Chapter 3
Axioms of Set Theory

Every mathematical science relies upon demonstration rather than argument and opinion. Certain principles, called premises, are granted, and a demonstration is made which resolves everything easily and clearly.

GIOSEFFO ZARLINO
Le Istitutioni Harmoniche, 1558

Why Axioms?

In the middle and late 19th century, members of the then small mathematical community began the search for a rigorous foundation of Mathematics. In accordance with the Euclidean model of reasoning, the ideal foundation ought to consist of a few simple, clear principles, so-called *axioms*, on which the rest of knowledge can be built via firm and reliable thoughts free of contradiction. However, at the time it was not clear what assumptions should be made and what operations should be allowed in mathematical reasoning.

At the beginning of the last book of *Politeia* [79], Plato develops his theory of ideas. Translated into the mathematical setting, Plato's theory of ideas reads as follows: Even though there may be more than one human approach to Mathematics, there is only *one* true idea of Mathematics (*i.e.*, a unique mathematical world), and from this idea alone can one attain real knowledge. In particular, the mathematical world already exists and is just waiting to be discovered. So, from a Platonic point of view it would make sense to search for the unique set of true axioms for Set Theory—also because the axioms of Set Theory are supposed to describe the world of "real" Mathematics.

However, if we consider Set Theory as a mathematical discipline, then, like in any other field of Mathematics, there is no true axiom system, and moreover, we are even allowed to weaken the axioms or to extend them by additional assumptions in order to get weaker or stronger theories. This is done, for example, in Group Theory when we study *semigroups* or *monoids*, or focus on *abelian groups*.

It is often the case that a mathematical theory is developed long before its formal axiomatisation, and in rare instances, mathematical theories were already partially developed before mathematicians were aware of them. This happened with Group Theory: Around the year 1600 in England it was discovered that by altering the fittings around each bell in a bell tower, it was possible for each ringer to maintain precise control of when his (there were no female ringers then) bell sounded. This enabled the ringers to ring the bells in any particular order, and either maintain that order or permute the order in a precise way. (For technical reasons, not every permutation is allowed. In fact, just products of mutually disjoint elementary transpositions may be used, which means that two bells can exchange their places only if they are adjacently rung before-hand.) So, in the first half of the 17th century the ringers tried to continuously change the order of the bells for as long as possible, while not repeating any particular order. This game evolved into a challenge to ring the bells in every possible order, without any repeats, and return to the initial order at the end. Thus, bell-ringers began to investigate permutations and Stedman's work *Campanologia* [94] can fairly be said to be the first work in which Group Theory was successfully applied to a "musical" situation and consequently, Stedman can be regarded as the first group theorist. This also shows that permutations—the prototype of finite groups—were first studied in the 17th century in the context of change-ringing, and therefore had a practical application long before they were used in Lagrange's work of 1770–1771 on the theory of algebraic equations.

Let us now turn back to Set Theory. The history of Set Theory is rather different from the history of most other areas of Mathematics. Usually a long process can be traced in which ideas evolve until an ultimate flash of inspiration, often by a number of mathematicians almost simultaneously, produces a discovery of major importance. Set Theory, however, is the creation of only one person, namely of Georg Cantor (1845–1918), who first discovered that infinite sets may have different sizes, *i.e.*, cardinalities. In fact, the birth of Set Theory dates to 1873 when Cantor proved that the set of real numbers is uncountable. Until then, no one envisioned the possibility that infinities come in different sizes, and moreover, mathematicians had no use for the *actual infinite*—in contrast to the *potential infinite*, as it is introduced by Aristotle in *Physics* [1] Book III. The difference between actual and potential infinite is that the latter just means "unlimited" or "arbitrarily large" (*e.g.*, there are arbitrarily large—and therefore arbitrarily many—prime numbers), whereas the former means that there are infinite objects which actually exist (*e.g.*, there exists a set containing all, *i.e.*, infinitely many, prime numbers). Moreover, Cantor also showed that for every infinite set, there is a set of larger cardinality, which implies that there is no largest set. Cantor never introduced formal axioms for Set Theory, even though he was tacitly using most of the axioms introduced later by Zermelo and Fraenkel. However, Cantor considered a set as any collection of well-distinguished objects of our mind, which leads directly to RUSSELL'S PARADOX: On the one hand, there are collections which contain themselves as a member (*i.e.*, the collection of all sets is a set which is a member of itself). On the other hand, there are collections which do not contain themselves as a member (*i.e.*, the set of negative natural numbers, since it is empty, cannot be a member of itself). Now, call a set x *good* if x is not a member of itself and let C be the collection of all sets which are good. Is C, as a set, good or not? If C is good, then C is *not* a member of itself, but since C contains all sets which are good, C *is* a member of C, a contradiction. Otherwise, if C is a member of itself, then C must be good, again a contradiction. In order to avoid this paradox we have to exclude the collection C from being a set, but then, we have to give reasons why certain collections are sets and others are not. The axiomatic way to do this is described by Zermelo as follows: *Starting with the historically grown Set Theory, one has to search for the principles required for the foundations of this mathematical discipline. In solving the problem we must, on the one hand, restrict these principles sufficiently to exclude all contradictions and, on the other hand, take them sufficiently wide to retain all the features of this theory.*

The principles, which are called axioms, will tell us how to get new sets from already existing ones. In fact, most of the axioms of Set Theory are constructive to some extent, *i.e.*, they tell us how new sets are *constructed* from already existing ones and what elements they contain.

The Axioms of Zermelo–Fraenkel Set Theory

In 1905, Zermelo began to axiomatise Set Theory and in 1908 he published his first axiomatic system consisting of seven axioms. In 1922, Fraenkel and Skolem independently improved and extended Zermelo's original axiomatic system, and the final version was again presented by Zermelo in 1930.

In this section we give Zermelo's axiomatic system of Set Theory, called *Zermelo–Fraenkel Set Theory*, denoted ZF. This axiomatic system contains all axioms of modern Set Theory except the Axiom of Choice, which will be introduced later in this chapter and discussed in great detail in Chapter 6. Alongside the axioms of Set Theory we shall develop the theory of ordinals and give various notations which will be used throughout this book.

Before we begin to present the axioms of Set Theory, let us say a few words about Set Theory in general: The language of Set Theory contains only one non-logical symbol, namely the binary **membership relation**, denoted by $\in$, and there exists just one type of object, namely *sets*. In other words, every object in the domain is a set and there are no other objects than sets. However, to make life easier, instead of $\in(a, b)$ we write $a \in b$ (or on rare occasions also $b \ni a$) and say that "a is an element of b", or that "a belongs to b". Furthermore, we write $a \notin b$ as an abbreviation of $\neg(a \in b)$. Later we will extend the language of Set Theory by defining some constants (like "$\emptyset$" and "ω"), relations (like "$\subseteq$"), and operations (like the power set operation "$\mathscr{P}$"), but in fact, all that can be formulated in Set Theory can be written as a formula containing only the non-logical relation "$\in$" (but for obvious reasons, we will usually not do so).

After these general remarks, let us now start to present the axioms of Set Theory.

0. The Axiom of Empty Set

$$\exists x \forall z (z \notin x)$$

This axiom not only postulates the existence of a set without any elements, *i.e.*, an empty set, it also shows that the set-theoretic universe is non-empty, because it contains at least an empty set (of course, the logical axioms L_{14} and L_{11} already incorporate this fact).

1. The Axiom of Extensionality

$$\forall x \forall y \big(\forall z (z \in x \leftrightarrow z \in y) \to x = y\big)$$

This axiom says that any sets x and y having the same elements are equal. Notice that the converse—which is $x = y$ implies that x and y have the same elements—is just a consequence of the logical axiom L_{15}.

The Axiom of Extensionality also shows that the empty set, postulated by the Axiom of Empty Set, is unique. For assume that there are two empty sets x_0 and x_1, then we have $\forall z (z \notin x_0 \wedge z \notin x_1)$, which implies that $\forall z (z \in x_0 \leftrightarrow z \in x_1)$, and therefore, $x_0 = x_1$.

Let us introduce the following notation: If $\varphi(x)$ is any first-order formula with free variable x (*i.e.*, x occurs at a particular place in the formula φ where it is not in the range of any logical quantifier), then

$$\exists! x \, \varphi(x) \quad :\Longleftrightarrow \quad \exists x \big(\varphi(x) \wedge \forall z \big(\varphi(z) \to z = x\big)\big).$$

With this definition we can reformulate the Axiom of Empty Set as follows:

$$\exists! x \forall z (z \notin x) \, ,$$

and this unique empty set is denoted by $\emptyset$.

We say that y is a **subset** of x, denoted $y \subseteq x$, if $\forall z (z \in y \to z \in x)$. Notice that the empty set is a subset of every set. If y is a **proper subset** of x, *i.e.*, $y \subseteq x$ and $y \neq x$, then this is sometimes denoted by $y \subsetneq x$.

One of the most important concepts in Set Theory is the notion of *ordinal number*, which can be seen as a transfinite extension of the natural numbers. In order to define the concept of ordinal numbers, we must first give some definitions: Let $z \in x$. Then z is called an **$\in$-minimal element** of x if $\forall y (y \notin z \vee y \notin x)$, or equivalently, $\forall y (y \in z \to y \notin x)$. A set x is **ordered by** $\in$ if for any sets $y_1, y_2 \in x$ we have $y_1 \in y_2$, or $y_1 = y_2$, or $y_1 \ni y_2$, but we do not require the three cases to be mutually exclusive. Now, a set x is called **well-ordered by** $\in$ if it is ordered by $\in$ and every non-empty subset of x has an $\in$-minimal element. Further, a set x is called **transitive** if $\forall y(y \in x \rightarrow y \subseteq x)$. Notice that if x is transitive and $z \in y \in x$, then this implies $z \in x$. A set is called an **ordinal number**, or just an **ordinal**, if it is transitive and well-ordered by "$\in$". Ordinal numbers are usually denoted by Greek letters like α, β, γ, λ, *et cetera*, and the collection of all ordinal numbers is denoted by Ω. We will see later, when we know more properties of ordinals, that Ω is not a set. However, we can consider "$\alpha \in \Omega$" as an abbreviation of "α is an ordinal", which is just a property of α, and thus, there is no harm in using the symbol Ω in this way, even though Ω is *not* an object of the set-theoretic universe.

FACT 3.1. *If* $\alpha \in \Omega$*, then either* $\alpha = \emptyset$ *or* $\emptyset \in \alpha$.

Proof. Since $\alpha \in \Omega$, α is well-ordered by $\in$. Thus, either $\alpha = \emptyset$, or, since $\alpha \subseteq \alpha$, α contains an $\in$-minimal element, say $x_0 \in \alpha$. If $x_0 \neq \emptyset$, then we find a $z \in x_0$, and by transitivity of α, we get $z \in \alpha$, which implies that z is not an $\in$-minimal element of α. Hence, we must have $x_0 = \emptyset$, which shows that $\emptyset \in \alpha$. $\dashv$

Notice that until now, we cannot prove the existence of any ordinal—or even of any set—beside the empty set, postulated by the Axiom of Empty Set. This will change with the following axiom.

2. The Axiom of Pairing

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

Informally, we just write

$$\forall x \forall y \exists u \big(u = \{x, y\}\big),$$

where $\{x, y\}$ denotes the set which contains just the elements x and y.

In particular, for $x = y$ we get the set $u = \{x, x\}$, which is, by the Axiom of Extensionality, the same as $\{x\}$. Thus, by the Axiom of Pairing, if x is a set, then $\{x\}$ is also a set. Starting with $\emptyset$, an iterated application of the Axiom of Pairing yields for example the sets $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}, \dots$, and $\{\emptyset, \{\emptyset\}\}$, $\{\{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \dots$. Among these sets, $\emptyset$, $\{\emptyset\}$, and $\{\emptyset, \{\emptyset\}\}$ are ordinals but, for example, $\{\{\emptyset\}\}$ is not an ordinal.

So far, we have not excluded the possibility that a set may be an element of itself, and in fact, we need the Axiom of Foundation in order to do so. However, we can already show that no ordinal is an element of itself:

FACT 3.2. *If* $\alpha \in \Omega$*, then* $\alpha \notin \alpha$.

Proof. Assume towards a contradiction that $\alpha \in \alpha$. Then $\{\alpha\}$ is a non-empty subset of α and therefore contains an $\in$-minimal element. Now, since $\{\alpha\}$ just contains the element α, the $\in$-minimal element of $\{\alpha\}$ must be α, but on the other hand, $\alpha \in \alpha$ implies that α is not $\in$-minimal, a contradiction. $\dashv$

For any sets x and y, the Axiom of Extensionality implies that $\{x, y\} = \{y, x\}$. So, it does not matter in which order the elements of a 2-element set are written down. However, with the Axiom of Pairing we can easily define **ordered pairs**, denoted $\langle x, y\rangle$, as follows:

$$\langle x, y\rangle := \big\{\{x\}, \{x, y\}\big\}$$

Notice that $\langle x, y\rangle = \langle x', y'\rangle$ *iff* $x = x'$ *and* $y = y'$, and further notice that this definition also makes sense in the case when x and y are equal—at least as long as we know that $\{\{x\}\}$ is supposed to denote an ordered pair. By a similar trick, one can also define ordered triples by stipulating, for example, $\langle x, y, z\rangle := \langle x, \langle y, z\rangle\rangle$, ordered quadruples, *et cetera*, but the notation becomes hard to read and it requires additional methods to distinguish, for instance, between ordered pairs and ordered triples. However, once we have more axioms at hand, we will be able to define arbitrarily large tuples.

3. The Axiom of Union

$$\forall x \exists u \forall z \big(z \in u \leftrightarrow \exists w \in x (z \in w)\big)$$

More informally, for all sets x there exists the union of x, denoted $\bigcup x$, consisting of all sets which belong to at least one element of x.

For sets x and y, let $x \cup y := \bigcup\{x, y\}$ denote the **union** of x and y. Notice that $x = \bigcup\{x\}$. For $x \cup y$, where x and y are disjoint (*i.e.*, do not have any common elements), we sometimes write $x \dot{\cup} y$, and for $x = \{y_\iota : \iota \in I\}$ we sometimes write $\bigcup_{\iota \in I} y_\iota$ instead of $\bigcup x$.

Now, using the Axiom of Union and the Axiom of Pairing, and by defining the successor x^+ of x by stipulating $x^+ := x \cup \{x\}$, we can build, for example, the following sets (which are in fact ordinals): $0 := \emptyset$, $1 := 0^+ = 0 \cup \{0\} = \{0\}$, $2 := 1^+ = 1 \cup \{1\} = \{0, 1\}$, $3 := 2^+ = 2 \cup \{2\} = \{0, 1, 2\}$, and so on. This construction leads to the following definition:

A set x such that $\forall y (y \in x \to (y \cup \{y\}) \in x)$ is called **inductive**. Obviously, the empty set $\emptyset$ is inductive, but of course, this definition is only useful if some other inductive sets exist. However, in order to guarantee that non-empty inductive sets exist we need the following axiom.

4. The Axiom of Infinity

$$\exists I \big(\emptyset \in I \wedge \forall y \big(y \in I \to (y \cup \{y\}) \in I\big)\big)$$

Informally, the Axiom of Infinity postulates the existence of a non-empty inductive set containing $\emptyset$. All the sets $0, 1, 2, \ldots$ constructed above—which we recognise as natural numbers—must belong to every inductive set. So, if there were a set which contains just the natural numbers, it would be the "smallest" inductive set containing the empty set. In order to construct this set, we need the following axiom.

5. The Axiom Schema of Separation

For each first-order formula $\varphi(z, p_1, \ldots, p_n)$ with $\mathrm{free}(\varphi) \subseteq \{z, p_1, \ldots, p_n\}$, the following formula is an axiom:

$$\forall x \forall p_1 \cdots \forall p_n \exists y \forall z \big(z \in y \leftrightarrow \big(z \in x \wedge \varphi(z, p_1, \ldots, p_n)\big)\big)$$

One can think of the sets $p_1, \ldots, p_n$ as parameters of φ, which are usually some fixed sets. Informally, for each set x and every first-order formula $\varphi(z)$, $\{z \in x : \varphi(z)\}$ is a set. Notice that the Axiom Schema of Separation allows us to separate all elements z of a given set x that satisfy a certain property φ into a new set, without allowing us to build the collection of all sets z that satisfy φ.

