0 \subset \alpha$. Hence every element $\delta \in \gamma_0$ belongs to α but does not meet the condition for a. That is:

(*) If $\delta \in \gamma_0$ then we have for any $\beta \in \alpha$ that $\beta = \delta$ or $\beta \in \delta$ or $\delta \in \beta$.

On the other hand, γ_0 belongs to a, hence the set

$$b = \big\{\beta | \beta \in \alpha, (\beta \neq \gamma_0) \wedge \neg(\beta \in \gamma_0) \wedge \neg(\gamma_0 \in \beta)\big\}$$

is non-empty and has a minimal element β_0. Because α is transitive, every element $\rho \in \beta_0$ belongs to α but does not meet the condition for b. That is

(**) If $\rho \in \beta_0$ then we have $\rho = \gamma_0$ or $\rho \in \gamma_0$ or $\gamma_0 \in \rho$.

We are going to deduce a contradiction from (*) and (**). First, we derive $\beta_0 \subseteq \gamma_0$. Let $\rho \in \beta_0$; we use (**) to conclude $\rho \in \gamma_0$. If we had $\rho = \gamma_0$, then $\gamma_0 \in \beta_0$ would contradict $\beta_0 \in b$. If we had $\gamma_0 \in \rho$ then $\rho \in \beta_0$ together with the transitivity of $\beta_0 \in \alpha$ would yield $\gamma_0 \in \beta_0$, contradicting again $\beta_0 \in b$. Now, $\beta_0 \subseteq \gamma_0$, and $\beta_0 = \gamma_0$ cannot hold because of $\beta_0 \in b$. This is $\beta_0 \subset \gamma_0$ and we may pick some $\delta \in \gamma_0 \setminus \beta_0$. By (*) we have $\beta_0 = \delta$ or $\beta_0 \in \delta$ or $\delta \in \beta_0$ where we have already excluded $\delta \in \beta_0$. Now, $\beta_0 = \delta$ yields $\beta_0 \in \gamma_0$ which contradicts $\beta_0 \in b$; $\beta_0 \in \delta$ and $\delta \in \gamma_0$ yields by transitivity of the set $\gamma_0 \in \alpha$ again the contradiction $\beta_0 \in \gamma_0$. □

We have already introduced the cumulative ZF-Hierarchy of Sets:

$$V_{\underline{0}} = \emptyset \,; V_{\alpha^+} = \mathcal{P}(V_\alpha) \,; V_\alpha = \bigcup \big\{ V_\beta | \beta < \alpha \big\}, \text{ if } \alpha \text{ is a limit ordinal.}$$

and the class $V(x) = \exists z \big[Ord(z) \wedge x \in V_z \big]$ of sets which belong to that hierarchy. We are going to show that under the assumption of AF, every set belongs to that hierarchy. Actually, AF is equivalent to this statement. Not making any use of AF, we are going to prove a few useful facts about $V(x)$. But first a rather trivial

Lemma 5.3 *Assume that the set a is transitive. Then $\mathcal{P}(a)$ is transitive. If $a \notin a$ then $\mathcal{P}(a) \notin \mathcal{P}(a)$.*

Proof. Let a be transitive and assume $b \in \mathcal{P}(a)$; that is $b \subseteq a$. But then, if $c \in b$ one has that $c \in a$. Because a is transitive, we have that $c \subseteq a$. Hence, $c \in \mathcal{P}(a)$. That is, $b \subseteq \mathcal{P}(a)$.
Now assume that we have $a \notin a$ but $\mathcal{P}(a) \in \mathcal{P}(a)$, that is, $\mathcal{P}(a) \subseteq a$. But then $a \in \mathcal{P}(a) \subseteq a$, hence $a \in a$, which is a contradiction. □

Lemma 5.4 *Each V_α is transitive and one has that $V_\alpha \notin V_\alpha$. If $\alpha < \beta$ then $V_\alpha \subset V_\beta$ and $V_\alpha \in V_\beta$*

Proof. Assume for all $\alpha < \gamma$ that V_α is transitive.
If γ is a limit ordinal then we have $V_\gamma = \bigcup \big\{ V_\alpha | \alpha < \gamma \big\}$. Assume that $a \in V_\gamma$. Then $a \in V_\alpha$ for some $\alpha < \gamma$. But by transitivity of V_α we have $a \subseteq V_\alpha$ and therefore $a \subseteq V_\gamma$.
If γ is a successor, then $\gamma = \alpha^+$ where $V_\gamma = \mathcal{P}(v_\alpha)$ and the claim follows from Lemma 5.3.

Assume $V_\alpha \notin V_\alpha$ for all $\alpha < \gamma$.
Assume $V_\gamma \in V_\gamma$. If γ is a limit ordinal then $V_\gamma \in V_\alpha$ for some $\alpha < \gamma$. But V_α is transitive, so $V_\gamma \subseteq V_\alpha$. But $V_\alpha \in V_{\alpha^+} \subseteq V_\gamma \subseteq V_\alpha$. Hence, $V_\alpha \in V_\alpha$, a contradiction.
If γ is a successor, then $\gamma = \alpha^+$ and the claim follows from Lemma 5.3.

Assume for all $\alpha < \beta < \gamma$ that $V_\alpha \subset V_\beta$ and $V_\alpha \in V_\beta$.
If γ is a limit ordinal then $V_\beta \subseteq V_\gamma$ and $V_\beta \in V_{\beta^+} \subseteq V_\gamma$. This is $V_\beta \in V_\gamma$. We have $V_\beta \subset V_{\beta^+} \subseteq V_\gamma$, thus $V_\beta \subset V_\gamma$.
If γ is a successor, then $V_\gamma = V_{\beta^+}$. We have $V_\beta \in V_{\beta^+} = V_\gamma$, i.e., $V_\beta \in V_\gamma$. Because V_γ is transitive, we have already $V_\beta \subseteq V_\gamma$. If we had $V_\beta = V_\gamma$, then $V_\gamma = V_\beta \in V_\gamma$, but we have already $V_\gamma \in V_\gamma$ ruled out.
□

Assume that $V(a)$ holds for the set a. Then let γ be the smallest ordinal such that $a \in V_\gamma$ holds. Because $V_{\underline{0}} = \emptyset$, we have $\gamma > \underline{0}$. It is also clear that γ must be a successor ordinal. For a limit ordinal γ, each element a of V_γ belongs to some V_α for some $\alpha < \gamma$. Thus $\gamma = \rho^+$ and we call $\rho(a)$ the *rank* of a: $\rho(a) = \alpha$ if $a \in V_\gamma$ where $\gamma = \alpha^+$ is the smallest ordinal γ such that $a \in V_\gamma$, or $a \subseteq V_\alpha$

Assume $a \in V_\beta$. If β is a limit ordinal then $\rho(a) < \beta$. If $\beta = \alpha^+$ one has that $\rho(a) \leq \alpha < \beta$. Thus, $a \in V_\beta$ iff $\rho(a) < \beta$.

Proposition 5.5 *$V(a)$ holds if and only if $V(b)$ holds for every $b \in a$. Moreover, $\rho(b) < \rho(a)$ for every $b \in a$. If c is a subset of a then $V(c)$ holds and $\rho(c) \leq \rho(a)$.*

PROOF. Let $\rho(a) = \alpha$. Then $a \in V_{\alpha^+} = \mathcal{P}(V_\alpha)$, i.e., $a \subseteq V_\alpha$. If $b \in a$ then $b \in V_\alpha$. Hence, $V(b)$ and $\rho(b) < \alpha = \rho(a)$.
Now assume $V(b)$ for every $b \in a$. Let $\alpha = \bigcup \{\rho(b)^+ | b \in a\}$. Then $b \in V_{\rho(b)^+} \subseteq V_\alpha$ for every $b \in a$. Hence, $b \in V_\alpha$ for every $b \in a$ which is $a \subseteq V_\alpha$, i.e., $a \in V_{\alpha^+}$.
We have $c \subseteq a \subseteq V_\alpha$, hence $c \subseteq V_\alpha$ which is $\rho(c) \leq \alpha = \rho(a)$ □

Proposition 5.6 *$V(\alpha)$ holds for every ordinal α and one has the $\rho(\alpha) = \alpha$.*