Before we give some applications of the Axiom Schema of Separation, let us introduce the following notation: If $\varphi(z)$ is any first-order formula with free variable z, and x is any set, then

$$\forall z \in x\big(\varphi(z)\big) \quad :\Longleftrightarrow \quad \forall z\big((z \in x) \to \varphi(z)\big),$$

and similarly

$$\exists z \in x\big(\varphi(z)\big) \quad :\Longleftrightarrow \quad \exists z\big((z \in x) \wedge \varphi(z)\big).$$

As a first application of the Axiom Schema of Separation we define the *intersection* of two sets x_0 and x_1: We use x_0 as a parameter and let $\varphi(z, x_0)$ be the formula $z \in x_0$, denoted $\varphi(z, x_0) \equiv z \in x_0$. Then, by the Axiom Schema of Separation, there exists a set $y = \{z \in x_1 : \varphi(z, x_0)\}$, *i.e.*,

$$z \in y \leftrightarrow (z \in x_1 \wedge z \in x_0)\,.$$

In other words, for any sets x_0 and x_1, the collection of all sets which belong to both x_0 and x_1 is a set. This set is called the **intersection** of x_0 and x_1 and is denoted by $x_0 \cap x_1$. In general, for non-empty sets x we define

$$\bigcap x \;:= \big\{z \in \bigcup x : \forall y \in x\,(z \in y)\big\}\,,$$

which is the intersection of all sets which belong to x. In order to see that $\bigcap x$ is a set, let $\varphi(z, x) \equiv \forall y \in x\,(z \in y)$ and apply the Axiom Schema of Separation to $\bigcup x$. Notice also that $x \cap y = \bigcap\{x, y\}$. Furthermore, for $x = \{y_\iota : \iota \in I\}$ we sometimes write $\bigcap_{\iota \in I} y_\iota$ instead of $\bigcap x$.

Another example is when $\varphi(z, y) \equiv z \notin y$, where y is a parameter. In this case, $\{z \in x : z \notin y\}$ is a set, denoted $x \setminus y$, which is called the **set-theoretic difference** of x and y.

Let us now turn back to ordinal numbers:

THEOREM 3.3.

(a) *If* $\alpha, \beta \in \Omega$, *then* $\alpha \in \beta$ *or* $\alpha = \beta$ *or* $\alpha \ni \beta$, *where these three cases are mutually exclusive.*

(b) *If* $\alpha \in \beta \in \Omega$, *then* $\alpha \in \Omega$.

(c) *If* $\alpha \in \Omega$, *then also* $(\alpha \cup \{\alpha\}) \in \Omega$.

(d) Ω *is transitive and is well-ordered by* $\in$. *More precisely,* Ω *is transitive, is ordered by* $\in$, *and every non-empty collection* $C \subseteq \Omega$ *has an* $\in$*-minimal element.*

Proof. (a) First, notice that by FACT 3.2 the three cases $\alpha \in \beta$, $\alpha = \beta$, $\alpha \ni \beta$, are mutually exclusive.

Let $\alpha, \beta \in \Omega$ be given. If $\alpha = \beta$, then we are done. So, let us assume that $\alpha \neq \beta$. Without loss of generality we may assume that $\alpha \setminus \beta \neq \emptyset$.

We first show that $\alpha \cap \beta$ is the $\in$-minimal element of $\alpha \setminus \beta$: Let γ be an $\in$-minimal element of $\alpha \setminus \beta$. Since α is transitive and $\gamma \in \alpha$, $\forall u(u \in \gamma \to u \in \alpha)$, and since γ is an $\in$-minimal element of $\alpha \setminus \beta$, $\forall u(u \in \gamma \to u \in \beta)$, which implies $\gamma \subseteq \alpha \cap \beta$. On the other hand, if there is a $w \in (\alpha \cap \beta) \setminus \gamma$, then, since α is ordered by $\in$ and $\gamma \neq w$ ($\gamma \notin \beta \ni w$), we must have $\gamma \in w$, and since β is transitive and $w \in \beta$, this implies that $\gamma \in \beta$, which contradicts the fact that $\gamma \in (\alpha \setminus \beta)$. Hence, $\gamma = \alpha \cap \beta$ is the $\in$-minimal element of $\alpha \setminus \beta$. Now, if also $\beta \setminus \alpha \neq \emptyset$, then we would find that $\alpha \cap \beta$ is also the $\in$-minimal element of $\beta \setminus \alpha$, which is obviously a contradiction.

Thus, $\alpha \setminus \beta \neq \emptyset$ implies that $\beta \setminus \alpha = \emptyset$, or in other words, $\beta \subseteq \alpha$, which is the same as saying $\beta = \alpha \cap \beta$. Consequently we see that β is the $\in$-minimal element of $\alpha \setminus \beta$, in particular, $\beta \in \alpha$.

(b) Let $\alpha \in \beta \in \Omega$. Since β is transitive, α is ordered by $\in$. So, it remains to show that α is transitive and well-ordered by $\in$.

well-ordered by $\in$: Because β is transitive, every subset of α is also a subset of β and consequently contains an $\in$-minimal element.

transitive: Let $\delta \in \gamma \in \alpha$. We have to show that $\delta \in \alpha$. Since β is transitive, $\delta \in \beta$, and since β is ordered by $\in$, we have either $\delta \in \alpha$ or $\delta = \alpha$ or $\alpha \in \delta$. If $\delta \in \alpha$, we are done, and if $\delta = \alpha$ or $\alpha \in \delta$, then the set $\{\alpha, \gamma, \delta\} \subseteq \beta$ does not have an $\in$-minimal element, which contradicts the fact that β is well-ordered by $\in$.

(c) We have to show that $\alpha \cup \{\alpha\}$ is transitive and well-ordered by $\in$.

transitive: If $\beta \in (\alpha \cup \{\alpha\})$, then either $\beta \in \alpha$ or $\beta = \alpha$, and in both cases we have $\beta \subseteq (\alpha \cup \{\alpha\})$.

well-ordered by $\in$: Since α is an ordinal, $\alpha \cup \{\alpha\}$ is ordered by $\in$. Let now $x \subseteq (\alpha \cup \{\alpha\})$ be a non-empty set. If $x = \{\alpha\}$, then α is obviously an $\in$-minimal element of x. Otherwise, $x \cap \alpha \neq \emptyset$, and since $\alpha \in \Omega$, $x \cap \alpha$ has an $\in$-minimal element, say γ. Since α is transitive we have $x \cap \gamma = \emptyset$ (otherwise, γ would not be $\in$-minimal in $x \cap \alpha$), which implies that γ is $\in$-minimal in x.

(d) Ω *is transitive and ordered by* $\in$: This follows immediately from part (b) and part (a), respectively.

Ω *is well-ordered by* $\in$: Let $C \subseteq \Omega$ be a non-empty collection of ordinals. If $C = \{\alpha\}$ for some $\alpha \in \Omega$, then α is the $\in$-minimal element of C. Otherwise, C contains an ordinal δ_0 such that $\delta_0 \cap C \neq \emptyset$ and let $x := \delta_0 \cap C$. Then x is a non-empty *set* of ordinals. Now, let $\alpha \in x$ and let γ be an $\in$-minimal element of $x \cap (\alpha \cup \{\alpha\})$. By definition, $\gamma \in (\alpha \cup \{\alpha\})$, and since $(\alpha \cup \{\alpha\}) \in \Omega$, $\gamma \subseteq (\alpha \cup \{\alpha\})$. Thus, every ordinal $\gamma' \in \gamma$ belongs to $\alpha \cup \{\alpha\}$, but by the definition of γ, γ' cannot belong to $x \cap (\alpha \cup \{\alpha\})$, which implies that γ is also $\in$-minimal in x, and consequently in C. $\dashv$

By THEOREM 3.3 (d) we find that Ω is transitive and well-ordered by $\in$. Thus, if Ω were a set, Ω would be an ordinal number and therefore would belong to itself, but this is a contradiction to FACT 3.2.

In general, a collection of sets, satisfying for example a certain formula, which is not necessarily a set is called a **class**. For example, Ω is a class which is *not* a set (it consists of all transitive sets which are well-ordered by $\in$). Even though proper classes (*i.e.*, classes which are not sets) do not belong to the set-theoretic universe, it is sometimes convenient to handle them like sets, *e.g.*, taking intersections or extracting certain subsets or subclasses from them.

By THEOREM 3.3 (c) we know that if $\alpha \in \Omega$, then also $(\alpha \cup \{\alpha\}) \in \Omega$. Now, for ordinals $\alpha \in \Omega$ let $\alpha + 1 := \alpha \cup \{\alpha\}$. Part (a) of the following result—which is just a consequence of THEOREM 3.3—motivates this notation.

COROLLARY 3.4.

(a) *If* $\alpha, \beta \in \Omega$ *and* $\alpha \in \beta$, *then* $\alpha + 1 \subseteq \beta$. *In other words,* $\alpha + 1$ *is the least ordinal which contains* α.

(b) *For every ordinal* $\alpha \in \Omega$ *we have either* $\alpha = \bigcup \alpha$ *or there exists a* $\beta \in \Omega$ *such that* $\alpha = \beta + 1$.

Proof. (a) Assume $\alpha \in \beta$, then $\{\alpha\} \subseteq \beta$, and since β is transitive, we also have $\alpha \subseteq \beta$; thus, $\alpha + 1 = \alpha \cup \{\alpha\} \subseteq \beta$.

(b) Since α is transitive, $\bigcup \alpha \subseteq \alpha$. Thus, if $\alpha \neq \bigcup \alpha$, then $\alpha \setminus \bigcup \alpha \neq \emptyset$. Let β be $\in$-minimal in $\alpha \setminus \bigcup \alpha$. Then $\beta \in \alpha$ and $\beta + 1 \in \Omega$, and by part (a) we have $\beta + 1 \subseteq \alpha$. On the one hand, $\alpha \in \beta + 1$ would imply that $\alpha \in \alpha$, a contradiction to FACT 3.2. On the other hand, $\beta + 1 \in \alpha$ would imply that $\beta \in \bigcup \alpha$, which contradicts the choice of β. Thus, we must have $\beta + 1 = \alpha$. $\dashv$

This leads to the following definitions: An ordinal α is called a **successor ordinal** if there exists an ordinal β such that $\alpha = \beta + 1$; otherwise, it is called a **limit ordinal**. In particular, $\emptyset$ is a limit ordinal.

We are now going to construct the smallest inductive set containing the empty set, which will turn out to be the least non-empty limit ordinal. By the Axiom of Infinity we know that there exists an inductive set I with $\emptyset \in I$. In order to isolate the smallest inductive set containing $\emptyset$, we consider first the set $I_\Omega = I \cap \Omega$. More precisely, let

$$I_\Omega = \{\alpha \in I : \alpha \text{ is an ordinal}\}.$$

Then I_Ω is a set of ordinals and by THEOREM 3.3 (c), I_Ω is even an inductive set. Now, let $C := \Omega \setminus I_\Omega$; then C is a proper class of ordinals. Thus, by THEOREM 3.3 (d), C contains an $\in$-minimal ordinal, say λ_0. Since every $\alpha \in \lambda_0$ belongs to I_Ω and I_Ω is inductive, λ_0 must be inductive. Furthermore, since $\emptyset \in I_\Omega$ (*i.e.*, $\emptyset \notin C$), $\lambda_0 \neq \emptyset$, and since λ_0 is inductive, λ_0 must be a non-empty limit ordinal. Finally, let $\Lambda := \big\{\alpha \in \lambda_0 : \alpha \neq \emptyset \wedge \bigcup \alpha = \alpha\big\}$. If $\Lambda = \emptyset$ let $\omega := \lambda_0$; otherwise, let ω be the $\in$-minimal element of $\Omega \setminus \Lambda$. In both cases, ω is the least non-empty limit ordinal, *i.e.*, $\bigcup \omega = \omega$. In particular, for each $\alpha \in \omega$ we have $\alpha + 1 \in \omega$, which shows that ω is inductive. Furthermore, because ω is the smallest limit ordinal besides $\emptyset$, each ordinal $\alpha \in \omega$ except $\emptyset$ is a successor.

The ordinals belonging to ω are called **natural numbers**. Since ω is the smallest non-empty limit ordinal, all natural numbers, except 0, are successor ordinals. Thus, for each $n \in \omega$ we have either $n = 0$ or there is an $m \in \omega$ such that $n = m + 1$. Furthermore, if we define the binary ordering relation "$<$" on ω by stipulating

$$k < n \;:\Longleftrightarrow\; k \in n$$

then for each $n \in \omega$ we find $n = \{k \in \omega : k < n\}$.

The following theorem is a consequence of the fact that Ω is transitive and well-ordered by $\in$ (which is just THEOREM 3.3 (d)).

THEOREM 3.5 (TRANSFINITE INDUCTION THEOREM). *Let $C \subseteq \Omega$ be a class of ordinals and assume that:*

(a) *if $\alpha \in C$, then $\alpha + 1 \in C$,*

(b) *if α is a limit ordinal and $\forall \beta \in \alpha(\beta \in C)$, then $\alpha \in C$.*

Then C is the class of all ordinals.

Proof. Notice first that by (b) we have $0 \in C$ which shows that $C \neq \emptyset$.

Assume towards a contradiction that $C \neq \Omega$ and let α_0 be the $\in$-minimal ordinal which does not belong to C (such an ordinal exists by THEOREM 3.3 (d)). Now, α_0 can be *neither* a successor ordinal, since this would contradict (a), *nor* a limit ordinal, since this would contradict (b). Thus, α_0 does not exist, which implies that $\Omega \setminus C = \emptyset$, *i.e.*, $C = \Omega$. ⊣

The following result is just a reformulation of the TRANSFINITE INDUCTION THEOREM.

COROLLARY 3.6. *For any first-order formula $\varphi(x)$ with free variable x we have*

$$\forall\alpha \in \Omega\big(\forall\beta \in \alpha\big(\varphi(\beta)\big) \to \varphi(\alpha)\big) \to \forall\alpha \in \Omega\big(\varphi(\alpha)\big).$$

Proof. Let $C \subseteq \Omega$ be the class of all ordinals $\alpha \in \Omega$ such that $\varphi(\alpha)$ holds and apply the TRANSFINITE INDUCTION THEOREM 3.5. ⊣

When some form of COROLLARY 3.6 is involved we usually do not mention the corresponding formula φ and just say "by induction on ..." or "by transfinite induction".

6. The Axiom of Power Set

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

Informally, the Axiom of Power Set states that for each set x there is a set $\mathscr{P}(x)$, called the **power set** of x, which consists of all subsets of x.

With the Axiom of Power Set (and other axioms like the Axiom of Union or the Axiom Schema of Separation) we can now define notions like functions, relations, and sequences: For arbitrary sets A and B we define the **Cartesian product** $A \times B$ by stipulating

$$A \times B := \big\{\langle x, y\rangle : x \in A \wedge y \in B\big\}$$

where $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$. Thus, the Cartesian product of two sets A and B is a subsets of $\mathscr{P}(\mathscr{P}(A \cup B))$. Further, let

$${}^{A}B \ := \big\{f \subseteq A \times B : \forall x \in A\, \exists! y \in B\big(\langle x, y\rangle \in f\big)\big\}.$$

An element $f \in {}^{A}B$, usually denoted by $f : A \to B$, is called a **function** or **mapping** from A to B, where A is called the **domain** of f, denoted $\operatorname{dom}(f)$.

For $f : A \to B$ we usually write $f(x) = y$ instead of $\langle x, y\rangle \in f$. If S is a set, then the **image** of S under f is denoted by $f[S] = \{f(x) : x \in S\}$ and $f|_S = \{\langle x, y\rangle \in f : x \in S\}$ is the restriction of f to S. Furthermore, for a function $f : A \to B$, $f[A]$ is called the **range** of f, denoted $\operatorname{ran}(f)$.

A function $f : A \to B$ is **surjective**, or **onto**, if $\forall y \in B\, \exists x \in A(f(x) = y)$. We sometimes emphasise the fact that f is surjective by writing $f : A \twoheadrightarrow B$.

A function $f : A \to B$ is **injective**, also called **one-to-one**, if we have

$$\forall x_1 \in A\, \forall x_2 \in A\, \big(f(x_1) = f(x_2) \to x_1 = x_2\big).$$

To emphasise the fact that f is injective we sometimes write $f : A \hookrightarrow B$.

A function $f : A \to B$ is **bijective** if it is injective and surjective. If $f : A \to B$ is bijective, then

$$\forall y \in B\, \exists! x \in A\, \big(\langle x, y\rangle \in f\big)$$

which implies that

$$f^{-1} := \big\{\langle y, x\rangle : \langle x, y\rangle \in f\big\} \in {}^{B}A$$

is a function which is also bijective. So, if there is a bijective function from A to B, then there is also one from B to A and we sometimes just say that there is a **bijection between** A and B. Notice that if $f : A \hookrightarrow B$ is injective, then f is a bijection between A and $f[A]$.

Now, we are ready to define the notion of finiteness and countability: A set A is called **finite** if there exists a surjection from a natural number $n \in \omega$ onto A, otherwise, A is called **infinite**. In particular, each natural number is finite and ω is the smallest infinite (*i.e.*, not finite) ordinal number. Furthermore, a set C is called **countable** if there exists a surjection from ω onto C, otherwise, C is called **uncountable**.