PROOF. By Proposition 5.5 we have that $V(\alpha)$ holds in case that $V(\beta)$ holds for all $\beta < \alpha$. So $V(\alpha)$ holds for every ordinal α.
If we have $\beta \in V_{\beta^+}$ for each $\beta \in \alpha$, then $\beta \subseteq V_\beta$ and because of $V_\alpha = \bigcup \{\mathcal{P}(V_\beta) | \beta < \alpha\}$, one has $\beta \in V_\alpha$ for each $\beta \in \alpha$. Hence $\alpha \subseteq V_\alpha$, i.e., $\alpha \in V_{\alpha^+}$. This is $\alpha \in V_{\alpha^+}$ for each ordinal α. Assume that there is an ordinal α where $\rho(\alpha) < \alpha$, i.e., an ordinal for which $\alpha \in V_\alpha$. We pick the smallest such α. Then $\alpha \in V_\alpha = \bigcup \{\mathcal{P}(V_\beta) | \beta < \alpha\}$ yields some $\beta < \alpha$ such that $\beta \in \alpha \subseteq V_\beta$, i.e., $\beta \in V_\beta$ for some $\beta < \alpha$. However, α was the smallest such ordinal. □

Proposition 5.7 *Each set a for which $V(a)$ holds contains an element b such that $a \cap b = \emptyset$*

PROOF. $\{\rho(b) | b \in a\}$ is a set of ordinals and has a smallest element β. Hence $\rho(b) = \beta$ for some $b \in a$ and according to Proposition 5.5, $\rho(c) < rho(b)$ for each $c \in b$. Hence $c \notin a$ for each $c \in b$. □

We are going to show that the axiom of foundation implies that each set a of the universe belongs to the class $V(x)$. A possible argument runs roughly like this: Assume that there is a set a such that $\neg V(a)$ holds. Then a must contain some element a_1 , such that $\neg V(a_1)$. But then a_1 must contain some element a_2 such that $\neg V(a_2)$ etc. That is, we create a sequence $a \ni a_1 \ni a_2 \ni \ldots$ which violates AF. A complete proof along these lines requires obviously AC. We are tempted to define recursively a sequence a_ν , $\nu \in \omega$, such that $a_0 = a$, and $a_{\nu+\underline{1}} = f\big(\{b | b \in a_\nu, \neg V(b)\}\big)$ for some choice function f. The trouble with this definition of the a_ν is that we first have to specify a suitable family of non-empty sets[1] in order to have a choice function for the generation the a_ν. There is a somewhat more elementary approach possible, which is based on the transitive closure of a set.

Lemma 5.8 *For any set a there is a set $tr(a)$ which contains a and which is transitive. Moreover, any transitive set which contains a contains also $tr(a)$.*

PROOF. We define sets $a_{\underline{0}} = a, a_{\underline{1}} = \bigcup \{b | b \in a_{\underline{0}}\}, a_{\underline{2}} = \bigcup \{b | b \in a_{\underline{1}}\}, \ldots a_{\nu+\underline{1}} = \bigcup \{b | b \in a_\nu\}$ and $tr(a) = \bigcup \{a_\nu | \nu \in \omega\}$.
We have $tr(a) \supseteq a$ and if $b \in tr(a)$ then $b \in a_\nu$ for some ν. But then $b \subseteq a_{\nu+\underline{1}} \subseteq tr(a)$, i.e., $tr(a)$ is transitive.
Now assume that $c \supseteq a$ where c is transitive. Assume that $a_\nu \subseteq c$. Let $b \in a_\nu \subseteq c$. Then $b \in c$ and $b \subseteq c$, because c is transitive. Hence $a_{\nu+\underline{1}} = \bigcup \{b | b \in a_\nu\} \subseteq c$. □

Theorem 5.9 *The axiom of foundation holds if and only if every set belongs to the ZF-Hierarchy: $AF \equiv \forall x V(x)$*

PROOF. We already showed that every set a which belongs to $V(x)$ has an element b which is disjoint to a. So $\forall x V(x)$ implies the axiom of foundation.

[1] However, **Hilbert** postulated the existence of a universal choice (class-)function ε for which $a \neq \emptyset \to \varepsilon(a) \in a$ holds. He called it the *logical choice function*. However, nobody assumes this strong version of AC anymore

For the converse, assume that we have AF but that there is a set a which does not belong to $V(x)$. Then $tr(a)$ does not satisfy $V(x)$ as a set which contains a. Let $b = \{c|c \subseteq tr(a), \neg V(c)\}$.We are going to show that b violates AF. Let $c \in b$. Then, because of $\neg V(c)$, c must contain an element d such that $\neg V(d)$. Now, $d \in c \subseteq tr(a)$ yields $d \in tr(a)$. But $tr(a)$ is transitive, thus $d \subseteq tr(a)$. Together with $\neg V(d)$ this yields $d \in b$. Hence, $c \cap b \neq \emptyset$. Because this holds for every element $c \in b$, the AF cannot hold. □

Chapter 6

Cardinals

6.1 Equivalence of Sets

Two sets a and b are said to be *equivalent* if there is some bijection from a onto b. This is obviously an equivalence relation whose domain is the class of all sets. We write $a \approx b$ for equivalent sets a and b. Intuitively, the sets a and b are equivalent if they have the same number of elements. That a has not more elements than b can be formalized by defining: $a \leq_{in} b$ iff there is an injection from a to b; or quite similarly: $a \leq_{pr} b$ iff there is a surjection from b onto a. Both relations are quasi orders (i.e., reflexive and transitive relations) on the class of all sets. Clearly, $\emptyset \leq_{in} a$ for every set $a \neq \emptyset$. In the following we always assume that $a \neq \emptyset$.

Proposition 6.1 *Let $f : a \to b$ and $g : b \to a$ be maps. Assume that $g \circ f = id_a$. Then f is injective and g is surjective.*

Proposition 6.2 *Assume that $f : a \to b$ is injective. Then there is some map $g : b \to a$ such that $g \circ f = id_a$. That is, every injective map has at least one left inverse.*

Proposition 6.3 (AC) *Assume that $g : b \to a$ is surjective. Then there is some map $f : a \to b$ such that $g \circ f = id_a$. That is, under the assumption of AC, every surjective map has at least one right inverse.*

The proofs are very easy. The map f for Proposition 6.3 is defined with the help of a choice function on $\mathcal{P}(b) \setminus \{\emptyset\}$ which picks for every $c \in a$ some element $d \in g^{-1}(a) = \{d | g(d) = a\}$.

Hence, $a \leq_{in} b$ always yields $a \leq_{pr} b$ but the converse needs the AC. Thus $a \leq_{in} b$ iff $a \leq_{pr} b$ holds under the assumption of the axiom of choice.

For every map $f : a \to b$ the *equivalence kernel*, or just the kernel, is defined by $c_1 \sim_f c_2$ iff $f(c_1) = f(c_2)$. This is an equivalence relation on the set a where the classes are the largest subsets of a on which the map f is constant. As usual, $a/\sim_f$ denotes the set of equivalence classes and $c \mapsto [c]$ is the *canonical projection* q_f . The map $[c] \mapsto f(c)$ then is the *canonical injection* $\dot{f}$.