Let us turn back to Cartesian products: Assume that for each $\iota \in I$ (for some set I) we have assigned a non-empty set A_ι. For $A = \bigcup_{\iota \in I} A_\iota$, where we define $\bigcup_{\iota \in I} A_\iota := \bigcup\{A_\iota : \iota \in I\}$, the set

$$\bigtimes_{\iota \in I} A_\iota = \Big\{f \in {}^{I}A : \forall \iota \in I\big(f(\iota) \in A_\iota\big)\Big\}$$

is called the Cartesian product of the sets A_ι ($\iota \in I$). Notice that if all sets A_ι are equal to a given set A, then $\bigtimes_{\iota \in I} A_\iota = {}^{I}A$. If $I = n$ for some $n \in \omega$, in abuse of notation we also write A^n instead of ${}^{n}A$ by identifying ${}^{n}A$ with the set

$$A^n = \underbrace{A \times \ldots \times A}_{n\text{-times}}.$$

Closely related to functions is the notion of sequence: For $\alpha \in \Omega$ we can identify a function $f \in {}^{\alpha}A$ with the **sequence** $\langle f(0), f(1), \ldots, f(\beta), \ldots\rangle_\alpha$ of length α, and vice versa. Sequences (of length α) are usually denoted using angled brackets (and by using α as a subscript), *e.g.*, $\langle s_0, \ldots, s_\beta, \ldots\rangle_\alpha$ or $\langle s_\beta : \beta < \alpha\rangle$. In general, for any set A, let $\operatorname{seq}(A)$ be the set of all finite sequences which can be formed with elements of A, or more formally:

$$\operatorname{seq}(A) := \bigcup_{n \in \omega} {}^{n}A$$

Furthermore, let $\operatorname{seq}^{1\text{-}1}(A)$ be those sequences of $\operatorname{seq}(A)$ in which no element appears twice. Again more formally, this reads as follows:

$$\operatorname{seq}^{1\text{-}1}(A) := \Big\{\sigma \in \bigcup_{n \in \omega} {}^{n}A : \sigma \text{ is injective}\Big\}$$

Similar to finite sequences, we can also define finite subsets: For any $n \in \omega$ and any set S, let $[S]^n$ denote the set of all n-element subsets of S (*e.g.*, $[S]^0 = \{\emptyset\}$). Slightly more formally, for $n \in \omega$ and an arbitrary set S,

$$[S]^n := \big\{x \in \mathscr{P}(S) : \textsf{there exists a bijection between } x \textsf{ and } n\big\}\,.$$

Further, the set of all finite subsets of a set S is denoted by $[S]^{<\omega}$ or just $\operatorname{fin}(S)$. In other words,

$$\operatorname{fin}(S) := \bigcup_{n\in\omega} [S]^n.$$

Finally, for any set S, $[S]^\omega$ denotes the set of all countably infinite subsets of S, in particular, since every infinite subset of ω is countable, $[\omega]^\omega$ is the set of *all* infinite subsets of ω.

Let us turn back again to Cartesian products, or more precisely, to subsets of finite Cartesian products:

- For any set A and any $n \in \omega$, a set $R \subseteq A^n$ is called an **n-ary relation** on A.
- If $n = 2$, then $R \subseteq A \times A$ is also called a **binary relation**. For binary relations R we usually write xRy instead of $\langle x, y\rangle \in R$.
- A binary relation R on A is a **linear ordering** on A, if for any elements $x, y \in A$ we have

  $$\textsf{either} \quad xRy \quad \textsf{or} \quad x = y \quad \textsf{or} \quad yRx$$

  where these three cases are mutually exclusive.
- A linear ordering R on A is a **well-ordering** on A if every non-empty subset $S \subseteq A$ has an R-minimal element, *i.e.*, there exists a $x_0 \in S$ such that for each $y \in S$ we have x_0Ry. Notice, that since R is a linear ordering, the R-minimal element x_0 is unique.
- If there is a well-ordering R on A, then we say that A is **well-orderable**. The discussion of whether each set is well-orderable has to be postponed until we have the Axiom of Choice.

Other important binary relations are the so-called equivalence relations: Let S be an arbitrary non-empty set. A binary relation "$\sim$" on S is an **equivalence relation** if it is

- *reflexive* (*i.e.*, for all $x \in S$: $x \sim x$),
- *symmetric* (*i.e.*, for all $x, y \in S$: $x \sim y \leftrightarrow y \sim x$), and
- *transitive* (*i.e.*, for all $x, y, z \in S$: $x \sim y \wedge y \sim z \ \rightarrow \ x \sim z$).

The **equivalence class** of an element $x \in S$, denoted $[x]\tilde{}$, is the set $\{y \in S : x \sim y\}$. We would like to recall the fact that for any $x, y \in S$ we have *either* $[x]\tilde{} = [y]\tilde{}$ *or* $[x]\tilde{} \cap [y]\tilde{} = \emptyset$. A set $A \subseteq S$ is a set of **representatives** if for each equivalence class

$[x]\tilde{}$, A has exactly one element in common with each equivalence class. We would like to mention that the existence of a set of representatives relies in general on the Axiom of Choice.

With the axioms we have so far, we could already construct a model of Peano Arithmetic. However, we postpone this construction and proceed with the axioms of Zermelo–Fraenkel Set Theory. The next axiom we present is Fraenkel's axiom schema of replacement, which allows us to build, for example, sets like $\{\mathscr{P}^n(\omega) : n \in \omega\}$, where $\mathscr{P}^0(\omega) := \omega$ and $\mathscr{P}^{n+1}(\omega) := \mathscr{P}(\mathscr{P}^n(\omega))$.

7. The Axiom Schema of Replacement

For every first-order formula $\varphi(x, y, p)$ with $\mathrm{free}(\varphi) \subseteq \{x, y, p\}$, where p can be an ordered n-tuple of parameters, the following formula is an axiom:

$$\forall A\, \forall p\big(\forall x \in A\, \exists! y\, \varphi(x, y, p) \rightarrow \exists B\, \forall x \in A\, \exists y \in B\, \varphi(x, y, p)\big)$$

In other words, for every set A and for each class function F (*i.e.*, a certain class of ordered pairs of sets) defined on A, $F[A] = \{F(x) : x \in A\}$ is a set. Or even more informally, images of sets under functions are sets.

As a first application of the Axiom Schema of Replacement we prove a result which will be used, for example, to define ordinal addition (see THEOREM 3.8) or to build the cumulative hierarchy of sets (see THEOREM 3.12).

THEOREM 3.7 (TRANSFINITE RECURSION THEOREM). *Let F be a class function which is defined for all sets. Then there is a unique class function G defined on Ω such that for each $\alpha \in \Omega$ we have*

$$G(\alpha) = F(G|_\alpha), \quad \text{where } G|_\alpha = \big\{\big\langle \beta, G(\beta)\big\rangle : \beta \in \alpha\big\}.$$

Proof. If such a class function G exists, then, by the Axiom Schema of Replacement, for *every* ordinal α, $\mathrm{ran}(G|_\alpha)$ is a set, and consequently, $G|_\alpha$ is a function with $\mathrm{dom}(G|_\alpha) = \alpha$. This leads to the following definition: For $\delta \in \Omega$, a function g with $\mathrm{dom}(g) = \delta$ is called a δ-*approximation* if

$$\forall \beta \in \delta\big(g(\beta) = F(g|_\beta)\big).$$

In other words, g is an δ-approximation if and only if g has the following properties:

(a) If $\beta + 1 \in \delta$, then $g(\beta + 1) = F(g|_\beta \cup \{\langle \beta, g(\beta)\rangle\})$.

(b) If $\beta \in \delta$ is a limit ordinal, then $g(\beta) = F(g|_\beta)$.

In particular, by (b) we get $g(0) = F(\emptyset)$. For example $g_1 = \{\langle 0, F(\emptyset)\rangle\}$ is a 1-approximation; in fact, g_1 is the unique 1-approximation. Similarly,

$$g_2 = \{\langle 0, F(\emptyset)\rangle, \langle 1, F(\{\langle 0, F(\emptyset)\rangle\})\rangle\}$$

is the unique 2-approximation.

First, notice that for all ordinals δ and δ', if g is a δ-approximation and g' is a δ'-approximation, then $g|_{\delta\cap\delta'} = g'|_{\delta\cap\delta'}$. Otherwise, there would be a smallest ordinal β_0 such that $g(\beta_0) \neq g'(\beta_0)$, but by (a) and (b), β_0 would be neither a successor ordinal nor a limit ordinal.

Now we show that for each ordinal δ there exists a δ-approximation: Otherwise, by THEOREM 3.3 (d), there would be a smallest ordinal δ_0 such that there is no δ_0-approximation. In particular, for each $\delta \in \delta_0$ there would be a δ-approximation, *i.e.*,

$$\forall\delta \in \delta_0 \, \exists! d \, (\text{“}d \textit{ is a } \delta\textit{-approximation”}) .$$

Hence, by the Axiom Schema of Replacement,

$$\exists D \, \forall\delta \in \delta_0 \, \exists d \in D \, (\text{“}d \textit{ is a } \delta\textit{-approximation”}) ,$$

and $\bigcup D$ is a δ'-approximation for some $\delta' \in \Omega$. Now, if δ_0 is a limit ordinal, then $\delta' = \delta_0$ and we get a δ_0-approximation, and if δ_0 is a successor ordinal, then $\delta_0 = \delta' + 1$ and we get a δ_0-approximation by (a). So, in both cases we get a contradiction to our assumption that there is no δ_0-approximation.

Now, for each $\alpha \in \Omega$ define $G(\alpha) := g(\alpha)$, where g is the δ-approximation for any δ such that $\alpha \in \delta$. ⊣

When the TRANSFINITE RECURSION THEOREM is involved in some construction or proof, we usually just say “by transfinite recursion...” without defining the corresponding class function F.

As a first application of the TRANSFINITE RECURSION THEOREM we show how we can define addition, multiplication, and exponentiation of arbitrary ordinal numbers:

ORDINAL ADDITION. For arbitrary ordinals $\alpha \in \Omega$ we define:

(a) $\alpha + 0 := \alpha$.

(b) $\alpha + (\beta + 1) := (\alpha + \beta) + 1$, for all $\beta \in \Omega$.

(c) If $\beta \in \Omega$ is non-empty and a limit ordinal, then $\alpha + \beta := \bigcup_{\delta\in\beta}(\alpha + \delta)$.

Notice that addition of ordinals is in general not commutative (*e.g.*, $1 + \omega = \omega$ but $\omega + 1 \neq \omega$).

ORDINAL MULTIPLICATION. For arbitrary ordinals $\alpha \in \Omega$ we define:

(a) $\alpha \cdot 0 := 0$.

(b) $\alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha$, for all $\beta \in \Omega$.

(c) If $\beta \in \Omega$ is a limit ordinal, then $\alpha \cdot \beta := \bigcup_{\delta \in \beta}(\alpha \cdot \delta)$.

Notice that multiplication of ordinals is in general not commutative (*e.g.*, $2 \cdot \omega = \omega$ but $\omega \cdot 2 = \omega + \omega \neq \omega$).

ORDINAL EXPONENTIATION. For arbitrary ordinals $\alpha \in \Omega$ we define:

(a) $\alpha^0 := 1$.

(b) $\alpha^{\beta+1} := \alpha^\beta \cdot \alpha$, for all $\beta \in \Omega$.

(c) If $\beta \in \Omega$ is non-empty and a limit ordinal, then $\alpha^\beta := \bigcup_{\delta \in \beta} \alpha^\delta$.

Notice that, for example, $2^\omega = \omega$. This should not be confused with *cardinal exponentiation*, which will be defined later in this chapter.

THEOREM 3.8. *Addition, multiplication, and exponentiation of ordinals are well-defined binary operations on* Ω.

Proof. We will prove this result for addition (the proof for the other operations is similar): For each $\alpha \in \Omega$ define a class function F_α by stipulating $F_\alpha(x) := \emptyset$ if x is *not* a function; if x *is* a function, then let

$$F_\alpha(x) = \begin{cases} \alpha & \text{if } x = \emptyset, \\ x(\beta) \cup \{x(\beta)\} & \text{if } \operatorname{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega, \\ \bigcup_{\delta \in \beta} x(\delta) & \text{if } \operatorname{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal,} \\ \emptyset & \text{otherwise.} \end{cases}$$

By the TRANSFINITE RECURSION THEOREM 3.7, for each $\alpha \in \Omega$ there is a unique class function G_α defined on Ω such that for each $\beta \in \Omega$ we have $G_\alpha(\beta) = F_\alpha(G_\alpha|_\beta)$.

By definition of F_α, $G_\alpha(\emptyset) = F_\alpha(\emptyset) = \alpha$, which shows that $G_\alpha(\emptyset) = \alpha + \emptyset$. If, for some $\beta \in \Omega$, we already have $G_\alpha(\beta) = \alpha + \beta$, then, by definition of F_α,

$$G_\alpha(\beta + 1) = F_\alpha(G_\alpha|_{\beta+1}) = \alpha + \beta \cup \{\alpha + \beta\},$$

which shows that $G_\alpha(\beta + 1) = (\alpha + \beta) + 1$, hence, by part (b) of the definition of ordinal addition, we get $G_\alpha(\beta + 1) = \alpha + (\beta + 1)$. Finally, if $\beta \in \Omega$ is a non-empty limit ordinal and for all $\delta \in \beta$ we have $G_\alpha(\delta) = \alpha + \delta$, then, again by definition of F_α,

$$G_\alpha(\beta) = F_\alpha(G_\alpha|_\beta) = \bigcup_{\delta \in \beta} G_\alpha(\delta),$$

which shows that $G_\alpha(\beta) = \bigcup_{\delta \in \beta} G_\alpha(\delta) = \alpha + \beta$. Thus, for each $\beta \in \Omega$ we get $G_\alpha(\beta) = \alpha + \beta$. $\dashv$

Notice that even though addition and multiplication of ordinals are not commutative, they are still associative.

PROPOSITION 3.9. *Addition and multiplication of ordinals defined as above are associative operations.*

Proof. We have to show that for all $\alpha, \beta, \gamma \in \Omega$, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ and $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$. We give the proof for addition and leave the proof for multiplication as an exercise to the reader.

Let α and β be arbitrary ordinals. The proof is by induction on $\gamma \in \Omega$. For $\gamma = 0$ we obviously have $(\alpha + \beta) + 0 = \alpha + \beta = \alpha + (\beta + 0)$. Now, let us assume that $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for some γ. Then:

$$\begin{aligned}(\alpha + \beta) + (\gamma + 1) &= \big((\alpha + \beta) + \gamma\big) + 1 && \text{(by definition of “ + ”)}\\ &= \big(\alpha + (\beta + \gamma)\big) + 1 && \text{(by our assumption)}\\ &= \alpha + \big((\beta + \gamma) + 1\big) && \text{(by definition of “ + ”)}\\ &= \alpha + \big(\beta + (\gamma + 1)\big) && \text{(by definition of “ + ”)}\end{aligned}$$

Finally, let γ be a limit ordinal. Notice first that $\beta + \gamma$ is a limit ordinal and that

$$\alpha + (\beta + \gamma) \;=\; \bigcup_{\delta \in (\beta+\gamma)} \alpha + \delta \;=\; \bigcup_{(\beta+\gamma') \in (\beta+\gamma)} \alpha + (\beta + \gamma') \;=\; \bigcup_{\gamma' \in \gamma} \alpha + (\beta + \gamma')\,.$$

Thus, if $(\alpha + \beta) + \gamma' = \alpha + (\beta + \gamma')$ for all $\gamma' \in \gamma$, then

$$(\alpha + \beta) + \gamma = \bigcup_{\gamma' \in \gamma} (\alpha + \beta) + \gamma' = \bigcup_{\gamma' \in \gamma} \alpha + (\beta + \gamma') = \alpha + (\beta + \gamma).$$

⊣

As another application of the TRANSFINITE RECURSION THEOREM we introduce the notion of the *transitive closure* of a set S, which is the smallest transitive set which contains S. Informally, we could define the transitive closure $\mathrm{TC}(S)$ of S by stipulating

$$\mathrm{TC}(S) = \bigcap \big\{T : T \supseteq S \textit{ and } T \textit{ is transitive}\big\}\,,$$

but we will do it in a more explicit way: Let S be an arbitrary set. Define the class function F_S by stipulating $F_S(x) := \emptyset$ if x is *not* a function; and if x *is* a function, then let

$$F_S(x) = \begin{cases} S & \text{if } x = \emptyset,\\ \bigcup x(\beta) & \text{if } \mathrm{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega,\\ \bigcup_{\delta \in \beta} x(\delta) & \text{if } \mathrm{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal},\\ \emptyset & \text{otherwise.}\end{cases}$$

By the TRANSFINITE RECURSION THEOREM 3.7, there is a unique class function G_S defined on Ω such that for each $\beta \in \Omega$ we have $G_S(\beta) = F_S(G_S|_\beta)$. In particular, we have $G_S(\emptyset) = S$ and for $n \in \omega$ we get $G_S(n + 1) = \bigcup G_S(n)$.