Proposition 6.4 *Every map $f : a \to b$ decomposes into a surjection followed by an injection: $\dot{f} \circ q_f = f$.*

Theorem 6.5 (Cantor-Bernstein) *$a \leq_{in} b$ and $b \leq_{in} a$ if and only if $a \approx b$.*

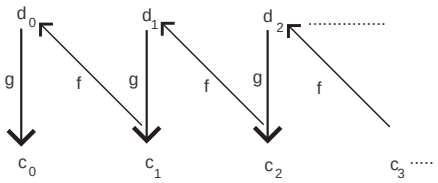

Figure 6.1: A moving element

PROOF. Let $f : a \to b$ and $g : b \to a$ be injections. We need to find a bijection from a to b. We call an element $c_0 \in a$ moving if it allows for an infinite diagram as in figure 6.1. That is, we can define two sequences $c_\nu = g(d_\nu)$ and $d_\nu = f(c_{\nu+1})$, $\nu \in \omega$, where c_ν has a (unique) counter image d_ν in b and where d_ν then has a (unique) counter image $c_{\nu+1}$ in a. We call an element $c \in a$ stationary if it is not moving. An element c is stationary if for a first ν we don't have a d_ν , i.e., c_ν is not in the range of the map g (c gets stopped in a) or we don't have a $c_{\nu+1}$, i.e., d_ν is not in the range of f (c gets stopped in b). Let a_1 be the subset of a consisting of all moving elements and the elements which are stopped in b; the set a_2 then is the complement of a_1, i.e., the set of all elements of a which are stopped in a. We define a map $h : a \to b$ by pieces. On a_1 an element c is mapped to d, where $g(d) = c$. This makes sense. If c is moving, then clearly $c \in ran(g)$. If c is not moving, then it got stopped in b, so again we must have that at least $c \in ran(g)$. On a_2 an element c is mapped to $f(c)$.
The map h is injective because f and g are injective, and $g(d_0) = c_0$ and $d_0 = f(c)$ for $c_0 \in a_1$ and $c \in a_2$ cannot happen. If c_0 is moving then $d_0 = f(c_1)$ for some $c_1 \in a_1$. If c_0 is stationary, it is stopped in b; hence again we conclude from $d_0 = f(c)$ that $c = c_1 \in a_1$.
Let $d \in b$. If $g(d)$ is moving, then $h(g(d)) = d$. If $g(d)$ is not moving and got stopped in b, then again $h(g(d)) = d$. If $g(d) = c$ got stopped in a, then we must have some c_1 in a such that $f(c_1) = d$, otherwise c would have been stopped at d in b. Clearly $c_1 \in a_2$ and $h(c_1) = d$. □

Corollary 6.6 *If $a \leq_{in} b \leq_{in} c$ and $a \approx c$ then $a \approx b$.*

PROOF. We have $a \leq_{in} b$ and $b \leq_{in} c \approx a$, yields also $b \leq_{in} a$. The claim follows now from Cantor-Bernstein. □

Theorem 6.7 (Cantor) *For any set a one has that $a <_{in} \mathcal{P}(a)$. That is $a \leq_{in} \mathcal{P}(a)$ but not $\mathcal{P}(a) \leq_{in} a$.*

PROOF. We have that $c \mapsto \{c\}$ establishes an injective map from a into $\mathcal{P}(a)$. We need to show that there is no surjection from a to $\mathcal{P}(a)$. Let $h : a \to \mathcal{P}(a)$ be any map. The set $r = \{c | c \in a, c \notin h(c)\}$ then is not within the range of h: $h(c_0) = r$ yields the Russel Paradox $c_0 \in h(c_0)$ iff $c_0 \in r$ iff $c_0 \notin h(c_0)$. Hence h is not surjective. □

For any quasi-order $\leq^*$ the relation $a \approx b$ which is defined by ($a \leq^* b$ and $b \leq^* a$) is an equivalence and $[a] \leq [b]$ iff $a \leq^* b$ is a partial order for the equivalence classes. It is called the *contraction* of $\leq^*$. The equivalence for $\leq_{in}$ is according to the Cantor-Bernstein Theorem the bijective equivalence $\approx$ and $[a] \leq [b]$ iff$a \leq_{in} b$ defines a partial order on the classes of equivalent sets. It is quite obvious that many set operations are invariant under $\approx$, e.g., if $a \approx b$ then $\mathcal{P}(a) \approx \mathcal{P}(b)$. However, some familiar facts, e.g., cartesian products are equivalent if the factors are equivalent, require (in case of infinite products) the help of the axiom of choice.

Proposition 6.8 *Let a,b,c be sets. Then:*

(i) $\mathcal{P}(a) \approx \underline{2}^a$;

(ii) $(a^b)^c \approx a^{(b\times c)}$*;*

(iii) $a^{(b\cup c)} \approx a^b \times a^c$ *in case that* $b \cap c = \emptyset$*;*

(iv) $a^c \times b^c \approx (a \times b)^c$

PROOF. Subsets can be identified with their characteristic functions. Actually $\mathcal{P}(a)$ and $\underline{2}^a$ are isomorphic as *boolean algebras*. Families of maps in one variable are just maps in two variables. This is (ii). The statement (iii) says that maps on disjoint domains can be patched together; (iv) is the basis of the *categorical* characterization of the cartesian product, i.e., pairs of maps (f, g) on a common domain correspond to maps h into the direct product of the codomains. □

Let r be the set of real numbers and (a, b) and (c, d) any two (proper) open intervals. Then $(a, b) \approx (c, d)$ by means of a simple linear equation. Clearly $(-1, 1) \leq_{in} [-1, 1] \leq_{in} (-2, 2)$ and $(-1, 1) \approx (-2, 2)$ then yields $(-1, 1) \approx [-1, 1]$. *Hence any two proper intervals, whether open, closed or half-open, are equivalent.* The arctangent function maps r bijectively onto the open interval $(-\pi/2, \pi/2)$. *Hence r is equivalent to any of its proper intervals.* The function $1/x$ maps $(0, 1]$ to $[1, \infty)$ and, as before, any two improper intervals are equivalent. Thus r *is equivalent to any of its intervals.* On the other hand, with the help of the binary representation of real numbers, one easily establishes $[0, 1] \approx \underline{2}^\omega$.

Proposition 6.9 *The set r of real numbers, the continuum, is equivalent to the powerset of the set ω of natural numbers.* □

Proposition 6.10 (Cantor) $\omega \approx \omega \times \omega$

PROOF. Every $\nu \in \omega$, $\nu \geq 1$, is of the form $\underline{2}^x \cdot (\underline{2} \cdot y + \underline{1})$ with uniquely determined x and y. This follows from the Prime Factorization Theorem. The following map: $f(x, y) = \underline{2}^x \cdot (\underline{2} \cdot y + \underline{1}) - \underline{1}$ then provides a bijection between ω and $\omega \times \omega$. □

Cantor's well-known proof based on an enumeration of the tableau $\omega \times \omega$ is easy to remember but a bit more difficult to formalize.

Corollary 6.11 $\omega^\nu \approx \omega$ *for every* $\nu \in \omega \setminus \{\underline{0}\}$. □

Corollary 6.12 $r^\nu \approx r^\omega \approx \omega^\omega \approx r$, $\nu \in \omega \setminus \{\underline{0}\}$.

PROOF. We have $\underline{2}^\omega \leq (\underline{2}^\omega)^\nu \approx \underline{2}^{\omega\times\nu} \leq \underline{2}^{\omega\times\omega} \leq \omega^{\omega\times\omega} \approx \omega^\omega \leq (\underline{2}^\omega)^\omega \approx \underline{2}^{\omega\times\omega} \approx \underline{2}^\omega$ and $r^\nu \approx (\underline{2}^\omega)^\nu$ and $r^\omega \approx (\underline{2}^\omega)^\omega$. We wrote $\leq$ instead of $\leq_{in}$ and we will from now on continue to do so for similar calculations. □

Hence the continuum r is equivalent to the finite dimensional spaces r^ν and even equivalent to the infinite dimensional space of all real sequences. The set q^+ of non-negative rational numbers is equivalent to a subset of $\omega\times\omega$, by means of the unique representation of a rational number as a reduced fraction. Hence, $q^+ \leq \omega \times \omega \approx \omega$. But $\omega \leq q^+$, therefore $q^+ \approx \omega$. *It follows that the set q of rational numbers is equivalent to the set of finite ordinals:* $q^+ \leq q \leq \{1, -1\} \times q^+ \approx \underline{2} \times \omega \leq \omega \times \omega \approx \omega$. Two continuous real functions which agree on the set q are equal. Let c be the set of all continuous functions. Then $r \leq c$ because every constant is continuous, and $c \leq r^q$ with $f \mapsto f \restriction q$. So $r \leq c \leq r^q \approx r^\omega \approx r$. *Hence r is equivalent to the set of all continuous functions.*

We have $\underline{2}^\omega \leq \omega \times \underline{2}^\omega \leq \underline{2}^\omega \times \underline{2}^\omega \approx r \times r \approx r$, thus $\omega \times \underline{2}^\omega \approx \underline{2}^\omega$ and therefore

Corollary 6.13 $r < \underline{2}^r \leq \omega^r \leq (\underline{2}^\omega)^r \approx \underline{2}^{\omega\times r} \approx \underline{2}^{\omega\times\underline{2}^\omega} \approx \underline{2}^r$. □

Therefore, the set of all real valued functions is equivalent to the set of functions which take on only two values and this set is strictly larger than the continuum.