PROPOSITION 3.10. *The set $G_S(\omega)$ is the smallest transitive set which contains S, i.e., $G_S(\omega) = \mathrm{TC}(S)$.*

Proof. To simplify the notation, for each $n \in \omega$ let $S_n = G_S(n)$ and let $\bar{S} := \bigcup_{n\in\omega} S_n$. Then $S_0 = S$, $S_{n+1} = \bigcup S_n$, and $G_S(\omega) = \bar{S}$. We have to show that $\bar{S}$ is the smallest transitive set which contains S.

Since $S = S_0$, S is contained in $\bar{S}$. Furthermore, if $x \in y \in \bar{S}$, then there is an $n \in \omega$ such that $y \in S_n$, which implies that $x \in S_{n+1}$. Hence, $x \in \bar{S}$, which shows that $\bar{S}$ is transitive. It remains to show that $\bar{S}$ is the smallest transitive set containing S. For this, choose a proper transitive subset $T \subsetneq \bar{S}$ and let $x_0 \in \bar{S} \setminus T$. Since $x_0 \in \bar{S}$, there is an $n \in \omega$ such that $x_0 \in S_n$; let

$$n_0 := \bigcap \big\{n \in \omega : x_0 \in S_n\big\} .$$

If $n_0 = 0$, then $x_0 \in S$ which implies that S is not a subset of T. Otherwise, let $Y_0 = \{x_0\}$. Since T is transitive and $x_0 \notin T$, $Y_0 \cap T = \emptyset$. By induction on ω, starting with n_0 and Y_0, we define sequences $\langle n_k : k \in \omega\rangle$ and $\langle Y_k : k \in \omega\rangle$ as follows: If $n_k = 0$, then $n_{k+1} := 0$ and $Y_{k+1} := Y_k$. Otherwise,

$$Y_{k+1} := \big\{y \in \bar{S} : \exists x \in Y_k (x \in y)\big\}$$

and

$$n_{k+1} = \bigcap \big\{n \in \omega : \exists y \in Y_{k+1}(y \in S_n)\big\} .$$

By construction we get $n_{k+1} \in n_k$ (in fact $n_{k+1} + 1 = n_k$) and $Y_{k+1} \neq \emptyset$. Furthermore, since T is transitive, by construction we get $Y_k \cap T = \emptyset$ implies $Y_{k+1} \cap T = \emptyset$. Thus, by induction we get $Y_k \cap T = \emptyset$ for all $k \in \omega$. Consider now the set $N = \{n_k : k \in \omega\} \subseteq \omega$. Since N is non-empty and ω is well-ordered by $\in$, there exists an $\in$-minimal element n_{k_0} in N, and by construction, $n_{k_0} = 0$. For the corresponding set Y_{k_0} we get that Y_{k_0} is a non-empty subset of S which is disjoint from T, and hence, $S \nsubseteq T$. $\dashv$

8. The Axiom of Foundation

$$\forall x\big(x \neq \emptyset \to \exists y \in x(y \cap x = \emptyset)\big)$$

As a consequence of the Axiom of Foundation we see that there is no infinite descending sequence $x_0 \ni x_1 \ni x_2 \ni \cdots$ since otherwise, the set $\{x_0, x_1, x_2, \ldots\}$ would contradict the Axiom of Foundation: To see this, assume there is a set $x = \{x_n : n \in \omega\}$ such that for each $n \in \omega$ we have $x_n \in x_{n+1}$. Then, for each $y \in x$ we find an $n \in \omega$ such that $y = x_n$, which implies $x_{n+1} \in y \cap x$. So, for each $y \in x$, $y \cap x \neq \emptyset$, which contradicts the Axiom of Foundation. Moreover, there is no set x such that $x \in x$ (consider the set $\{x\}$) and there are also no cycles like $x_0 \in x_1 \in \cdots \in x_n \in x_0$ (consider the set $\{x_0, \ldots, x_n\}$). As a matter of fact, we would like to mention that if one assumes the Axiom of Choice, then the non-existence of infinite descending sequences $x_0 \ni x_1 \ni x_2 \ni \cdots$ can be proved to be equivalent to the Axiom of Foundation.

The axiom system containing the axioms 0–8 is called **Zermelo–Fraenkel Set Theory** and is denoted by ZF. In fact, ZF contains all axioms of Set Theory except the Axiom of Choice.

Even though the Axiom of Foundation is irrelevant outside Set Theory, it is extremely useful in the metamathematics of Set Theory, since it allows us to arrange all sets in a cumulative hierarchy—as we shall do after discussing the consistency of Set Theory.

On the Consistency of ZF

Zermelo writes in [102, p. 262] that he was not able to show that the seven axioms for Set Theory given in that article are consistent. Even though it is essential to know whether a theory is consistent or not, we will see that there is no way to prove the consistency of ZF without the aid of some metamathematical assumptions.

By GÖDEL'S INCOMPLETENESS THEOREM 2.4 we know that for every consistent theory which is sufficiently strong to prove the axioms of Peano Arithmetic PA, there is a sentence σ which is independent of that theory. To apply this result for Set Theory, we first have to show that ZF is sufficiently strong to define the concept of natural numbers. In other words, we have to show that ZF is strong enough to provide a model $\mathbf{N}$ of PA.

We do this by constructing an $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbf{N}$ with domain ω, and show that $\mathbf{N}$ is a model of PA. Recall that $\mathscr{L}_{\mathsf{PA}} = \{\mathsf{0}, \mathsf{s}, +, \cdot\}$. The $\mathscr{L}_{\mathsf{PA}}$-structure is defined by the following assignments:

$$
\begin{array}{rlccc}
\mathsf{0}^{\mathbf{N}} := & \emptyset & & \\
\mathsf{s}^{\mathbf{N}} : & \omega & \to & \omega \\
 & n & \mapsto & n+1 \\
+^{\mathbf{N}} : & \omega \times \omega & \to & \omega \\
 & \langle n, m\rangle & \mapsto & n+m \\
\cdot^{\mathbf{N}} : & \omega \times \omega & \to & \omega \\
 & \langle n, m\rangle & \mapsto & n \cdot m
\end{array}
$$

Before we show that the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbf{N}$ is a model of Peano Arithmetic, we first recall the axioms of PA:

PA_0: $\neg\exists x(\mathtt{s}x = 0)$

PA_1: $\forall x\forall y(\mathtt{s}x = \mathtt{s}y \to x = y)$

PA_2: $\forall x(x + 0 = x)$

PA_3: $\forall x\forall y(x + \mathtt{s}y = \mathtt{s}(x + y))$

PA_4: $\forall x(x \cdot 0 = 0)$

PA_5: $\forall x\forall y(x \cdot \mathtt{s}y = (x \cdot y) + x)$

If φ is any $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \mathrm{free}(\varphi)$, then:

PA_6: $\big(\varphi(0) \land \forall x(\varphi(x) \to \varphi(\mathtt{s}(x)))\big) \to \forall x\varphi(x)$

Let us now show that $\mathbf{N} \vDash \mathsf{PA}$:

- PA_0: Since $\bigcup\emptyset = \emptyset$, by COROLLARY 3.4 (b) we get that there is no $n \in \omega$ such that $n + 1 = \emptyset$.

- PA_1: If $n, m \in \omega$ and $n \neq m$, then, by THEOREM 3.3 (a), we have either $n \in m$ or $m \in n$, and in both cases we get $n + 1 \neq m + 1$.

- PA_2 and PA_3: Follow immediately from (a) and (b) of ORDINAL ADDITION.

- PA_4 and PA_5: Follow immediately from (a) and (b) of ORDINAL MULTIPLICATION.

- PA_6: Let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \mathrm{free}(\varphi)$ and assume
$$\varphi(\emptyset) \land \forall n \in \omega\big(\varphi(n) \to \varphi(n+1)\big)\,.$$
Furthermore, let $E := \{n \in \omega : \neg\varphi(n)\}$. Obviously, E is a subset of ω. Assume towards a contradiction that $E \neq \emptyset$ and let m be the $\in$-minimal element of E. Now, m can *neither* be $\emptyset$, since we assumed $\varphi(\emptyset)$, *nor* a successor ordinal (*i.e.*, of the form $n + 1$), since we assumed $\varphi(n) \to \varphi(n + 1)$ which is equivalent to $\neg\varphi(n + 1) \to \neg\varphi(n)$. Thus, there is no $\in$-minimal element of E, which is only possible when E is empty. Now, $E = \emptyset$ implies that there is no $n \in \omega$ such that $\neg\varphi(n)$, *i.e.*, $\forall n \in \omega\big(\varphi(n)\big)$.

Since every model of ZF contains a model $\mathbf{N}$ of PA (with domain ω), by GÖDEL'S COMPLETENESS THEOREM 2.4 we get that the consistency of ZF implies the consistency of PA.

On the other hand, one can show that, without additional assumptions, no model of ZF contains a model of ZF: If we can prove from ZF that a certain set V_0 is the domain of a model of ZF, then we can prove in ZF that V_0 contains a set V_1 which is again the domain of a model of ZF. Proceeding this way, this would result in an infinite, strictly decreasing sequence $V_0 \ni V_1 \ni V_2 \ni \cdots$ of sets, which is a contradiction to the Axiom of Foundation (compare with RELATED RESULT 91). Thus, ZF cannot provide a model of ZF. In other words, there is no way to construct or to define a model of ZF without the aid of some concepts that go beyond what is provided in ordinary Mathematics. More formally, any proof of $\mathrm{Con}(\mathsf{ZF})$ has to be carried out in some extension of ZF, which contains some information that is not provable from ZF.

Even though ZF cannot provide a model of ZF, we know that if ZF is consistent, then it has a model—such a model will be constructed in the next section.

Models of ZF

Let us assume that ZF is consistent. Then, by GÖDEL'S COMPLETENESS THEOREM 2.4 we know that there must be a model $\mathbf{M}$ of ZF. Surprisingly, the model $\mathbf{M}$ can be easily described as a hierarchy of sets. For this, we construct within the model $\mathbf{M}$ a certain hierarchy of sets and show that every set in $\mathbf{M}$ belongs to this hierarchy.

First, we define the following sets:

$$\begin{aligned} \mathrm{V}_0 &:= \emptyset \\ \mathrm{V}_\alpha &:= \bigcup_{\beta\in\alpha} \mathrm{V}_\beta \quad \text{if } \alpha \text{ is a limit ordinal} \\ \mathrm{V}_{\alpha+1} &:= \mathscr{P}(\mathrm{V}_\alpha) \end{aligned}$$

and then we define the class $\mathbf{V}$ by stipulating

$$\mathbf{V} := \bigcup_{\alpha\in\Omega} \mathrm{V}_\alpha .$$

To carry out the construction in the framework of ZF, we define the class function F by stipulating $F(x) := \emptyset$ if x is *not* a function; and if x *is* a function, then let

$$F(x) = \begin{cases} \emptyset & \text{if } x = \emptyset, \\ \mathscr{P}\big(x(\beta)\big) & \text{if } \mathrm{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega, \\ \bigcup_{\delta\in\beta} x(\delta) & \text{if } \mathrm{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal}, \\ \emptyset & \text{otherwise.} \end{cases}$$

By the TRANSFINITE RECURSION THEOREM 3.7, there is a unique class function G defined on Ω such that for each $\alpha \in \Omega$ we have $G(\alpha) = F(G|_\alpha)$. In particular, for each $\alpha \in \Omega$ we have $G(\alpha) = \mathrm{V}_\alpha$.

Notice that by construction, or more precisely by the Axiom Schema of Replacement, for each $\alpha \in \Omega$, V_α is a *set*. Moreover, we can easily prove the following

FACT 3.11. *For any $\alpha, \beta \in \Omega$ we have:*

(a) V_α *is transitive.*

(b) *The class* $\mathbf{V}$ *is transitive.*

(c) *If* $\alpha \in \beta$, *then* $\mathrm{V}_\alpha \subsetneq \mathrm{V}_\beta$.

(d) $\alpha \subseteq \mathrm{V}_\alpha$ *and* $\alpha \in \mathrm{V}_{\alpha+1}$.

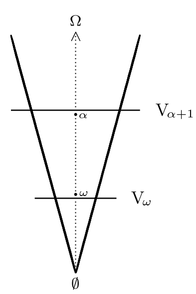
Proof. (a) First, notice that V_0 is transitive. Now, if α is a non-empty limit ordinal and for each $\beta \in \alpha$, V_β is transitive, then $\bigcup_{\beta\in\alpha} \mathrm{V}_\beta$ is obviously transitive, too. Finally, if $\alpha = \beta + 1$ and V_β is transitive, then $\mathrm{V}_\alpha = \mathscr{P}(\mathrm{V}_\beta)$, and for all sets x and y, such that $y \in x \in \mathrm{V}_\alpha$ we have $y \in x \subseteq \mathrm{V}_\beta$. Hence, $y \in \mathrm{V}_\beta$, and since V_β is transitive, we get $y \subseteq \mathrm{V}_\beta$ which shows that $y \in \mathrm{V}_\alpha$. Therefore, by transfinite induction we get that for each $\alpha \in \Omega$, V_α is transitive.

(b) Since arbitrary unions of transitive sets are transitive, this follows immediately from (a).

(c) If $\beta = \alpha + 1$, then, since V_α is transitive, $\mathrm{V}_\alpha \subseteq \mathrm{V}_\beta$, and since $\mathrm{V}_\alpha \in \mathrm{V}_\beta$ but $\mathrm{V}_\alpha \notin \mathrm{V}_\alpha$, we get $\mathrm{V}_\alpha \subsetneq \mathrm{V}_\beta$. Now, if β is a limit ordinal and $\alpha \in \beta$, then $\mathrm{V}_\alpha \subsetneq \mathrm{V}_{\alpha+1} \subsetneq \mathrm{V}_\beta$.

(d) If $\alpha = 0$, then we obviously have $\alpha \subseteq \mathrm{V}_\alpha$ and $\alpha \in \mathrm{V}_{\alpha+1}$. Now, let α_0 be an ordinal and assume that for all $\beta \in \alpha_0$ we have $\beta \subseteq \mathrm{V}_\beta$ and $\beta \in \mathrm{V}_{\beta+1}$. If α_0 is a limit ordinal, then the assumption implies $\alpha_0 \subseteq \mathrm{V}_{\alpha_0}$, and consequently we get $\alpha_0 \in \mathrm{V}_{\alpha_0+1}$. Finally, if $\alpha_0 = \beta + 1$, then, by the assumption, $\beta \in \mathrm{V}_{\alpha_0}$, and since V_{α_0} is transitive, $\alpha_0 \subseteq \mathrm{V}_{\alpha_0}$, and consequently $\alpha_0 \in \mathrm{V}_{\alpha_0+1}$. Therefore, by transfinite induction we get that for each $\alpha \in \Omega$, $\alpha \subseteq \mathrm{V}_\alpha$ and $\alpha \in \mathrm{V}_{\alpha+1}$. ⊣

The previous fact is visualised by the following figure:



Now we prove that the so-called **cumulative hierarchy** $\mathbf{V}$, which we have constructed within the model $\mathbf{M}$, contains *all* sets of $\mathbf{M}$, hence $\mathbf{M} = \mathbf{V}$.

THEOREM 3.12. *For every set x in $\mathbf{M}$ there is an ordinal α such that $x \in \mathrm{V}_\alpha$. In particular, every model $\mathbf{M}$ of ZF has the structure of a cumulative hierarchy $\mathbf{V}$.*

Proof. Assume towards a contradiction that there exists a set x in $\mathbf{M}$ which does not belong to $\mathbf{V}$. Let $\bar{x} := \mathrm{TC}(\{x\})$ (*i.e.*, $\bar{x}$ is the transitive closure of $\{x\}$), and let $w := \{z \in \bar{x} : z \notin \mathbf{V}\}$, *i.e.*, $w = \bar{x} \setminus \{z' \in \bar{x} : \exists \alpha \in \Omega\,(z' \in \mathrm{V}_\alpha)\}$. Notice that $w \in \mathbf{M}$. Since $x \in w$ we have $w \neq \emptyset$, and by the Axiom of Foundation there is a $z_0 \in w$ such that $(z_0 \cap w) = \emptyset$. Since $z_0 \in w$ we have $z_0 \notin \mathbf{V}$, which implies that $z_0 \neq \emptyset$, but for all $u \in z_0$ there is a least ordinal α_u such that $u \in \mathrm{V}_{\alpha_u}$. By the Axiom Schema of Replacement, $\{\alpha_u : u \in z_0\}$ is a set, and moreover, $\alpha = \bigcup\{\alpha_u : u \in z_0\} \in \Omega$. This implies that $z_0 \subseteq \mathrm{V}_\alpha$ and consequently we get $z_0 \in \mathrm{V}_{\alpha+1}$, which contradicts the fact that $z_0 \notin \mathbf{V}$ and completes the proof. $\dashv$

So, the set-theoretic universe $\mathbf{V}$ constructed above contains all sets; but does this imply that $\mathbf{V}$ is unique? Surprisingly, the answer is no, as we will show now.