For any sets a and b where $a \notin a$ and $b \notin b$ one has that $a \approx b$ iff $a \cup \{a\} = a^+ \approx b^+ = b \cup \{b\}$. Let $g : a^+ \to b^+$ be a bijection. If $g(a) = b$, then $g \restriction a$ is a bijection between a and b. If $g(c) = b$, then with the transposition (a, c) we get in $g \circ (a, c)$ a bijection $g' : a^+ \to b^+$ which maps a to b. In particular, $\underline{0} <_{in} \underline{1}$, leads to $\underline{1} <_{in} \underline{2}$. It follows, by induction, that for any finite ordinal ν one has that $\nu <_{in} \nu^+$. Hence, for finite ordinals one has that $\nu < \mu$ iff $\nu <_{in} \mu$. (Assume $\nu < \mu$. Then $\nu^+ \leq \mu$ and certainly $\nu^+ \leq_{in} \mu$, by means of the inclusion map. Hence $\nu <_{in} \nu^+ \leq_{in} \mu$. So $\nu < \mu$ yields $\nu <_{in} \mu$. Assume $\nu <_{in} \mu$. Then $\nu < \mu$, because otherwise $\mu \leq \nu$ and therefore, $\mu \leq_{in} \nu$.) A set a is called *finite* if it is equivalent to some finite ordinal ν. This ordinal ν is then unique and called the *cardinal* of a; $\nu = card(a)$ is a canonical representative of the class of all sets which are equivalent to the set a. Sets which are not finite are called *infinite*.

For the ordinal ω we have that $\omega \approx \omega^+$. The map $\omega \mapsto \underline{0}, \beta \mapsto \beta^+$ is a bijection between ω and ω^+. *Hence ω is an infinite set.* Our next goal is to show that for any ordinal α one has an ordinal β such that $\alpha <_{in} \beta$. Then automatically $\alpha < \beta$. Otherwise $\beta \subseteq \alpha$, so $\beta \leq_{in} \alpha$, thus $\alpha \approx \beta$. We are going to show somewhat more:

Proposition 6.14 *For any set a there is an ordinal α which is **not** equivalent to any subset b of a. The smallest such α is called the **Hartogs number** of a*

PROOF. Let $k(a) = \{w | w \in \mathcal{P}(a \times a),\ w \text{ a well-order of a subset } b \text{ of } a\}$. If w is in $k(a)$ then w is isomorphic to a unique ordinal β. That is, we have a functional relation $H(w, \beta), w \in k(a)$. If β is any ordinal which is equivalent to some subset b of a, then b carries a well-ordering w such that $H(w, \beta)$. The range of $H(w, \beta)$ is by replacement a set and obviously a segment. So it is of the form $S(\alpha)$ where α is the first ordinal not equivalent to any subset of a. □

Definition 6.1 An ordinal α is called an *initial* ordinal if $\beta < \alpha$ yields $\beta <_{in} \alpha$; that is, no ordinal smaller than α is equivalent to α. Initial ordinals are also called *cardinals*.

Proposition 6.15 *For any set a the Hartogs number $h(a) = \alpha$ is a cardinal.*

PROOF. Assume for some $\beta < \alpha$ that $\beta \approx \alpha$. Then β is equivalent to some subset b of a because α was the smallest ordinal for which this is not the case. But $\beta \approx \alpha$ would make b well-orderable by α. This is a contradiction. □

Proposition 6.16 *The class of cardinals is not a set.*

PROOF. Let a be any set of cardinals. Then $\beta = \bigcup a$ is an upper bound for a. But then $h(b) > \beta \geq \alpha$ holds for each $\alpha \in a$. That is, the cardinal $h(\beta)$ does not belong to the set a. □

Proposition 6.17 *All infinite cardinals are limit ordinals.*

PROOF. Assume that α is an ordinal and $\alpha \geq \omega$. We define a bijection $f : \alpha \to \alpha \cup \{\alpha\}$ by setting $f(\underline{0}) = \alpha$; $f(\nu) = \nu - \underline{1}, \nu \in \omega \setminus \{\underline{0}\}$; $f(\gamma) = \gamma$ for $\omega \leq \gamma < \alpha$. Hence α^+ is not a cardinal. □

We have that all finite ordinals are cardinals. If α is any ordinal then there is a smallest ordinal equivalent to it: $card(\alpha) = min\{\beta | \beta \approx \alpha\}$. For example, $\omega = card(\omega) = card(\omega + \underline{1}) = card(\omega + \underline{2}) = \ldots$ and we have already shown that there are cardinals larger than the first infinite cardinal, which is ω. An ordinal α is *countably infinite*, or *denumerable*, if $\omega = card(\alpha)$. Let ω_1 be the first cardinal which is not countable. Then $\omega_1 = \{\beta | \beta < \omega_1\}$ tells us that ω_1 is just the set of all countable ordinals, i.e., the ordinals which are either finite or denumerable.

Let a be a set which can be well-ordered. We then define:

$$card(a) = min\{\gamma|\gamma \approx a\}$$

In particular, $card(a) \approx a$. If α is an ordinal then $card(\alpha) \leq \alpha$ because α is well-ordered by α by the identity map. If α is a cardinal then $card(\alpha) = \alpha$ because no smaller ordinal is equivalent to α. If α and β are cardinals such that $\alpha \approx \beta$, then $\alpha = \beta$.

Lemma 6.18 *Let s be a subset of the ordinal α. Then $(s, \in)$ is well-ordered and $card(s) \leq \alpha$.*

PROOF. Indeed, we have that $(s, \in) \cong \alpha$ or $(s, \in) \cong S(\gamma) = \gamma$ for some $\gamma \in \alpha$, or $\alpha \cong \{\beta|\beta \in s, \beta < \gamma\}$, for some γ in s. The last case is impossible because it would give us some strictly increasing map $f : \alpha \to \alpha$ and we know from Proposition 3.9 on Ordinals that for any such map $f(\delta) \geq \delta$ holds. In particular, $f(\gamma) \geq \gamma$, so $f(\gamma)$ is not in the range of f which consists only of ordinals less than γ. The first two possibilities tell us that s can be well-ordered by α or by an ordinal less than α. Hence, $card(s) \leq \alpha$. □
In particular, $h(\omega) = \{\beta|\beta \approx s \subseteq \omega\} = \{\beta|\beta \approx \omega \text{ or } \beta \in \omega\} = \omega_1$.

Corollary 6.19 *If $b \leq_{in} a$ where a can be well-ordered, then b can be well-ordered and one has that $card(b) \leq card(a)$.*

PROOF. We have $b \approx s'$ for some subset s' of a and because of $a \approx card(a)$ we get $b \approx s$ for some subset s of $card(a)$. Hence, $card(b) = card(s) \leq card(a)$. □

Corollary 6.20 *Assume for the ordinal α that $card(\alpha) < \kappa$ where κ is a cardinal. Then $\alpha < \kappa$.*

PROOF. Indeed, if we had $\alpha \geq \kappa$, then κ would be a subset of α. The smallest ordinal that well-orders α, i.e., $card(\alpha)$, would also well-order the set κ. This well-ordering of κ is order-isomorphic to an ordinal β where $\beta \leq card(\alpha)$. But κ was a cardinal, thus $\kappa \leq \beta$, and therefore $\kappa \leq \beta \leq card(\alpha)$ which contradicts $card(\alpha) < \kappa$. □

Thus, an ordinal α is a cardinal if and only if for every ordinal β one has that

$$\beta < \alpha \text{ iff } card(\beta) <_{in} \alpha$$

The Continuum Hypothesis (CH) is the assertion that any subset s of real numbers which is not finite is equivalent to ω or to r.

We can avoid the reals in the formulation of CH by asserting that every subset of $\underline{2}^\omega$ which is infinite is equivalent to ω or to the whole set $\underline{2}^\omega$.