First we extend the signature $\mathscr{L}_{\mathsf{ZF}} = \{\in\}$ by adding countably many new constant symbols $\mathsf{c}_0, \mathsf{c}_1, \mathsf{c}_2, \ldots$, *i.e.*, the new signature is $\{\in, \mathsf{c}_0, \mathsf{c}_1, \mathsf{c}_2, \ldots\}$. Now, we extend the axioms ZF by adding the formulae

$$\underbrace{\mathsf{c}_0 = \{\mathsf{c}_1\}}_{\varphi_0}, \quad \underbrace{\mathsf{c}_1 = \{\mathsf{c}_2\}}_{\varphi_1}, \quad \underbrace{\mathsf{c}_2 = \{\mathsf{c}_3\}}_{\varphi_2}, \quad \ldots$$

and let Ψ be the set of these formulae. Now, if ZF has a model $\mathbf{V}$ and Φ is any finite subset of Ψ, then, by interpreting the finitely many c_n's appearing in Ψ in a suitable way, $\mathbf{V}$ is also a model of $\mathsf{ZF} \cup \Phi$, which implies that $\mathsf{ZF} \cup \Phi$ is consistent. Thus, by the COMPACTNESS THEOREM 2.7, $\mathsf{ZF} \cup \Psi$ is also consistent and therefore has a model, say $\mathbf{V}^*$. Since $\mathbf{V}^* \vDash \mathsf{ZF} \cup \Psi$, we get that the Axiom of Foundation holds in $\mathbf{V}^*$. In particular, there must be a set $z \in \mathrm{TC}\left(\mathsf{c}_0^{\mathbf{V}^*}\right)$ such that

$$\mathbf{V}^* \vDash z \cap \mathrm{TC}\left(\mathsf{c}_0^{\mathbf{V}^*}\right) = \emptyset\,,$$

which implies that z must be different from all the sets $\mathsf{c}_n^{\mathbf{V}^*}$. On the other hand, by the Axiom of Foundation, the length of a decreasing sequence of the form

$$\mathsf{c}_0^{\mathbf{V}^*} \ni \mathsf{c}_1^{\mathbf{V}^*} \ni \mathsf{c}_2^{\mathbf{V}^*} \ni \cdots \ni z$$

must be finite in the sense of $\mathbf{V}^*$. In other words, the length of such a decreasing sequence must be an element of $\omega^{\mathbf{V}^*}$, which shows that $\omega^{\mathbf{V}^*}$ cannot be the same set as $\omega^{\mathbf{V}}$; hence, the models $\mathbf{V}^*$ and $\mathbf{V}$ are essentially different. In fact, $\omega^{\mathbf{V}^*}$ is the domain of a non-standard model of PA (like the one we constructed in the last section of Chapter 2). In particular, $\omega^{\mathbf{V}^*}$ is different from the set $\mathbb{N}$ of standard natural numbers.

Since the model we got is just a so-called non-standard model of ZF, we may ask whether there are also different standard models of ZF, (*i.e.*, models of ZF in which the set ω is the same). This is indeed the case, but we have to postpone the technique to build proper models of Set Theory to Part III.

Cardinals and Ordinals in ZF

It is natural to ask whether there exists some kind of upper bound or ceiling for the set-theoretic universe $\mathbf{V}$ or if there exist arbitrarily large sets. In order to address this questions we have to introduce the notion of cardinal numbers.

Two sets A and B are said to have the same **cardinality**, denoted $|A| = |B|$, if there is a bijection between A and B. For example, $|\omega + \omega| = |\omega|$, *e.g.*, define the bijection $f : \omega + \omega \to \omega$ by stipulating

$$f(\alpha) = \begin{cases} \beta + \beta & \text{if } \alpha = \omega + \beta, \\ \alpha + \alpha + 1 & \text{otherwise.} \end{cases}$$

Notice that since every composite of bijections is a bijection, the cardinality equality is an equivalence relation.

If $|A| = |B'|$ for some $B' \subseteq B$, then the cardinality of A is less than or equal to the cardinality of B, denoted $|A| \leq |B|$. Notice that $|A| \leq |B|$ *iff* there is an injection from A into B. Finally, if $|A| \neq |B|$ but $|A| \leq |B|$, then the cardinality of A is said to be strictly less than the cardinality of B, denoted $|A| < |B|$. Notice that the relation "$\leq$" is reflexive and transitive. The notation suggests that $|A| \leq |B|$ and $|B| \leq |A|$ implies $|A| = |B|$. This is indeed the case and is consequence of the following result.

LEMMA 3.13. *Let A_0, A_1, A be sets such that $A_0 \subseteq A_1 \subseteq A$. If $|A| = |A_0|$, then $|A| = |A_1|$.*

Proof. If $A_1 = A$ or $A_1 = A_0$, then the statement is trivial. So, let us assume that $A_0 \subsetneq A_1 \subsetneq A$ and let $C = A \setminus A_1$, *i.e.*, $A \setminus C = A_1$. Further, let $f : A \to A_0$ be a bijection and define $g : \mathscr{P}(A) \to \mathscr{P}(A_0)$ by stipulating $g(D) := f[D]$. With the Axiom Schema of Separation we can define the following set:

$$H = \big\{h \in {}^{\omega}\mathscr{P}(A) : h(0) = C \wedge \forall n \in \omega \,\big(h(n+1) = g\big(h(n)\big)\big)\big\}.$$

By induction on n we show that the set H contains at most one function: If $h, h' \in H$, then by definition $h(0) = C = h'(0)$. Assume now that $h(n) = h'(n)$ for some $n \in \omega$, then

$$h(n+1) = g\big(h(n)\big) = g\big(h'(n)\big) = h'(n+1)\,,$$

and consequently we get $h = h'$. In order to show that H contains at least one function h_0 we proceed as follows: Let $h_0 \subseteq \omega \times \mathscr{P}(A)$ be the set of ordered pairs $\langle n, D\rangle$ for which we have:

- $\langle 0, C\rangle \in h_0$
- $\forall n\, \forall D\, \forall n' \big(\langle n, D\rangle \in h_0 \wedge n' < n \to \exists D' \left(\langle n', D'\rangle \in h_0\right)\big)$
- $\forall n\, \forall D\, \forall n'\, \forall D' \big(\langle n, D\rangle \in h_0 \wedge \langle n', D'\rangle \in h_0 \wedge n' = n+1 \to D' = g(D)\big)$

It remains to show that $h_0 \in {}^{\omega}\mathscr{P}(A)$: By construction, $h_0(0) = C$. Furthermore, if $h_0(n) = D$, for some $n \in \omega$ and $D \subseteq A$, then, by construction, $h_0(n+1) = g(D)$. So, by induction on n we get $h_0 : \omega \to \mathscr{P}(A)$ and $h_0 \in H$. In particular, for all $n \in \omega$ we have $h_0(n+1) = f[h_0(n)]$. Now, let

$$\bar{C} = \bigcup \{h_0(n) : n \in \omega\}$$

and define the function $\tilde{f} : A \to A$ by stipulating

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \bar{C}, \\ x & \text{otherwise.} \end{cases}$$

By definition of $\tilde{f}$ and since f is a bijection which maps C into A_0, $\tilde{f}[\bar{C}] = \bar{C} \setminus C$. Moreover, the function $\tilde{f}$ is injective. To see this, let $x, y \in A$ be *distinct* and consider the following three cases:

(1) If $x, y \in \bar{C}$, then $\tilde{f}(x) = f(x)$ and $\tilde{f}(y) = f(y)$, and since f is injective we get $\tilde{f}(x) \neq \tilde{f}(y)$.

(2) If $x, y \in A \setminus \bar{C}$, then $\tilde{f}(x) = x$ and $\tilde{f}(y) = y$, and hence, $\tilde{f}(x) \neq \tilde{f}(y)$.

(3) If $x \in \bar{C}$ and $y \in A \setminus \bar{C}$, then $\tilde{f}(x) = f(x) \in \bar{C}$ and $\tilde{f}(y) = y \notin \bar{C}$, and therefore, $\tilde{f}(x) \neq \tilde{f}(y)$.

We already know that $\tilde{f}[\bar{C}] = \bar{C} \setminus C$ and by definition we have $\tilde{f}[A \setminus \bar{C}] = A \setminus \bar{C}$. Hence,

$$\tilde{f}[A] = (A \setminus \bar{C}) \,\dot{\cup}\, (\bar{C} \setminus C) = A \setminus C = A_1$$

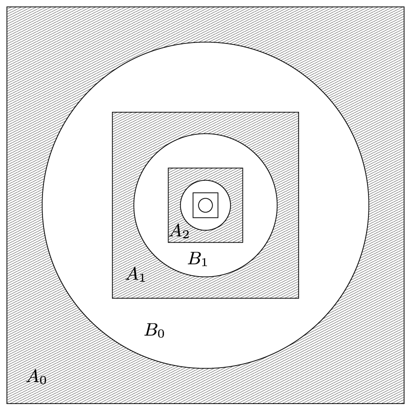
which shows that $|A| = |A_1|$. $\dashv$

THEOREM 3.14 (CANTOR–BERNSTEIN THEOREM). *Let A and B be any sets. If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.*

Proof. Let $f : A \hookrightarrow B$ be a one-to-one mapping from A into B, and $g : B \hookrightarrow A$ be a one-to-one mapping from B into A. Further, let $A_0 := (g \circ f)[A]$ and $A_1 := g[B]$. Then $|A_0| = |A|$ and $A_0 \subseteq A_1 \subseteq A$, hence, by LEMMA 3.13, $|A| = |A_1|$, and since $|A_1| = |B|$ we have $|A| = |B|$. $\dashv$

As an alternative proof of the CANTOR–BERNSTEIN THEOREM 3.14 we give Bernstein's original proof:

Bernstein's proof of the Cantor–Bernstein Theorem. Let A and B be two arbitrary sets and let $f : A \hookrightarrow B$ and $g : B \hookrightarrow A$ two injections. Further, let $A_0 := A$, $B_0 := g[B]$, and for $n \in \omega$ let $A_{n+1} := (g\circ f)[A_n]$ and $B_{n+1} := (g\circ f)[B_n]$, and finally let $D := \bigcap_{n\in\omega} A_n$. This construction is visualised by the following picture.



We leave it as an exercise to reader to show that the sets A_n and B_n have the following properties:

1. $A_0 = D \cup (A_0 \setminus B_0) \cup (B_0 \setminus A_1) \cup (A_1 \setminus B_1) \cup (B_1 \setminus A_2) \cup \dots$
2. $B_0 = D \cup (B_0 \setminus A_1) \cup (A_1 \setminus B_1) \cup (B_1 \setminus A_2) \cup (A_2 \setminus B_2) \cup \dots$
3. For all $n \in \omega$, $|A_n \setminus B_n| = |A_{n+1} \setminus B_{n+1}|$.

Since the sets $(A_n \setminus B_n)$, $(B_n \setminus A_{n+1})$, and D, are pairwise disjoint, by (c) and by regrouping the representation of B_0 in (b), we get $|A_0| = |B_0|$. $\dashv$

As an application of the CANTOR–BERNSTEIN THEOREM 3.14 let us show that the set of **real numbers**, denoted by $\mathbb{R}$, has the same cardinality as $\mathscr{P}(\omega)$.

PROPOSITION 3.15. $|\mathbb{R}| = |\mathscr{P}(\omega)|$.

Proof. Cantor showed that every real number $r > 1$ can be written in a unique way as a product of the form

$$r = \prod_{n\in\omega} \left(1 + \frac{1}{q_n}\right)$$

where all q_n's are positive integers and for all $n \in \omega$ we have $q_{n+1} \geq q_n^2$. Such products are called *Cantor products*. So, for every real number $r > 1$ there exists a unique infinite sequence $q_0(r), q_1(r), \ldots, q_n(r), \ldots$ of positive integers with $q_{n+1} \geq q_n^2$ (for all $n \in \omega$) such that $r = \prod_{n\in\omega}(1 + \frac{1}{q_n})$.

Let us first show that $|\mathbb{R}| \leq |\mathscr{P}(\omega)|$: For $r \in \mathbb{R}$ let

$$f(r) = \left\{\sum_{j\leq n} q_j(r)(2^j + 1) : n \in \omega\right\}.$$

Define the function $h : \mathbb{R} \to \mathbb{R}$ by stipulating $h(x) := 1 + e^x$, where e is the *Euler number* and $e^x = \sum_{n\in\omega}(x^n/n!)$. Then h is a bijection between $\mathbb{R}$ and the set of real numbers $r > 1$. We leave it as an exercise to the reader to verify that the composition $f \circ h$ is an injective mapping from $\mathbb{R}$ into $\mathscr{P}(\omega)$.

To see that $|\mathscr{P}(\omega)| \leq |\mathbb{R}|$, consider for example the function $g(x) = \sum_{n\in x} 3^{-n}$, where $g(\emptyset) := 0$, which is obviously a injective mapping from $\mathscr{P}(\omega)$ into $\mathbb{R}$ (or more precisely, into the interval $[0, \frac{3}{2}]$).

So, by the CANTOR–BERNSTEIN THEOREM 3.14, $|\mathbb{R}| = |\mathscr{P}(\omega)|$. $\dashv$

So far, we have used the notion of cardinality in order to compare the sizes of two sets. Now, we will show that the cardinality of a set can also be considered as a set. First, the cardinality of a set A, denoted $|A|$, can be defined as the class of all sets B which have the same cardinality as A (*i.e.*, for which there exists a bijection between A and B), but this would have the disadvantage that except for $A = \emptyset$, $|A|$ would not belong to the set-theoretic universe. However, with the Axiom of Foundation the cardinality of a set A can be defined as a proper set:

$$|A| = \{B \in \mathrm{V}_{\beta_0} : \text{there exists a bijection between } B \text{ and } A\}$$

where β_0 is the least ordinal number for which there is a $B \in \mathrm{V}_{\beta_0}$ such that B has the same cardinality as A. Notice that, for example, $|\emptyset| = \{\emptyset\}$, where $\{\emptyset\} \subseteq \mathrm{V}_1$ (in this case, $\beta_0 = 1$). The set $|A|$ is called a **cardinal number**, or just a **cardinal**. Notice that A is not necessarily a member of $|A|$. Further, notice that $|A| = |B|$ *iff* there is a bijection between A and B, and as above we write $|A| \leq |B|$ if $|A| = |B'|$ for some $B' \subseteq B$. Cardinal numbers are usually denoted by Fraktur letters like $\mathfrak{m}$ and $\mathfrak{n}$. A cardinal number is **finite** if it is the cardinality of a natural number $n \in \omega$, otherwise it is **infinite**. Finite cardinals are usually denoted by letters like $n, m, \ldots$. An infinite cardinal which contains a well-orderable set is traditionally called an **aleph** and consequently denoted by an "$\aleph$", *e.g.*, $\aleph_0 := |\omega|$. The following fact summarises some simple properties of alephs.

FACT 3.16. *All sets which belong to an aleph can be well-ordered and the cardinality of any ordinal is an aleph. Further, for any ordinals $\alpha, \beta \in \Omega$ we have $|\alpha| < |\beta|$ or $|\alpha| = |\beta|$ or $|\alpha| > |\beta|$, and these three cases are mutually exclusive.*

A non-empty set A is called **uncountable** if there is no enumeration of the elements of A, or equivalently, no mapping from ω to A is surjective.

By the Axiom of Infinity we know that there is an infinite set and we have seen that there is even a smallest infinite ordinal, namely ω, which is of course a countable set. Now, the question arises whether every infinite set is countable. We answer this question in two steps: First we show that the set of real numbers is uncountable, and then we show that in general, for every set A there exists a set of strictly greater cardinality than A—which implies that there is no largest cardinal.

PROPOSITION 3.17. *The set of real numbers is uncountable.*

Proof. By PROPOSITION 3.15 we already know that there is a bijection between $\mathbb{R}$ and $\mathscr{P}(\omega)$. Further, we have $|\mathscr{P}(\omega)| = |{}^{\omega}2|$. Indeed, for every $x \in \mathscr{P}(\omega)$ let $\chi_x \in {}^{\omega}2$ be such that

$$\chi_x(n) = \begin{cases} 1 & \text{if } n \in x, \\ 0 & \text{otherwise.} \end{cases}$$

So, it is enough to show that no mapping from ω to ${}^{\omega}2$ is surjective. Let

$$\begin{aligned} g : \omega &\longrightarrow {}^{\omega}2 \\ n &\longmapsto f_n \end{aligned}$$

be any mapping from ω to ${}^{\omega}2$. Define the function $f \in {}^{\omega}2$ by stipulating

$$f(n) = 1 - f_n(n).$$

Then for each $n \in \omega$ we have $f(n) \neq f_n(n)$, so, f is distinct from every function f_n ($n \in \omega$), which shows that g is not surjective. $\dashv$

For cardinals $\mathfrak{m} = |A|$ let $2^{\mathfrak{m}} := |\mathscr{P}(A)|$. By modifying the proof above we can show the following result:

THEOREM 3.18 (CANTOR'S THEOREM). *For every cardinal* $\mathfrak{m}$, $2^{\mathfrak{m}} > \mathfrak{m}$.