Proposition 6.21 *If $\underline{2}^\omega$ can be well-ordered then $\omega_1 = card(\underline{2}^\omega)$ is equivalent to the Continuum Hypothesis.*

PROOF. If $\underline{2}^\omega$ can be well-ordered then $\omega_1 \leq card(\underline{2}^\omega)$ because $\omega <_{in} \underline{2}^\omega \approx card(\underline{2}^\omega)$ but ω_1 is the smallest cardinal γ such that $\omega < \gamma$. It follows that $\underline{2}^\omega$ contains a subset s which is equivalent to ω_1.
If we assume CH then $s \approx \underline{2}^\omega$, hence $\omega_1 \approx \underline{2}^\omega$ and $\omega_1 = card(\omega_1) \approx card(\underline{2}^\omega)$. But equivalent cardinals are equal, so $\omega_1 = card(\underline{2}^\omega)$.
Assume that $\underline{2}^\omega$ can be well-ordered and that $\omega_1 = card(\underline{2}^\omega)$. Then let s be any infinite subset of $\underline{2}^\omega$. If $card(s) > \omega$, then $\omega_1 \leq card(s) \leq card(\underline{2}^\omega)$. Hence, $card(s) = card(\underline{2}^\omega)$ and it follows that $s \approx \underline{2}^\omega$, i.e., CH. □

Cantor stated the Continuum Hypothesis in 1878 and it was listed by Hilbert in his famous address of 1900 as the most important mathematical problem of the twentieth century. Hilbert believed in the

truth of CH and called it actually the Continuum Theorem in his 1925 paper "About Infinity". **Gödel** proved in 1938 that AC and CH are simultaneously consistent with ZF. Within any universe he formed the hierarchy of *constructible* sets: Roughly speaking, sets are constructed in stages. The critical case is the one from stage λ to λ^+. If the set a has already been constructed at stage λ then at λ^+ all subsets of a are collected which are given by formulas $Q(x)$ where all occurring parameters are sets already constructed at stages $\leq \lambda$. This constructible universe L settled many other open problems in the affirmative and there are mathematicians who firmly believe in $V = L$, i.e., the ZF-Hierarchy V should consist of constructible sets only. However, **P. Cohen** showed in 1963 that CH does not follow from AC and he even constructed a model where the reals cannot be well-ordered, i.e., where AC is not true. And CH can be true without having $V = L$. He invented a new method which he called *forcing* in order to obtain these models. His research revolutionized set theory and gave new impetus to systems of non-classical logics. *The CH and its generalization to arbitrary infinite sets implies, however, AC.* It is quite interesting to note that P. Cohen believes that CH is obviously false. He argues that ω_1 as the set of all types of well-orderings realizable on ω is, after all, a set one gets by adding one element at a time: $\omega_1 = \{\underline{1}, \underline{2}, \ldots, \omega, \omega + \underline{1}, \omega + \underline{2}, \ldots, \omega + \omega, \ldots\}$ while forming the power set should be perceived as a much stronger tool to generate a set of a larger cardinal.

6.2 Cardinals

If we assume AC, and we will do so from now on, then any set has *a* cardinal. Let $(\kappa)_j, j \in i$, be any family of cardinals. We define the *cardinal sum* as the cardinal of the disjoint union:

$$\sum_{j \in i} \kappa_j = card(\bigcup_{j \in i} \{j\} \times \kappa_j)$$

If a_j is a family of pairwise disjoint sets where $card(a_j) = \kappa_j$, then the sum of the κ_j is the cardinal of the union of the a_j. This is very easy to see. The *cardinal product* of the κ_j is defined as the cardinal of the cartesian product:

$$\prod_{j \in i} \kappa_j = card(\prod_{j \in i} \kappa_j)$$

If a_j is any family of sets where $card(a_j) = \kappa_j$, then the product of the κ_j is the cardinal of the cartesian product of the a_j. Again, this is very easy to see. If κ and λ are cardinals then $\kappa + \lambda$ and $\kappa \cdot \lambda$ denote the sum and product, respectively. It is obvious that sum and product are commutative and associative and that the multiplication is distributive over addition; for exponentiation κ^λ we have the familiar rules:

$$\begin{aligned} \kappa^{\lambda \cdot \rho} &= (\kappa^\lambda)^\rho \\ \kappa^\lambda \cdot \mu^\lambda &= (\kappa \cdot \mu)^\lambda \\ \kappa^\lambda \cdot \kappa^\rho &= \kappa^{\lambda + \rho} \end{aligned}$$

Finite sums and products agree with ordinary addition and multiplication. These are the addition and multiplication rules of finite combinatorics. Whenever we use cardinal and ordinal addition or multiplication simultaneously, the ordinal operations are encircled, e.g., $\omega \oplus \omega = \omega \odot \underline{2}$ while $\omega + \omega = 2 \cdot \omega = \omega \cdot \omega = \omega$. Another convention is to use Hebrew letters for cardinals, like Aleph($\aleph$) or Beth($\beth$).

We already know that the cardinals form a class, well-ordered by $\in$. Hence there is a unique order isomorphic correspondence between all ordinals and all infinite cardinals. This correspondence was called by Cantor $Aleph(\alpha, \aleph)$ and for which he introduced the notation $\aleph_\alpha$. So in particular:

$$\omega = \aleph_{\underline{0}}, \aleph_{\underline{0}} + \aleph_{\underline{0}} = \aleph_{\underline{0}}, \aleph_{\underline{0}} \cdot \aleph_{\underline{0}} = \aleph_{\underline{0}}, , \aleph_{\underline{0}} < \underline{2}^{\aleph_{\underline{0}}}, \aleph_{\underline{0}}^{\aleph_{\underline{0}}} = \underline{2}^{\aleph_{\underline{0}}}$$

The single most useful fact about cardinal arithmetic is

Theorem 6.22 *Let κ be an infinite cardinal. Then $\kappa \cdot \kappa = \kappa$.*

PROOF. For the proof we restate two auxiliary Lemmas which can be easily proven ad hoc. Recall, however, Theorem 3.34 and Theorem 3.37.

Lemma 6.23 *Let $(\omega_1, <_1)$ and $(w_2, <_2)$ be two well-ordered systems. Then the cartesian product $\omega_1 \times \omega_2$ becomes a well-ordered set by defining: $(c, d) < (c', d')$ iff $c <_1 c'$ or if $c = c'$ then $d <_2 d'$. $(w, <) = (w_1, <_1) \times (w_2, <_2)$ is called the **lexicographic product**. The ordinal product $\beta \odot \alpha$ is isomorphic to the lexicographic product $\alpha \times \beta$ of α and β.*

Lemma 6.24 *Let $(w_\beta, <_\beta), \beta < \alpha$, be a family of well-ordered systems indexed by the ordinal α. Assume that the w_β are pairwise disjoint. Then $w = \bigcup\{w_\beta | \beta < \alpha\}$ becomes a well-ordered set by defining: $c < d$ iff $c <_\beta d$ in case that c and d belong to the same w_β or if $c \in w_\beta$ and $d \in w_\gamma$ then $\beta < \gamma$. The set w is called the **order sum**. The ordinal sum $\alpha \oplus \beta$ is isomorphic to the order sum of α and β.*

If we write $|a|$ for $card(a)$ then according to the last statement of the preceding Lemma we have that $|\alpha \oplus \beta| = |\alpha| + |\beta|$. For example, if α is an infinite ordinal then $|\alpha \oplus \underline{1}| = |\alpha \cup \{\alpha\}| = |\alpha|$ according to $\alpha \mapsto \underline{0}, \nu \mapsto \nu + \underline{1}$ for $\nu < \omega, \beta \mapsto \beta$, for $\omega \leq \beta < \alpha$.

The proof of the theorem generalizes a particular proof for $\omega \approx \omega \times \omega$. Define on $\omega \times \omega$ the following well-ordering $<^*$. Let $p_\rho = \{(\nu, \mu) | \nu + \mu = \rho\}$. Then $\omega \times \omega = \bigcup \{p_\rho | \rho \in \omega\}$ where, of course, the union is disjoint. Each p_ρ carries the lexicographic ordering which it inherited from $\omega \times \omega$. The order sum of the p_ρ defines a well-ordering $<^*$ on $\omega \times \omega$ which is order-isomorphic to ω. We have:

$$(\underline{0}, \underline{0}) <^* (\underline{0}, \underline{1}) <^* (\underline{1}, \underline{0}) <^* (\underline{0}, \underline{2}) <^* (\underline{1}, \underline{1}) <^* (\underline{2}, \underline{0}) <^* \ldots$$

It is plain that this enumeration of $\omega \times \omega$ defines an isomorphism f between $(\omega, \in)$ and $(\omega \times \omega, <^*)$. A formal, recursive definition of f is left as an easy exercise.