Proof. Let $A \in \mathfrak{m}$ be arbitrary. It is enough to show that there is an injection f from A into $\mathscr{P}(A)$, but there is no surjection g from A onto $\mathscr{P}(A)$.

First, for $A = \emptyset$ let $f = \emptyset$, and for $A \neq \emptyset$ define:

$$\begin{aligned} f : A &\longrightarrow \mathscr{P}(A) \\ x &\longmapsto \{x\} \end{aligned}$$

Then f is obviously an injection from A into $\mathscr{P}(A)$ and therefore we get $\mathfrak{m} \leq 2^{\mathfrak{m}}$.

Now, let $g : A \to \mathscr{P}(A)$ be an arbitrary function and let

$$\Gamma := \{x \in A : x \notin g(x)\}.$$

As a subset of A, the set Γ is an element of $\mathscr{P}(A)$. Let x be an arbitrary element of A. Then

$$x \in \Gamma \iff x \notin g(x)$$

which shows that $g(x) \neq \Gamma$, and hence, since x was arbitrary, g is not surjective. ⊣

As an immediate consequence of CANTOR'S THEOREM 3.18 we find that there are arbitrarily large cardinal numbers. Before we show that there are also arbitrarily large ordinal numbers, recall that a binary relation $R \subseteq A \times A$ is a well-ordering on A, if R is a linear ordering on A and every non-empty subset of A has an R-minimal element.

The following proposition is crucial in order to define the *order type* of well-ordered sets.

PROPOSITION 3.19. *If $\alpha, \beta \in \Omega$ and $f : \alpha \to \beta$ is an order-preserving bijection (i.e., for all $\gamma_1 \in \gamma_2 \in \alpha$ we have $f(\gamma_1) \in f(\gamma_2)$), then $\alpha = \beta$.*

Proof. Assume towards a contradiction that there is a pair of distinct ordinals $\alpha, \beta \in \Omega$ with $\beta \in \alpha$ such that there exists an order-preserving bijection $f : \alpha \to \beta$. Let α_0, β_0 be such a pair with least ordinal $\alpha_0 \in \Omega$ and let $f : \alpha_0 \to \beta_0$ be an order-preserving bijection. Since $\beta_0 \in \alpha_0$ and f is a bijection, $f(\beta_0) \in \beta_0$. Furthermore, since f is order-preserving, for all $\gamma, \delta \in \Omega$ we have

$$\gamma \in \beta_0 \to f(\gamma) \in f(\beta_0) \qquad \text{and} \qquad \delta \in f(\beta_0) \to f^{-1}(\delta) \in \beta_0$$

which shows that $f[\beta_0] = f(\beta_0)$. Hence, $f|_{\beta_0}$ is an order-preserving bijection between β_0 and $f(\beta_0)$, and since $\beta_0 \in \alpha_0$ and $f(\beta_0) \in \beta_0$, the pair $\beta_0, f(\beta_0)$ contradicts the choice of the pair α_0, β_0 with least ordinal α_0. ⊣

As a consequence of the previous result we show that each well-ordering R of a set A corresponds to exactly one ordinal. This ordinal number is called the **order type** of R and is denoted $\mathrm{o.t.}(R)$. The proof of the following result is essentially Zermelo's first proof of the so-called Well-Ordering Principle, which shall be discussed below.

PROPOSITION 3.20. *Let A be an arbitrary set and let R be a well-ordering on A. Then there exists a unique ordinal number $\mathrm{o.t.}(R) \in \Omega$ for which there is a bijection $w : \mathrm{o.t.}(R) \to A$ such that for all $\alpha, \beta \in \mathrm{o.t.}(R)$,*

$$\alpha \in \beta \iff w(\alpha) \, R \, w(\beta) \, .$$

Proof. Let R be a well-ordering on A. If $A = \emptyset$, then $R = \emptyset$, $w = \emptyset$, and we define $\mathrm{o.t.}(A) := \emptyset$. Otherwise, let $\mathscr{P}^*(A) := \mathscr{P}(A) \setminus \{\emptyset\}$ and let $f : \mathscr{P}^*(A) \to A$ be such that for each non-empty set $S \in \mathscr{P}^*(A)$, $f(S)$ is the R-minimal element of S.

A one-to-one function $w_\alpha : \alpha \hookrightarrow A$, where $\alpha \in \Omega$, is an *f-set* if for all $\gamma \in \alpha$:

$$w_\alpha(\gamma) = f\big(A \setminus \{w_\alpha(\delta) : \delta \in \gamma\}\big).$$

For example, $w_1(0) = f(A)$ is an f-set, in fact, w_1 is the unique f-set with domain $\{0\}$. For distinct f-sets w_α and w_β let $w_\alpha \prec w_\beta$ if $w_\beta|_\alpha = w_\alpha$. Notice that $w_\alpha \prec w_\beta$ implies $\alpha \in \beta$.

CLAIM. *For any two f-sets w_α and w_β we have*

$$w_\alpha \prec w_\beta \quad \text{or} \quad w_\alpha = w_\beta \quad \text{or} \quad w_\beta \prec w_\alpha\,,$$

where these three cases are mutually exclusive.

Proof of Claim. Let w_α and w_β be any two f-sets and let

$$\Gamma = \big\{\gamma \in (\alpha \cap \beta) : w_\alpha(\gamma) \neq w_\beta(\gamma)\big\}.$$

If $\Gamma \neq \emptyset$, then, for $\gamma_0 = \bigcap \Gamma$, we have $w_\alpha(\gamma_0) \neq w_\beta(\gamma_0)$. On the other hand, for all $\delta \in \gamma_0$ we have $w_\alpha(\delta) = w_\beta(\delta)$, thus, by the definition of f-sets, we get $w_\alpha(\gamma_0) = w_\beta(\gamma_0)$. Hence, $\Gamma = \emptyset$, and consequently we get:

$$\begin{aligned} w_\alpha \prec w_\beta &\iff \alpha \in \beta \\ w_\alpha = w_\beta &\iff \alpha = \beta \\ w_\beta \prec w_\alpha &\iff \beta \in \alpha \end{aligned}$$

By THEOREM 3.3 (a), this proves the CLAIM. $\dashv_{\text{Claim}}$

Now, let $W := \{\alpha \in \Omega : w_\alpha \textit{ is an } f\textit{-set}\}$ be a collection (not necessarily a set) of ordinals and let

$$\mathscr{C} = \big\{X \subseteq A : \exists \alpha \in W (w_\alpha[\alpha] = X)\big\}\,.$$

Obviously, $\mathscr{C}$ is a subset of $\mathscr{P}(A)$. Furthermore, by the properties of f-sets, for any $\alpha, \beta \in W$ we have

$$\alpha \neq \beta \iff w_\alpha[\alpha] \neq w_\beta[\beta]\,,$$

which implies that there exists a bijection between $\mathscr{C}$ and W. Therefore, since $\mathscr{C}$ is a set, by Axiom Schema of Replacement we get that W is also a set, and hence, $\gamma := \bigcup W$ is an ordinal. Moreover, w_γ is even an f-set.

Finally, let $A' := w_\gamma[\gamma]$. Then $A' = A$; otherwise, w_γ could be extended to the f-set

$$w_{\gamma+1} = w_\gamma \cup \big\{\langle \gamma, f(A \setminus A')\rangle\big\}\,,$$

and since $\gamma + 1 \notin W$, this would contradict the definition of W.

Thus, the f-set w_γ is a bijection between the ordinal γ and the set A, and by PROPOSITION 3.19 we get that γ—which is the order type of R—is unique. $\dashv$

THEOREM 3.21 (HARTOGS' THEOREM). *For every cardinal* $\mathfrak{m}$ *there is a smallest aleph, denoted* $\aleph(\mathfrak{m})$*, such that* $\aleph(\mathfrak{m}) \nleq \mathfrak{m}$*.*

Proof. Let $A \in \mathfrak{m}$ be arbitrary and let $\mathscr{R} \subseteq \mathscr{P}(A \times A)$ be the set of all well-orderings of subsets of A. For every $R \in \mathscr{R}$, $\mathrm{o.t.}(R)$ is an ordinal, and for every $R \in \mathscr{R}$ and any $\beta \in \mathrm{o.t.}(R)$ there is an $R' \in \mathscr{R}$ such that $\mathrm{o.t.}(R') = \beta$, which shows that

$$\alpha = \big\{\mathrm{o.t.}(R) : R \in \mathscr{R}\big\}$$

is an ordinal. By definition, for every $\beta \in \alpha$ there is a well-ordering R_S of some $S \subseteq A$ such that $\mathrm{o.t.}(R_S) = \beta$, which implies that $|\beta| \leq |A|$. On the other hand, $|\alpha| \leq |A|$ would imply that $\alpha \in \alpha$, which is obviously a contradiction. Let $\aleph(\mathfrak{m}) := |\alpha|$, then $\aleph(\mathfrak{m}) \nleq \mathfrak{m}$ and for each $\aleph < \aleph(\mathfrak{m})$ we have $\aleph \leq \mathfrak{m}$. $\dashv$

COROLLARY 3.22. *For every ordinal number* α *there exists an ordinal number* β *such that* $|\beta| > |\alpha|$*. Furthermore, for every cardinal number* $\mathfrak{m}$*, there exists an ordinal number* β *such that* $|\beta| \nleq \mathfrak{m}$*.*

Proof. For the first inequality let $\alpha \in \Omega$ and let $\mathfrak{n} = |\alpha|$. By HARTOGS' THEOREM 3.21 there is an aleph, namely $\aleph(\mathfrak{n})$, such that $\aleph(\mathfrak{n}) \nleq \mathfrak{n}$. Now, since $\mathfrak{n}$ and $\aleph(\mathfrak{n})$ both contain well-ordered sets we have $\mathfrak{n} < \aleph(\mathfrak{n})$. Let $w \in \aleph(\mathfrak{n})$ be a well-ordered set and let β be the order type of w. Then $\aleph(\mathfrak{n}) = |\beta| > |\alpha| = \mathfrak{n}$.

For the second inequality let β be the order type of a well-ordered set which belongs to $\aleph(\mathfrak{m})$. Then, by definition of $\aleph(\mathfrak{m})$, we have $|\beta| \nleq \mathfrak{m}$. $\dashv$

Zermelo's Axiom of Choice

In 1904, Zermelo published his first proof that every set can be well-ordered. The proof is based on the so-called Axiom of Choice, denoted AC, which, in Zermelo's words, states that *every product of an infinite totality of non-empty sets is non-empty*. The full theory ZF + AC, denoted ZFC, is called *Set Theory*.

In order to state the Axiom of Choice we first define the notion of a choice function: If $\mathscr{F}$ is a family of non-empty sets (*i.e.*, $\emptyset \notin \mathscr{F}$), then a **choice function** for $\mathscr{F}$ is a function $f : \mathscr{F} \to \bigcup \mathscr{F}$ such that for each $x \in \mathscr{F}$, $f(x) \in x$.

The Axiom of Choice—which completes the axiom system of Set Theory and which is in our counting the ninth axiom of ZFC—states as follows:

9. The Axiom of Choice

$$\forall \mathscr{F}\Big(\emptyset \notin \mathscr{F} \to \exists f\Big(f \in {}^{\mathscr{F}}\bigcup \mathscr{F} \wedge \forall x \in \mathscr{F}\big(f(x) \in x\big)\Big)\Big)$$

Informally, every family of non-empty sets has a choice function, or equivalently, every Cartesian product of non-empty sets is non-empty. To see this, notice that each element of $\bigtimes_{x \in \mathscr{F}} x$ is a choice function for $\mathscr{F}$.

A seemingly different statement, which is in fact equivalent to the Axiom of Choice (as we will see below), is the following principle:

Well-Ordering Principle. Every set can be well-ordered.

Before we show that the Axiom of Choice and the Well-Ordering Principle are equivalent, let us first illustrate the difficulties in finding a well-ordering on the sets $\mathbb{Q}$ and $\mathbb{R}$. Obviously, both sets are linearly ordered by "$<$". However, since for any elements x and y with $x < y$ there exists a z such that $x < z < y$, the ordering "$<$" is far away from being a well-ordering. Even though $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ have similar properties (at least from an order-theoretical point of view), when we try to well-order these sets they behave very differently. However, as we will show later (see FACT 5.1), $\mathbb{Q}$ is countable (*i.e.*, there is a bijection between $\mathbb{Q}$ and ω), and any bijection $f : \mathbb{Q} \to \omega$ allows us to define a well-ordering "$\prec$" on $\mathbb{Q}$ by stipulating $q \prec p \iff f(q) < f(p)$. Now, let us consider the set $\mathbb{R}$. For example, we could first well-order the rational numbers, or even the algebraic numbers, and then try to extend this well-ordering to all real numbers. However, this attempt—as well as all other attempts—to construct explicitly a well-ordering of the reals will end in failure (the reader is invited to verify this claim by writing down explicitly some orderings of $\mathbb{R}$).

As mentioned above, Zermelo proved in 1904 that the Axiom of Choice implies the Well-Ordering Principle. Just a few years later, he published a new proof, which we present now.

THEOREM 3.23. *The* Well-Ordering Principle *is equivalent to the* Axiom of Choice.

Proof. ($\Leftarrow$) Let us first try a naive attempt: Let S be a non-empty set on which we would like to define a well-ordering. In the first step, we choose an element a_0 from S. In the second step we choose an element a_1 from $S \setminus \{a_0\}$, then we choose an element a_2 from $S \setminus \{a_0, a_1\}$, and so on, until $S \setminus \{a_0, a_1, \ldots\}$ is empty. Furthermore, we define an ordering "$\prec$" on S by stipulating $a \prec b$ *iff in this process, a was selected earlier than b*, and finally we show that this ordering is a well-ordering on S. There are some obvious difficulties with this attempt. Firstly, why can we be sure that the process terminates? In other words, how can we prove that $S \setminus \{a_0, a_1, \ldots\}$ eventually becomes the empty set? Secondly, how can we show that the ordering we defined is a well-ordering on S? On the other hand, the easy part is to select an element from a non-empty subset of S, which we can do with a choice function for $\mathscr{P}(S) \setminus \{\emptyset\}$ (given by AC).

Let S be a set. If $S = \emptyset$, then S is already well-ordered. So, assume that $S \neq \emptyset$ and let $\mathscr{P}^*(S) := \mathscr{P}(S) \setminus \{\emptyset\}$. Further, let $f : \mathscr{P}^*(S) \to S$ be an arbitrary but fixed choice function for $\mathscr{P}^*(S)$ (which exists by AC). Notice that for every $A \in \mathscr{P}^*(S)$ we have $f(A) \in A$. For every $A \in \mathscr{P}^*(S)$, let $A' := A \setminus \{f(A)\}$.

Now, a set $\mathscr{W} \subseteq \mathscr{P}(S)$ is called a Θ-*chain* if it satisfies the following three conditions:

(a) $S \in \mathscr{W}$,

(b) for each non-empty $A \in \mathscr{W}$, $A' \in \mathscr{W}$,

(c) for each set $\mathscr{A} \subseteq \mathscr{W}$, $\bigcap \mathscr{A} \in \mathscr{W}$.

For example, $\mathscr{P}(S)$ is a Θ-chain, which shows that the set of Θ-chains $\mathscr{W} \subseteq \mathscr{P}(S)$ is non-empty. Furthermore, since $\bigcap \emptyset = \emptyset$, each Θ-chain contains $\emptyset$. As an immediate consequence of the definition we get that the intersection of Θ-chains is again a Θ-chain. In particular, the intersection $\mathscr{W}_0$ of *all* Θ-chains is a Θ-chain, namely the Θ-chain which does not properly contain any other Θ-chain (*i.e.*, no proper subset of $\mathscr{W}_0$ is a Θ-chain).

An element $A \in \mathscr{W}_0$ is called a *cut* if for all $X \in \mathscr{W}_0$ we have *either* $X \subsetneq A$ *or* $A \subseteq X$. For each cut $A \in \mathscr{W}_0$, let

$$U_A = \{X \in \mathscr{W}_0 : X \subsetneq A\} \quad \text{and} \quad V_A = \{X \in \mathscr{W}_0 : A \subseteq X\}.$$

Let $\mathscr{C}_0 \subseteq \mathscr{W}_0$ be the set of cuts. We show now that $\mathscr{C}_0$ is a Θ-chain: Since S is obviously a cut, $\mathscr{C}_0$ satisfies condition (a); conditions (b) & (c) are proved in the following two claims.