In order to prove the theorem, assume that there is a smallest cardinal κ such that $\kappa \cdot \kappa \neq \kappa$. Because of $\kappa \cdot \kappa \geq \kappa$ we may assume $\kappa \cdot \kappa > \kappa$ for a smallest κ and $\lambda \cdot \lambda = \lambda < \kappa$ for every $\omega \leq \lambda < \kappa$. We already know that $\kappa > \omega$, but we will not need this fact.
Let $\alpha, \beta \in \kappa$. If α and β are both finite, then $\alpha \oplus \beta < \kappa$. Otherwise, $|\alpha \oplus \beta| = |\alpha| + |\beta| \leq \lambda + \lambda = \underline{2} \cdot \lambda \leq \lambda \cdot \lambda = \lambda < \kappa$ where λ is the maximum of $|\alpha|$ and $|\beta|$. Hence, by Corollary 6.20, $\alpha \oplus \beta < \kappa$ for all $\alpha, \beta \in \kappa$.
Let $p = \kappa \times \kappa = \{(\alpha, \beta) | \alpha, \beta \in \kappa\}$. For each $\xi \in \kappa$, let $p_\xi = \{(\alpha, \beta) | \alpha \oplus \beta = \xi\}$. The p_ξ constitute a partition of p. Each p_ξ is well-ordered as a subset of the lexicographic product $\kappa \times \kappa$. We now take on p the well-ordering $<^*$ which is the order sum of the p_ξ.
We have $(\alpha, \beta) <^* (\alpha_0, \beta_0)$ where $\alpha_0 \oplus \beta_0 = \xi_0$ iff $\big((\alpha \oplus \beta = \xi_0$ and $\alpha < \alpha_0$ or if $\alpha = \alpha_0$ one has that $\beta < \beta_0)$ or $(\alpha \oplus \beta < \xi_0)\big)$. At any rate, if $(\alpha, \beta) \leq^* (\alpha_0, \beta_0)$ then, $\alpha, \beta \leq \xi_0$.
The well-ordered system $(p, <^*)$ is order isomorphic to some ordinal δ. Because $card(p) = \kappa \times \kappa > \kappa$, we have that $\delta > \kappa$. Hence κ is isomorphic to a segment of $(p, <^*)$, say $\kappa \cong S(\alpha_0, \beta_0)$ where $(\alpha_0, \beta_0) \in p_{\xi_0}$ for some $\xi_0 < \kappa$.
If $(\alpha, \beta) \in S(\alpha_0, \beta_0)$, then $\alpha, \beta < \xi_0 \oplus \underline{1}$. Hence $S(\alpha_0, \beta_0) \subseteq (\xi_0 \oplus \underline{1}) \times (\xi_0 \oplus \underline{1})$. Thus $\kappa = |S(\alpha_0, \beta_0)| \leq |\xi \oplus \underline{1}| \times |\xi \oplus \underline{1}| = |\xi \oplus \underline{1}| < \kappa$, which is a contradiction. (For the last step we needed that κ is an infinite cardinal: If $\xi_0 < \kappa$, then $|\xi_0| < \kappa$ as well as $|\xi \oplus \underline{1}| < \kappa$). □

Corollary 6.25 *Let κ and λ be cardinals where at least one is infinite. Then*

$$\kappa + \lambda = max(\kappa, \lambda) = \kappa \cdot \lambda$$

PROOF. We first assume that $\omega \leq \kappa \leq \lambda$. Then $\lambda \leq \kappa + \lambda \leq \underline{2} \cdot \lambda \leq \kappa \cdot \lambda \leq \lambda \cdot \lambda = \lambda$.

If κ is infinite and λ finite then $\kappa + \lambda = \kappa$: $\kappa \leq \kappa + \lambda \leq \kappa + \kappa = \kappa$; $\kappa \cdot \lambda = \kappa$: $\kappa \leq \kappa \cdot \lambda \leq \kappa \cdot \kappa = \kappa$. □

Corollary 6.26 *Let κ and λ be cardinals such that $\omega \leq \kappa \leq \lambda$. Then*

$$\kappa^\lambda = \underline{2}^\lambda$$

PROOF. We have that $\underline{2}^\lambda \leq \kappa^\lambda \leq (\underline{2}^\kappa)^\lambda = \underline{2}^{\kappa \cdot \lambda} = \underline{2}^\lambda$. □

Corollary 6.27 *Assume $\kappa \geq \omega$. Let $(a_j)_{j \in i}$ be a family of sets where $card(a_j) = \kappa_j \leq \kappa, j \in i$, and $card(i) \leq \kappa$. Then*

$$card(\bigcup_{j \in i} a_j) \leq \kappa.$$

PROOF. This follows from $card(\bigcup_{j \in i} a_j) \leq \sum_{j \in i} \kappa_j \leq card(i) \cdot \kappa \leq \kappa \cdot \kappa = \kappa$. □

In particular, a countable union of countable sets is countable. It is quite remarkable that one cannot prove this without using AC.

Let $\kappa_j, j \in i$, be a family of cardinals with supremum α. Then α is a cardinal. Indeed, $\alpha \geq \kappa_j$ yields $card(\alpha) \geq \kappa_j$ for all $j \geq i$. Hence $card(\alpha)$ is an upper bound for the κ_j; but then $\alpha \geq card(\alpha)$ yields $card(\alpha) = \alpha$.

Corollary 6.28 *Let $\kappa_j \geq \underline{1}, j \in i$, be a family of cardinals where $\kappa = sup\{\kappa_j | j \in i\}$ and $\lambda = card(i)$. Then, if κ or λ are infinite, one has that*

$$\sum_{j \in i} \kappa_j = \lambda \cdot \kappa$$

PROOF. $\sum_j \kappa_j \leq \sum_j \kappa \leq card(i) \cdot \kappa = \lambda \cdot \kappa$. On the other hand, $\lambda = \sum_j \underline{1} \leq \sum_j \kappa_j$. And for each $j \in i$ one has that $\kappa_j \leq \sum_j \kappa_j$, therefore $\kappa \leq \sum_j \kappa_j$. Hence $\lambda \cdot \kappa = max(\kappa, \lambda) \leq \sum \kappa_j$. □

Theorem 6.29 (Zermelo-König) *Let for $j \in i$, $a_j \subset b_j$ where $a_j <_{in} b_j$ and where the a_j are pairwise disjoint. Then*

$$\bigcup a_j <_{in} \prod b_j$$

PROOF. Because $c_j = b_j \setminus a_j$ are non-empty, we have a choice function $j \mapsto d_j$ where the d_j belong to c_j . This gives us an injective map $f : \bigcup a_j \to \prod b_j$ from the disjoint sum a of the a_j into the direct product b of the b_j : If $e \in a$, then $e \in a_k$ for a unique $k \in i$. We define $f(e) \in \prod b_j$ such that $p_j(f(e)) = d_j$ for $j \neq k$, and $p_k(f(e)) = e$. It is easy to see that f is injective: Assume that $e \neq e'$. If e and e' belong to the same a_k, then $f(e)$ and $f(e')$ have different $k-$components, namely e and e'. If $e \in a_k$ and $e' \in a_{k'}$, where $k \neq k'$, then the $k-$component of $f(e)$ is e, while the $k-$component of $f(e')$ is $d_k \in c_k$, so certainly different from e, because $d_k \notin a_k$. Hence again, $f(e) \neq f(e')$.
Now let $f : \bigcup a_j \to \prod b_j$ be any function. We are going to show that f is not surjective. Because of $a_j <_{in} b_j$ none of the maps $p_j \circ f \restriction a_j$ from a_j into b_j is surjective (again by AC, see Proposition 6.3). We have a choice function $j \mapsto d_j$ where $d_j \in b_j$ but d_j is not the $j-$component of any $f(e)$ where $e \in a_j$. Hence $(d_j)_{j \in i}$ is an element of b which is not in the range of the map f. □

Corollary 6.30 *Assume for the family of cardinals that $\kappa_j < \kappa'_j, j \in i$. Then $\sum \kappa_j < \prod \kappa'_j$.*

PROOF. We apply the previous theorem to $b_j = \{j\} \times \kappa'_j$ and $a_j = \{j\} \times \kappa_j$. □

A function whose domain is an ordinal is called a *transfinite sequence*. Let $(\alpha_\beta)_{\beta\in\delta}$ be an increasing sequence of ordinals. If δ is a successor ordinal then the sequence has a largest element, namely α_γ where $\gamma^+ = \delta$. At any rate, we can form $\alpha = \bigcup\{\alpha_\beta|\beta < \delta\}$ which is the supremum of the α_β. If $\alpha' < \alpha$, then one has some $\beta_0 < \delta$ such that for all $\beta_0 \leq \beta < \delta$ one has that $\alpha' < \alpha_\beta$. We call α therefore the *limit* of the increasing sequence $(\alpha_\beta)_{\beta<\delta}$.