CLAIM 1. *If A is a non-empty element of $\mathscr{C}_0$, then also $A' \in \mathscr{C}_0$.*

Proof of Claim 1. Let A be a non-empty cut. We have to show that for every $X \in \mathscr{W}_0$, either $X \in V_{A'}$ or $X \in U_{A'}$. For this, we first show that

$$\mathscr{W}_{A'} := U_{A'} \cup \{A'\} \cup V_A$$

is a Θ-chain. Obviously we have $S \in \mathscr{W}_{A'}$, which shows that $\mathscr{W}_{A'}$ satisfies condition (a). To see that $\mathscr{W}_{A'}$ satisfies condition (b), take an arbitrary $X \in \mathscr{W}_{A'}$. If $X \in V_A$ and $X \neq A$, then $A \subsetneq X$ and $X' \notin U_A$. Because A is a cut, this implies $X' \in V_A$, hence $X' \in \mathscr{W}_{A'}$. If $X = A$, then $X' = A'$, hence $X \in \mathscr{W}_{A'}$. Finally, if $X \in U_{A'}$ or $X = A'$, then $X' \in U_{A'}$, hence $X \in \mathscr{W}_{A'}$. So, for every $X \in \mathscr{W}_{A'}$ we have $X' \in \mathscr{W}_{A'}$, which shows that $\mathscr{W}_{A'}$ satisfies condition (b). For condition (c), take an arbitrary $\mathscr{A} \subseteq \mathscr{W}_{A'}$. If $\mathscr{A} \cap U_{A'} \neq \emptyset$, then $\bigcap \mathscr{A} \in U_{A'}$, hence $\bigcap \mathscr{A} \in \mathscr{W}_{A'}$. If $\mathscr{A} \cap U_{A'} = \emptyset$ and $A \notin \mathscr{A}$, then $\mathscr{A} \subseteq V_A$, which implies $\bigcap \mathscr{A} \in V_A$, hence $\bigcap \mathscr{A} \in \mathscr{W}_{A'}$. Finally, if $\mathscr{A} \cap U_{A'} = \emptyset$ and $A \in \mathscr{A}$, then $\bigcap \mathscr{A} = A$, hence

$\bigcap \mathscr{A} \in \mathscr{W}_{A'}$. So, for every $\mathscr{A} \subseteq \mathscr{W}_{A'}$ we have $\bigcap \mathscr{A} \in \mathscr{W}_{A'}$, which shows that $\mathscr{W}_{A'}$ satisfies condition (c).

Now, since $\mathscr{W}_{A'} \subseteq \mathscr{W}_0$ is a Θ-chain and $\mathscr{W}_0$ is the smallest Θ-chain, we must have $\mathscr{W}_{A'} = \mathscr{W}_0$, *i.e.*, $\mathscr{W}_0 = U_{A'} \cup \{A\} \cup V_A$, which shows that A' is a cut. $\dashv_{\text{Claim 1}}$

CLAIM 2. *If $\mathscr{A} \subseteq \mathscr{C}_0$ is a set of cuts, then $A_0 := \bigcap \mathscr{A}$ is also a cut.*

Proof of Claim 2. For each $C \in \mathscr{C}_0$ we have either $C \in V_A$ for some $A \in \mathscr{A}$, or for all $A \in \mathscr{A}$ we have $C \in U_A$. In the former case we get $C \in V_{A_0}$, and in the latter case we get $C \subseteq A_0$, which implies $C \in U_{A_0}$ or $C = A_0$ (*i.e.*, $C \in V_{A_0}$). So, for each $C \in \mathscr{C}_0$ we have either $C \in V_{A_0}$ or $C \in U_{A_0}$, which shows that A_0 is a cut. $\dashv_{\text{Claim 2}}$

Thus, the set of cuts $\mathscr{C}_0$ is a Θ-chain, and since $\mathscr{C}_0 \subseteq \mathscr{W}_0$, by definition of $\mathscr{W}_0$ we get $\mathscr{C}_0 = \mathscr{W}_0$. In other words, every element $A \in \mathscr{W}_0$ is a cut. In particular, for any distinct elements $A, B \in \mathscr{W}_0$ we have either $B \subsetneq A$ and $B \subseteq A'$ (if $B \in U_A$), or $A \subsetneq B$ and $A \subseteq B'$ (if $B \in V_A$).

Let now $P \in \mathscr{P}^*(S)$ be an arbitrary non-empty subset of S, and let

$$\bar{P} = \bigcap \{A \in \mathscr{W}_0 : P \subseteq A\}.$$

Since $\mathscr{W}_0$ is a Θ-chain, by condition (c) we get $\bar{P} \in \mathscr{W}_0$, and by definition of $\bar{P}$ we have $P \subseteq \bar{P}$. Since $\bar{P}'$ also belongs to $\mathscr{W}_0$, by definition of $\bar{P}$ we get $P \nsubseteq \bar{P}'$, which implies that

$$f(\bar{P}) \in P.$$

CLAIM 3. *The set $\bar{P}$ is the unique element of $\mathscr{W}_0$ such that*

$$P \subseteq \bar{P} \quad \text{and} \quad f(\bar{P}) \in P.$$

Proof of Claim 3. Let $\tilde{P} \in \mathscr{W}_0$ be such that $P \subseteq \tilde{P}$. If $\tilde{P} \neq \bar{P}$, then, since each element of $\mathscr{W}_0$ is a cut, we have either $\bar{P} \subsetneq \tilde{P}$ or $\tilde{P} \subsetneq \bar{P}$. In the former case we have $\bar{P} \subseteq \tilde{P}'$, which implies that $f(\tilde{P}) \notin P$, and in the latter case we have $\tilde{P} \subseteq \bar{P}'$, and since $f(\bar{P}) \in P$, we get $P \nsubseteq \tilde{P}$. Hence, if $\tilde{P} \neq \bar{P}$, then $f(\tilde{P}) \notin P$ or $P \nsubseteq \tilde{P}$, which completes the proof. $\dashv_{\text{Claim 3}}$

CLAIM 4. *For every $a \in S$ there exists a unique $W_a \in \mathscr{W}_0$ such that $f(W_a) = a$*

Proof of Claim 4. To see this, let $W_a := \overline{\{a\}}$, *i.e.*,

$$W_a = \bigcap \{A \in \mathscr{W}_0 : a \in A\}.$$

Since $f(W_a) \in \{a\}$, we get $f(W_a) = a$. For the uniqueness of W_a, let $\tilde{P} \in \mathscr{W}_0$ be such that $f(\tilde{P}) = a$. If $\tilde{P}$ is distinct from W_a, then, since W_a is a cut, we have either $\tilde{P} \subseteq W'_a$ or $W_a \subseteq \tilde{P}'$, and in both cases we have $f(\tilde{P}) \neq a$ (either because $a \in W'_a$ or because $a \in \tilde{P}'$). $\dashv_{\text{Claim 4}}$

Since for each $a \in S$, there exists a unique $W_a \in \mathscr{W}_0$ such that $a \in W_a$ and $f(W_a) = a$, and since each element of $\mathscr{W}_0$ is a cut, for any distinct elements $a, b \in S$ we have either $W_a \subsetneq W_b$ or $W_b \subsetneq W_a$. In the former case we get $a \in W_b$ and write $b \prec a$, and in the latter case we get $b \in W_a$ and write $a \prec b$, *i.e.*,

$$a \prec b \;:\iff\; W_b \subsetneq W_a\,.$$

By the properties of $\mathscr{W}_0$ (*i.e.*, if $a \neq b$ then either $W_b \subsetneq W_a$ or $W_a \subsetneq W_b$) we get that the binary relation "$\prec$" is a linear ordering on S.

Finally, let $P \in \mathscr{P}^*(S)$ be an arbitrary non-empty subset of S and let $\bar{P}$ be as above. Then $f(\bar{P})$ is the $\prec$-minimal element of P: To see this, let $a := f(\bar{P})$ and let $b \in P$ be distinct from a. Notice, that since $\bar{P} \in \mathscr{W}_0$ and $f(\bar{P}) = a$, we obtain $\bar{P} = W_a$ which implies

$$P \subseteq W_a\,.$$

Now, for $W_b \in \mathscr{W}_0$, we have $f(W_b) = b$ and either $W_b \subsetneq W_a$ or $W_a \subsetneq W_b$. In the former cases we get $a \prec b$. In the latter case we get $W_a \subseteq W_b \setminus \{b\}$, and since $P \subseteq W_a$, we get $b \notin P$, contrary to our assumption.

So, every non-empty subset of S has a $\prec$-minimal element, which shows that the binary relation "$\prec$" is a well-ordering on S, and since S was arbitrary, every set can be well-ordered.

($\Rightarrow$) Let $\mathscr{F}$ be any family of non-empty sets and let "$<$" be any well-ordering on $\bigcup \mathscr{F}$. Define $f : \mathscr{F} \to \bigcup \mathscr{F}$ by stipulating $f(x)$ being the $<$-minimal element of x. $\dashv$

Gödel's Model of ZFC

Before we proceed, we should address the question whether AC is consistent relative to the other axioms of Set Theory (*i.e.*, relative to ZF), which is indeed the case.

Assume that ZF is consistent, then, by PROPOSITION 2.5, ZF has a model, say $\mathbf{V}$. To obtain the relative consistency of AC with ZF, we have to show that ZF+AC also has a model. In 1935, Gödel informed von Neumann at the Institute for Advanced Study in Princeton that he had found such a model. In fact, he showed that there exists a smallest transitive subclass of $\mathbf{V}$ which contains all ordinals (*i.e.*, contains Ω as a subclass) in which AC as well as ZF holds. This unique submodel of $\mathbf{V}$ is called the **constructible universe** and is denoted by $\mathbf{L}$. According to Paul Bernays, Gödel originally used the old German script "$\mathscr{L}$" to denote the constructible universe, where "$\mathscr{L}$" is a capital "C" and not—as one could think—a capital "L". Roughly speaking, the model $\mathbf{L}$ consists of all "mathematically constructible" sets, or in other words, all sets which are "constructible" or "describable", but nothing else. To be more precise, let us give the following definitions:

Let M be a set and $\varphi(x_0, \ldots, x_n)$ be a first-order formula in the language $\{\in\}$. Then φ^M denotes the formula we obtain by replacing all occurrences of "$\exists x$" and "$\forall x$" with "$\exists x \in M$" and "$\forall x \in M$", respectively. A subset $y \subseteq M$ is **definable over** M if there is a first-order formula $\varphi(x_0, \ldots, x_n)$ in the language $\{\in\}$, and parameters $a_1, \ldots, a_n$ in M, such that $\{z : \varphi^M(z, a_1, \ldots, a_n)\} = y$. Finally, for any set M:

$$\operatorname{def}(M) = \{y \subseteq M : y \text{ is definable over } M\}.$$

Notice that for any set M, $\operatorname{def}(M)$ is a set, being itself a subset of $\mathscr{P}(M)$. Now, by induction on $\alpha \in \Omega$, define the following sets (compare with the cumulative hierarchy defined above):

$$\begin{aligned} \mathrm{L}_0 &:= \emptyset, \\ \mathrm{L}_{\alpha+1} &:= \operatorname{def}(\mathrm{L}_\alpha), \\ \mathrm{L}_\alpha &:= \bigcup_{\beta \in \alpha} \mathrm{L}_\beta \quad \text{if } \alpha \text{ is a limit ordinal}, \\ \mathbf{L} &:= \bigcup_{\alpha \in \Omega} \mathrm{L}_\alpha. \end{aligned}$$

For the cumulative hierarchy, one can show that for each $\alpha \in \Omega$, L_α is a transitive set, $\alpha \subseteq \mathrm{L}_\alpha$ and $\alpha \in \mathrm{L}_{\alpha+1}$, and that $\alpha \in \beta$ implies $\mathrm{L}_\alpha \subsetneq \mathrm{L}_\beta$.

Moreover, Gödel showed that $\mathbf{L} \vDash \mathsf{ZF} + \mathsf{AC}$, and that $\mathbf{L}$ is the smallest transitive class containing Ω as a subclass such that $\mathbf{L} \vDash \mathsf{ZFC}$. Thus, by starting with any model $\mathbf{V}$ of ZF we find a subclass $\mathbf{L}$ of $\mathbf{V}$ such that $\mathbf{L} \vDash \mathsf{ZFC}$. In other words, we find that if ZF is consistent then so is ZFC (roughly speaking, if ZFC is inconsistent, then AC cannot be blamed for it).

Let us now work in ZFC and let us turn back to cardinals, or more precisely, to cardinal arithmetic.

Cardinal Arithmetic in ZFC

In the presence of AC we are able to define cardinal numbers as ordinals: For any set A we define

$$|A| = \bigcap \left\{\alpha \in \Omega : \text{there is a bijection between } \alpha \text{ and } A\right\}.$$

In order to see that this definition makes sense, notice that by AC, every set A is well-orderable and that by PROPOSITION 3.20 every well-ordering on A corresponds to exactly one ordinal (which is the order type of the well-ordering). So, for each set A, the set of ordinals α such that there is a bijection between α and A is a non-empty set of ordinals. Hence, $|A|$ defined as above, is an ordinal.

For example, we have $|n| = n$ for every $n \in \omega$, and $|\omega| = \omega$, but in general, for $\alpha \in \Omega$, we do not have $|\alpha| = \alpha$. For example, $|\omega + 1| \neq \omega + 1$, since $|\omega + 1| = \omega$ and $\omega \neq \omega + 1$. However, there are also other ordinals α beside $n \in \omega$ and ω itself for which we have $|\alpha| = \alpha$, which leads to the following definition:

An ordinal number $\kappa \in \Omega$ such that $|\kappa| = \kappa$ is called a **cardinal number**, or just a **cardinal**. Cardinal numbers are usually denoted by Greek letters like κ, λ, μ, *et cetera*.

A cardinal κ is **infinite** if $\kappa \notin \omega$, otherwise, it is **finite**. In other words, a cardinal is finite if and only if it is a natural number.

Since cardinal numbers are just a special kind of ordinal, they are well-ordered by "$\in$". However, for cardinal numbers κ and λ we usually write $\kappa < \lambda$ instead of $\kappa \in \lambda$, thus,

$$\kappa < \lambda \iff \kappa \in \lambda.$$

Let κ be a cardinal. The smallest cardinal number which is greater than κ is denoted by κ^+, thus,

$$\kappa^+ = \bigcap \big\{\alpha \in \Omega : \kappa < |\alpha|\big\}.$$

Notice that by CANTOR'S THEOREM 3.18, for every cardinal κ there is a cardinal $\lambda > \kappa$, in particular, for every cardinal κ, $\bigcap\{\alpha \in \Omega : \kappa < |\alpha|\}$ is non-empty and therefore κ^+ exists.

A cardinal μ is called a **successor cardinal** if there exists a cardinal κ such that $\mu = \kappa^+$; otherwise, it is called a **limit cardinal**. In particular, every positive number $n \in \omega$ is a successor cardinal and ω is the smallest non-zero limit cardinal. By induction on $\alpha \in \Omega$ we define $\omega_{\alpha+1} := \omega_\alpha^+$, where $\omega_0 := \omega$, and $\omega_\alpha := \bigcup_{\delta \in \alpha} \omega_\delta$ for limit ordinals α; notice that $\bigcup_{\delta \in \alpha} \omega_\delta$ is a cardinal. In particular, ω_ω is the smallest uncountable limit cardinal and $\omega_1 = \omega_0^+$ is the smallest uncountable cardinal. Further, the collection $\{\omega_\alpha : \alpha \in \Omega\}$ is the class of all infinite cardinals, *i.e.*, for every infinite cardinal κ there is an $\alpha \in \Omega$ such that $\kappa = \omega_\alpha$. Notice that the collection of cardinals is—like the collection of ordinals—a proper *class* and not a *set*.

Cardinal addition, multiplication, and exponentiation are defined as follows:

Cardinal addition: For cardinals κ and μ, let $\kappa + \mu := |(\kappa \times \{0\}) \dot\cup (\mu \times \{1\})|$.

Cardinal multiplication: For cardinals κ and μ, let $\kappa \cdot \mu := |\kappa \times \mu|$.

Cardinal exponentiation: For cardinals κ and μ, let $\kappa^\mu := |{}^\mu\kappa|$.

Since for any set A, $|{}^{A}2| = |\mathscr{P}(A)|$, the cardinality of the power set of a cardinal κ is usually denoted by 2^κ. However, because 2^ω is the cardinality of the so-called *continuum* $\mathbb{R}$, it is usually denoted by $\mathfrak{c}$. Notice that by CANTOR'S THEOREM 3.18 for all cardinals κ we have $\kappa < 2^\kappa$.