A subset c of an ordinal is said to be *bounded* in α if there is some $\beta < \alpha$ such that $\gamma \leq \beta$ holds for all $\gamma \in c$.
The **increasing** sequence $\sigma = (\alpha_\beta)_{\beta\in\delta}$ is said to be *cofinal* in λ if the range of σ, $ran(\sigma) = \{\alpha_\beta|\beta < \delta\}$, is an unbounded subset of λ.
Assume that $\sigma = (\alpha_\beta)_{\beta\in\delta}$ is cofinal in λ. Then δ and λ have to be limit ordinals. If δ is a successor, i.e., $\delta = \beta_0^+$, then σ has maximum $\alpha_{\beta_0} < \lambda$, so $ran(\sigma)$ is bounded in λ. On the other hand, any sequence with values within a successor ordinal γ has to be bounded, namely by the predecessor of γ. So λ is a limit ordinal.
The increasing sequence $\sigma = (\alpha_\beta)_{\beta\in\delta}$ is cofinal in λ iff $\lambda = lim(\sigma)$. Indeed, the ordinal λ is an upper bound of $ran(\sigma)$. It is the least upper bound iff $ran(\sigma)$ is not bounded in λ, i.e., σ is cofinal in λ.

Proposition 6.31 *For any limit ordinal λ, the identity on λ is cofinal in λ. If $\tau : \delta \to \delta'$ is cofinal in δ' and $\sigma : \delta' \to \lambda$ is cofinal in λ then $\sigma \circ \tau : \delta \to \lambda$ is cofinal in λ.* □

We remark that the subset a of λ is unbounded in λ iff $a^* = \{\beta|\beta \leq \alpha, \alpha \in a\}$ is unbounded in λ. In this case one has that $\bigcup a^* = \bigcup a = \lambda$.

Definition 6.2 Let λ be a limit ordinal. The *cofinality* of λ, $cf(\lambda)$, is the least ordinal δ such that there is an increasing transfinite sequence $\sigma = (\alpha_\beta)_{\beta\in\delta}$ of ordinals within λ which is cofinal in λ.

Proposition 6.32 *Let λ be a limit ordinal and $\kappa = card(\lambda)$. Then $cf(\lambda) \leq \kappa$.*

PROOF. We have to find an increasing map $g : \delta \to \lambda$ which is cofinal in λ and where $\delta \leq \kappa$. We know that there is a bijection $f : \kappa \to \lambda$ because $\kappa \approx \lambda$. However, we will only need the fact that $ran(f)$ is unbounded in λ. The map f gives rise to the map $g' : \kappa \to \lambda \cup \{\lambda\}, g'(\beta) = \bigcup\{f(\alpha)|\alpha < \beta\}$. Clearly, g' is increasing. If $\gamma \in \lambda$, then $\gamma^+ < \lambda$ because λ is a limit ordinal. There is some $\beta \in \kappa$ such that $f(\beta) \geq \gamma^+$ and we may choose $\beta = \min\{\alpha|\alpha \in \kappa, f(\alpha) \geq \gamma^+\}$. Then $\gamma \in g'(\beta^+) = g'(\beta) \cup f(\beta)$. Hence, $\bigcup\{g'(\beta)|\beta \in \kappa\} = \lambda$. If $g'(\beta) \in \lambda$ for all $\beta \in \kappa$, then let $g' = g$ and $\delta = \kappa$. Otherwise, let $\delta \in \kappa$ be the smallest ordinal such that $g'(\delta) = \lambda$ and put $g = g' \restriction \delta$. If we had $\delta = \beta^+$, then $g'(\beta^+) = g'(\beta) \cup f(\beta) = \lambda$ which contradicts $g'(\beta) < \lambda, f(\beta) < \lambda$. Hence, δ is a limit ordinal and we conclude: $\bigcup\{g(\beta)|\beta \in \delta\} \supseteq \bigcup\{g(\beta^+)|\beta \in \delta\} \supseteq \{f(\beta)|\beta \in \delta\} = g'(\delta) = \lambda$. □

Remark 6.1 *Let λ and μ be limit ordinals where $\lambda < cf(\mu)$. Let $f : \lambda \to \mu$ be any function from λ into μ. Then f is bounded, i.e., there is some $\gamma \in \mu$ such that $f(\alpha) \leq \gamma$ holds for all $\alpha \in \lambda$.*

This follows from the proof of the proposition.

Corollary 6.33 *Let λ be a limit ordinal. Then $cf(\lambda)$ is a cardinal.*

PROOF. Let $\kappa = card(cf(\lambda))$. By the previous theorem we have a cofinal map $\tau : \delta \to cf(\lambda), \delta \leq \kappa$, and a cofinal map $\sigma : cf(\lambda) \to \lambda$. The composition $\tau \circ \sigma : \delta \to \lambda$ is cofinal and $\kappa < cf(\lambda)$ would contradict the definition of $cf(\lambda)$. □

Proposition 6.34 *Let λ be a limit ordinal. Then $cf(cf(\lambda)) = cf(\lambda)$.*

PROOF. We have $cf(cf(\lambda)) \leq cf(\lambda)$ and, as in the previous proof, $cf(cf(\lambda)) < cf(\lambda)$ would contradict the definition of $cf(\lambda)$. □

Definition 6.3 An infinite cardinal κ is called a *regular cardinal* if $\kappa = cf(\kappa)$. An infinite cardinal which is not regular is called *singular*.

For example, $cf(cf(\lambda))$ is for any limit ordinal some regular cardinal. The first infinite cardinal, ω, is obviously a regular cardinal.

Recall that the infinite cardinals form a well ordered class and that there is a unique order isomorphism between the class of all ordinals and the class of all infinite cardinals which we called $Aleph(\alpha, \kappa)$, i.e., $\kappa = \aleph_\alpha$. For example, $\aleph_{\underline{0}} = \omega$ and $card(\underline{2}^\omega) = \aleph_\beta$ for some β. We are going to show that β cannot be ω, one of the few definitive statements one can make on the basis of ZF with the AC.

Definition 6.4 A cardinal κ is called a *successor cardinal* if $\kappa = \aleph_\alpha$ for some successor ordinal α. Otherwise it is called a *limit cardinal*, i.e., $\kappa = \aleph_\lambda$ for some limit ordinal λ.

Proposition 6.35 *Let α be a limit ordinal. Then $\aleph \restriction \alpha : \alpha \to \aleph_\alpha, \beta \mapsto \aleph_\beta$ is cofinal.*

PROOF. Let $\gamma < \aleph_\alpha$. Then $card(\gamma) < \aleph_\alpha$, thus $card(\gamma) = \aleph_\beta$ for some $\beta < \alpha$. Because α is a limit ordinal, we have that $\beta^+ < \alpha$. Hence, $card(\gamma) = \aleph_\beta \leq \gamma < \aleph_{\beta^+} < \aleph_\alpha$. □

The second limit cardinal, $\aleph_{\omega \oplus \omega}$, is singular because of $cf(\omega \oplus \omega) = \omega$ by means of the map $\nu \mapsto \omega \oplus \nu$. By the previous proposition, $\delta : \omega \to \aleph_{\omega \oplus \omega}, \nu \mapsto \aleph_{\omega \oplus \nu}$ is cofinal. Hence, $cf(\aleph_{\omega \oplus \omega}) = \omega$.

Proposition 6.36 *All infinite successor cardinals are regular.*

PROOF. Let κ be an infinite cardinal and $\sigma : \delta \to \kappa^+$ be an increasing map where $\delta < \kappa^+$. For any $\beta \in \delta$ we have that $\sigma(\beta) < \kappa^+$. We have $card(\delta) \leq \kappa$ and $card(\sigma(\beta)) \leq \kappa, \beta \in \delta$. It follows that $card(\bigcup\{\sigma(\beta))|\beta \in \delta\}$ is at most $\kappa \circ \kappa = \kappa$, hence $a = \bigcup\{\sigma(\beta))|\beta \in \delta\} \subset \aleph^+$, and therefore σ cannot be cofinal. □

Theorem 6.37 (Hausdorff's Formula) *Let κ and λ be infinite cardinals. Then*

$$(\kappa^+)^\lambda = \kappa^\lambda \cdot \kappa^+$$

*Here κ^+ is the smallest **cardinal** larger than κ.*

PROOF. If $\lambda \geq \kappa^+$ then (by Corollary 6.26 of Theorem 6.22 one has $(\kappa^+)^\lambda = \underline{2}^\lambda$; $\kappa^\lambda = \underline{2}^\lambda$, $\kappa^+ \leq \lambda < \underline{2}^\lambda$. Hence, $(\kappa^+)^\lambda = \underline{2}^\lambda$ and $\kappa^\lambda \cdot \kappa^+ = \underline{2}^\lambda \cdot \underline{2}^\lambda = \underline{2}^\lambda$.
Recall that $(\kappa^+)^\lambda$ is the cardinal of the set of all functions f from λ into κ^+. If $\lambda < \kappa^+$, then by Remark 6.1 after the proof of Proposition 6.32 and because of $cf(\kappa^+) = \kappa^+$, every such function f is bounded by an element of κ^+. That is, $(\kappa^+)^\lambda = \bigcup\{\gamma^\lambda|\gamma \in \kappa^+\}$ as sets; but the cardinal of the right-hand side can be estimated by: $card(\bigcup\{\gamma^\lambda|\gamma \in \kappa^+\}) \leq \sum_{\gamma \in \kappa^+} \kappa^\lambda \leq \kappa^+ \cdot \kappa^\lambda$.
On the other hand, $\kappa^+ \leq (\kappa^+)^\lambda$, $\kappa^\lambda \leq (\kappa^+)^\lambda$ and therefore $\kappa^+ \cdot \kappa^\lambda \leq (\kappa^+)^\lambda$. □

Corollary 6.38 *Let κ be an infinite cardinal. Then $(\kappa^+)^\kappa = \underline{2}^\kappa$.*

PROOF. $(\kappa^+)^\kappa = \kappa^\kappa \cdot \kappa^+ = \underline{2}^\kappa \cdot \kappa^+ = \underline{2}^\kappa$. □

In Aleph notation, Hausdorff's formula reads

$$\left(\aleph_{\alpha \oplus \underline{1}}\right)^{\aleph_\beta} = \left(\aleph_\alpha^{\aleph_\beta}\right) \cdot \aleph_{\alpha \oplus \underline{1}}$$

Proposition 6.39 *Let λ be an infinite limit ordinal. Then $cf(\aleph_\lambda) = cf(\lambda)$.*

PROOF. Let $\sigma : \delta \to \lambda$ be cofinal. Then $\sigma' : \delta \to \aleph_\lambda, \beta \mapsto \aleph_{\sigma(\beta)}$ is increasing, because Aleph is strictly increasing. We show that σ' is cofinal. Let $\pi : \aleph_\lambda \to \lambda, \alpha \mapsto \mu$ if α is an infinite ordinal and $card(\alpha) = \aleph_\mu$, and $\alpha \mapsto \underline{0}$ if α is finite. (Here we use that $\lambda > \underline{0}$, there is no map from $\aleph_{\underline{0}}$ into the empty set.) Now let $\alpha \in \aleph_\lambda$ and $\mu = \pi(\alpha) \in \lambda$. Then $\mu < \sigma(\beta)$ for some $\beta \in \delta$. But then $\aleph_\mu < \aleph_{\sigma(\beta)}$ and because of $card(\alpha) \leq \aleph_\mu$ we conclude $\alpha \in \aleph_{\sigma(\beta)} = \sigma'(\beta)$.
Now let $\sigma' : \delta \to \aleph_\lambda$ be cofinal. The map $\sigma : \delta \to \lambda, \beta \mapsto \pi(\sigma'(\beta))$ is obviously increasing. In order to show cofinality, let $\alpha \in \lambda$. Then $\aleph_\alpha < \aleph_{\alpha^+} < \sigma'(\beta)$ for some $\beta \in \delta$ but then $\pi(\sigma'(\beta)) \geq \alpha^+ > \alpha$. □

For $\lambda = \underline{0}$, the last proposition is not true: $cf(\aleph_{\underline{0}}) = \aleph_{\underline{0}} > \underline{0}$; $\aleph_{\underline{0}}$ is a limit cardinal and regular.

Definition 6.5 An uncountable cardinal κ is called *weakly inaccessible* if it is

(i) a limit cardinal, i.e., $\kappa = \aleph_\lambda$ for some limit ordinal λ,

(ii) regular, i.e., $cf(\aleph_\lambda) = \aleph_\lambda$.

Corollary 6.40 *Let $\aleph_\lambda$ be a weakly inaccessible cardinal. Then $\aleph_\lambda = \lambda$.*

PROOF. By the previous proposition, $cf(\aleph_\lambda) = cf(\lambda)$ because $\aleph_\lambda$ is uncountable and a limit cardinal. We certainly have $\lambda \leq \aleph_\lambda$ because Aleph is strictly increasing on Ord(x) (compare Proposition 3.9 on ordinals). But if we assume that $\lambda < \aleph_\lambda$, then $cf(\aleph_\lambda) = cf(\lambda) \leq \lambda < \aleph_\lambda$ shows that $\aleph_\lambda$ cannot be regular. □

The existence of weakly inaccessible cardinals cannot be proven on the basis of ZF.

Definition 6.6 A cardinal κ_μ is called *inaccessible* if it is weakly inaccessible and if $\underline{2}^{\aleph_\lambda} < \aleph_\mu$ holds for all $\lambda < \mu$.

One can show that V_κ is a model of ZF in case that κ is an inaccessible cardinal.

Theorem 6.41 $cf(\underline{2}^{\aleph_{\underline{0}}}) > \omega$.

PROOF. Let $\sigma : \omega \to \underline{2}^{\aleph_{\underline{0}}}$ be any increasing map. Then $card\big(\bigcup\{\sigma(\nu)|\nu \in \omega\}\big) \leq \sum \kappa_\nu$, where $\kappa_\nu = card(\sigma(\nu)) < \underline{2}^{\aleph_{\underline{0}}}$. Hence, by Zermelo-König's theorem, we conclude that $\sum \kappa_\nu < \big(\underline{2}^{\aleph_{\underline{0}}}\big)^{\aleph_{\underline{0}}} = \underline{2}^{\aleph_{\underline{0}}}$. Hence, $\bigcup\{\sigma(\nu)|\nu \in \omega\} \subset \underline{2}^{\aleph_{\underline{0}}}$. Thus the map σ is not cofinal. □

Corollary 6.42 $\underline{2}^{\aleph_{\underline{0}}} \neq \aleph_\omega$ *, or more generally,* $\underline{2}^{\aleph_{\underline{0}}} \neq \aleph_\alpha$ *where α is a limit ordinal and $cf(\alpha) = \omega$.*

PROOF. We have already noticed that $cf(\aleph_\omega) = \omega$. The claim now follows from Theorem 6.41. □

If we set $\underline{2}^{\aleph_\alpha} = \aleph_{E(\alpha)}$ then $E(\alpha) = \alpha \oplus \underline{1}$ is the generalized continuum hypothesis. It is known that any monotone functional relation E on the ordinals subject to the conditions that $E(\alpha) \geq \alpha \oplus \underline{1}$ and $cf(\aleph_{E(\alpha)}) > \aleph_\alpha$ can solve $\underline{2}^{\aleph_\alpha} = \aleph_{E(\alpha)}$ in a suitable model of set theory.