As a consequence of the definition we get the following

FACT 3.24. *Addition and multiplication of cardinals is associative and commutative and we have the distributive law for multiplication over addition, and for all cardinals κ, λ, μ, we have*

$$\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu, \quad \kappa^{\mu\cdot\lambda} = (\kappa^\lambda)^\mu, \quad (\kappa\cdot\lambda)^\mu = \kappa^\mu \cdot \lambda^\mu.$$

Proof. It is obvious that addition and multiplication is associative and commutative and that we have the distributive law for multiplication over addition. Now, let κ, λ, μ, be any cardinal numbers. First, for every function $f : (\lambda\times\{0\})\cup(\mu\times\{1\}) \to \kappa$ let the functions $f_\lambda : (\lambda\times\{0\}) \to \kappa$ and $f_\mu : (\mu\times\{1\}) \to \kappa$ be such that for each $x \in (\lambda\times\{0\})\cup(\mu\times\{1\})$,

$$f(x) = \begin{cases} f_\lambda(x) & \text{if } x \in \lambda\times\{0\}, \\ f_\mu(x) & \text{if } x \in \mu\times\{1\}. \end{cases}$$

It is easy to see that each function $f : (\lambda\times\{0\})\cup(\mu\times\{1\}) \to \kappa$ corresponds to a unique pair $\langle f_\lambda, f_\mu\rangle$, and vice versa, each pair $\langle f_\lambda, f_\mu\rangle$ defines uniquely a function $f : (\lambda\times\{0\})\cup(\mu\times\{1\}) \to \kappa$. Thus, we have a bijection between $\kappa^{\lambda+\mu}$ and $\kappa^\lambda\cdot\kappa^\mu$.

Secondly, for every function $f : \mu \to {}^{\lambda}\kappa$, let $\tilde{f} : \mu\times\lambda \to \kappa$ be such that for all $\alpha \in \mu$ and all $\beta \in \lambda$ we have

$$\tilde{f}\big(\langle\alpha,\beta\rangle\big) = f(\alpha)(\beta).$$

We leave it as an exercise to the reader to verify that the mapping

$$\begin{aligned} {}^{\mu}\big({}^{\lambda}\kappa\big) &\longrightarrow {}^{\mu\times\lambda}\kappa \\ f &\longmapsto \tilde{f} \end{aligned}$$

is bijective, and therefore we have $\kappa^{\mu\cdot\lambda} = (\kappa^\lambda)^\mu$.

Thirdly, for every function $f : \mu \to \kappa\times\lambda$ let the functions $f_\kappa : \mu \to \kappa$ and $f_\lambda : \mu \to \lambda$ be such that for each $\alpha \in \mu$, $f(\alpha) = \langle f_\kappa(\alpha), f_\lambda(\alpha)\rangle$. We leave it again as an exercise to the reader to show that the mapping

$$\begin{aligned} {}^{\mu}(\kappa\times\lambda) &\longrightarrow {}^{\mu}\kappa \times {}^{\mu}\lambda \\ f &\longmapsto \langle f_\kappa, f_\lambda\rangle \end{aligned}$$

is a bijection. $\dashv$

The next result shows that addition and multiplication of infinite cardinals is quite simple:

THEOREM 3.25. *For any ordinal numbers $\alpha, \beta \in \Omega$ we have*

$$\omega_\alpha + \omega_\beta = \omega_\alpha \cdot \omega_\beta = \omega_{\alpha \cup \beta} = \max\{\omega_\alpha, \omega_\beta\}.$$

In particular, for every infinite cardinal κ we have $\kappa^2 = \kappa$.

Proof. It is enough to show that for all $\alpha \in \Omega$ we have $\omega_\alpha \cdot \omega_\alpha = \omega_\alpha$. For $\alpha = 0$ we already know that $|\omega \times \omega| = \omega$, thus, $\omega_0 \cdot \omega_0 = \omega_0$. Assume towards a contradiction that there exists an $\alpha \in \Omega$ such that $\omega_\alpha \cdot \omega_\alpha > \omega_\alpha$. Then there exists a least ordinal α_0 with this property, *i.e.*,

$$\alpha_0 = \bigcap\{\alpha \in \Omega : \omega_\alpha \cdot \omega_\alpha > \omega_\alpha\}.$$

On $\omega_{\alpha_0} \times \omega_{\alpha_0}$ we define an ordering "$<$" by stipulating

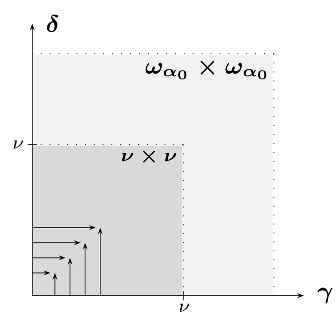
$$\langle \gamma_1, \delta_1 \rangle < \langle \gamma_2, \delta_2 \rangle \quad \iff \quad \begin{cases} (\gamma_1 \cup \delta_1) \in (\gamma_2 \cup \delta_2), \text{ or} \\ (\gamma_1 \cup \delta_1) = (\gamma_2 \cup \delta_2) \wedge \gamma_1 \in \gamma_2, \text{ or} \\ (\gamma_1 \cup \delta_1) = (\gamma_2 \cup \delta_2) \wedge \gamma_1 = \gamma_2 \wedge \delta_1 \in \delta_2. \end{cases}$$

With respect to the ordering "$<$", the first few elements of $\omega_{\alpha_0} \times \omega_{\alpha_0}$ are

$$\begin{aligned} &\langle 0,0 \rangle < \langle 0,1 \rangle < \langle 1,0 \rangle < \langle 1,1 \rangle \\ &\qquad < \langle 0,2 \rangle < \langle 1,2 \rangle < \langle 2,0 \rangle < \langle 2,1 \rangle < \langle 2,2 \rangle < \langle 0,3 \rangle < \cdots \end{aligned}$$

and in general, for $\alpha \in \beta \in \omega_{\alpha_0}$ we have $\langle \alpha, \beta \rangle < \langle \beta, \alpha \rangle$.

The ordering "$<$" on $\omega_{\alpha_0} \times \omega_{\alpha_0}$ is visualised by the following picture:



It is easily verified that "$<$" is a linear ordering on $\omega_{\alpha_0} \times \omega_{\alpha_0}$, and we leave it as an exercise to the reader to show that "$<$" is even a well-ordering.

Now, let $\eta \in \Omega$ be the order type of the well-ordering "$<$" on $\omega_{\alpha_0} \times \omega_{\alpha_0}$ and let $\Gamma : \eta \to \omega_{\alpha_0} \times \omega_{\alpha_0}$ be the unique order preserving bijection between η and $\omega_{\alpha_0} \times \omega_{\alpha_0}$. In particular, for any $\alpha, \alpha' \in \eta$ we get

$$\alpha \in \alpha' \quad \textit{if and only if} \quad \Gamma(\alpha) < \Gamma(\alpha') .$$

Because $\omega_{\alpha_0} < |\omega_{\alpha_0} \times \omega_{\alpha_0}|$ we have $\omega_{\alpha_0} < |\eta|$. Let now $\langle \gamma_0, \delta_0 \rangle := \Gamma(\omega_{\alpha_0})$. Then, since $\gamma_0, \delta_0 \in \omega_{\alpha_0}$, for $\nu = \max\{\gamma_0, \delta_0\}$ we have

$$|\nu| < \omega_{\alpha_0} \quad \textit{and} \quad \omega_{\alpha_0} \leq |\nu \times \nu| .$$

Thus, for $\omega_\beta = |\nu|$ we get $\omega_\beta < \omega_{\alpha_0}$ and $\omega_\beta \cdot \omega_\beta > \omega_\beta$, which is a contradiction to the choice of ω_{α_0}. ⊣

As a consequence of THEOREM 3.25 we get the following

COROLLARY 3.26. *If κ is an infinite cardinal, then*

(a) *for all $n \in \omega$, $\kappa^{n+1} = \kappa$,*

(b) $\operatorname{seq}(\kappa) = \kappa$, *and*

(c) $\kappa^\kappa = 2^\kappa$.

Proof. (a) The proof is by induction on $n \in \omega$: If $n = 0$, then by definition we have $\kappa^1 = \kappa$. If $n = 1$, by THEOREM 3.25 we get $\kappa^2 = \kappa$. Assume now that $\kappa^{n+1} = \kappa$ for some $n \in \omega$. Then $\kappa^{n+2} = \kappa \cdot \kappa^{n+1}$, which is, by assumption, equal to $\kappa \cdot \kappa$, and by THEOREM 3.25 we get $\kappa^{n+2} = \kappa$.

(b) Notice that $\operatorname{seq}(\kappa) = |\bigcup_{n\in\omega} \kappa^n| = 1 + \kappa + \kappa^2 + \ldots + \kappa^n + \ldots$, which is, by (a), equal to $1 + \kappa \cdot \omega = 1 + \kappa = \kappa$.

(c) Since κ is infinite, we obviously have $\kappa^\kappa = |{}^\kappa\kappa| \geq |{}^\kappa 2| = 2^\kappa$. On the other hand, by identifying each function $f \in {}^\kappa\kappa$ with its graph, which is a subset of $\kappa \times \kappa$, we get $|{}^\kappa\kappa| \leq |\mathscr{P}(\kappa \times \kappa)|$, and since $|\kappa \times \kappa| = \kappa$ we finally have $\kappa^\kappa \leq |\mathscr{P}(\kappa)| = 2^\kappa$. ⊣

Let λ be an infinite limit ordinal. A subset $\mathcal{C}$ of λ is called **cofinal** in λ if $\bigcup \mathcal{C} = \lambda$. The **cofinality** of λ, denoted $\operatorname{cf}(\lambda)$, is the cardinality of a smallest cofinal set $\mathcal{C} \subseteq \lambda$. In other words,

$$\operatorname{cf}(\lambda) = \min\bigl\{|\mathcal{C}| : \mathcal{C} \text{ is cofinal in } \lambda\bigr\}.$$

Notice that by definition, $\operatorname{cf}(\lambda)$ is always a cardinal number.

Let again λ be an infinite limit ordinal and let $\mathcal{C} = \{\beta_\xi : \xi \in \operatorname{cf}(\lambda)\} \subseteq \lambda$ be cofinal in λ. Now, for every $\nu \in \operatorname{cf}(\lambda)$ let $\alpha_\nu := \bigcup\{\beta_\xi : \xi \in \nu\}$ (notice that all the α_ν's belong to λ). Then $\langle \alpha_\nu : \nu \in \operatorname{cf}(\lambda)\rangle$ is an increasing sequence (not necessarily in the strict sense) of length $\operatorname{cf}(\lambda)$ with $\bigcup\{\alpha_\nu : \nu \in \operatorname{cf}(\lambda)\} = \lambda$. Thus, instead of cofinal *subsets* of λ we could equally well work with cofinal *sequences*.

Since every infinite cardinal is an infinite limit ordinal, $\mathrm{cf}(\kappa)$ is also defined for cardinals κ. An infinite *cardinal* κ is called **regular** if $\mathrm{cf}(\kappa) = \kappa$; otherwise, κ is called **singular**. For example, ω is regular and ω_ω is singular (since $\{\omega_n : n \in \omega\}$ is cofinal in ω_ω). In general, for non-zero limit ordinals λ we have $\mathrm{cf}(\omega_\lambda) = \mathrm{cf}(\lambda)$. For example, $\mathrm{cf}(\omega_\omega) = \mathrm{cf}(\omega_{\omega+\omega}) = \mathrm{cf}(\omega_{\omega_{\omega_\omega}}) = \omega$.

FACT 3.27. *For all infinite limit ordinals λ, the cardinal $\mathrm{cf}(\lambda)$ is regular.*

Proof. Let $\kappa = \mathrm{cf}(\lambda)$ and let $\langle \alpha_\xi : \xi \in \kappa \rangle$ be an increasing, cofinal sequence of λ. Further, let $\mathcal{C} \subseteq \kappa$ be cofinal in κ with $|\mathcal{C}| = \mathrm{cf}(\kappa)$. Now, $\langle \alpha_\nu : \nu \in \mathcal{C} \rangle$ is still a cofinal sequence of λ, which implies that $\mathrm{cf}(\lambda) \leq \mathrm{cf}(\kappa)$. On the other hand we have $\mathrm{cf}(\kappa) \leq \kappa = \mathrm{cf}(\lambda)$. Hence, $\mathrm{cf}(\kappa) = \kappa = \mathrm{cf}(\lambda)$, which shows that $\mathrm{cf}(\lambda)$ is regular. $\dashv$

The following result shows that all infinite successor cardinals are regular. As a matter of fact, we would like to mention that in the absence of the Axiom of Choice, ω_1 can be singular (see Chapter 17 | RELATED RESULT 95).

PROPOSITION 3.28. *If κ is an infinite cardinal, then κ^+ is regular.*

Proof. Assume towards a contradiction that there exists a subset $\mathcal{C} \subseteq \kappa^+$ such that $\mathcal{C}$ is cofinal in κ^+ and $|\mathcal{C}| < \kappa^+$, *i.e.*, $|\mathcal{C}| \leq \kappa$. Since $\mathcal{C} \subseteq \kappa^+$, for every $\alpha \in \mathcal{C}$ we have $|\alpha| \leq \kappa$. Now, by AC, for each $\alpha \in \mathcal{C}$ we can choose a one-to-one mapping $f_\alpha : \alpha \hookrightarrow \kappa$. In particular, for each $\alpha \in \mathcal{C}$ and each $\nu \in \alpha$, $f_\alpha(\nu) \in \kappa$. Furthermore, let $g : \mathcal{C} \hookrightarrow \kappa$ be a one-to-one mapping from $\mathcal{C}$ into κ; in particular, for each $\alpha \in \mathcal{C}$, $g(\alpha) \in \kappa$. Then,

$$\left\{ \left\langle g(\alpha), f_\alpha(\nu) \right\rangle : \alpha \in \mathcal{C} \wedge \nu \in \alpha \right\}$$

is a subset of $\kappa \times \kappa$, which shows that $|\bigcup \mathcal{C}| \leq |\kappa \times \kappa| = \kappa$. Thus, $\bigcup \mathcal{C} \neq \kappa^+$, which implies that $\mathcal{C}$ is not cofinal in κ^+. $\dashv$

For example, ω_1, ω_{17}, and $\omega_{\omega+5}$ are regular, since $\omega_1 = \omega_0^+$, $\omega_{17} = \omega_{16}^+$, and $\omega_{\omega+5} = \omega_{\omega+4}^+$.

We now consider arbitrary sums and products of cardinal numbers. For this, let I be a non-empty set and let $\{\kappa_\iota : \iota \in I\}$ be a family of cardinals. We define

$$\sum_{\iota \in I} \kappa_\iota := \left| \bigcup_{\iota \in I} A_\iota \right|$$

where $\{A_\iota : \iota \in I\}$ is a family of pairwise disjoint sets such that $|A_\iota| = \kappa_\iota$ for each $\iota \in I$, e.g., $A_\iota = \kappa_\iota \times \{\iota\}$ will do.

Similarly we define

$$\prod_{\iota\in I}\kappa_\iota := \Big|\bigtimes_{\iota\in I} A_\iota\Big|$$

where $\{A_\iota : \iota \in I\}$ is a family of sets such that $|A_\iota| = \kappa_\iota$ for each $\iota \in I$, *e.g.*, $A_\iota = \kappa_\iota$ will do.

THEOREM 3.29 (INEQUALITY OF KÖNIG–JOURDAIN–ZERMELO). *Let I be a non-empty set and let $\{\kappa_\iota : \iota \in I\}$ and $\{\lambda_\iota : \iota \in I\}$ be families of cardinal numbers such that $\kappa_\iota < \lambda_\iota$ for every $\iota \in I$. Then*

$$\sum_{\iota\in I}\kappa_\iota < \prod_{\iota\in I}\lambda_\iota .$$

Proof. Let $\{A_\iota : \iota \in I\}$ be a family of pairwise disjoint sets such that $|A_\iota| = \kappa_\iota$ for each $\iota \in I$. First, for each $\iota \in I$ choose a injection $f_\iota : A_\iota \hookrightarrow \lambda_\iota$ and an element $y_\iota \in \lambda_\iota \setminus f_\iota[A_\iota]$. Notice that since $|A_\iota| < \lambda_\iota$, the set $\lambda_\iota \setminus f_\iota[A_\iota]$ is non-empty, so for each $\iota \in I$ there is a y_ι as required.

In a first step we show that $\sum_{\iota\in I}\kappa_\iota \le \prod_{\iota\in I}\lambda_\iota$: For each $\iota \in I$ define the function $\bar f_\iota : \bigcup_{\iota\in I} A_\iota \to \lambda_\iota$ by stipulating

$$\bar f_\iota(x) = \begin{cases} f_\iota(x) & \text{if } x \in A_\iota, \\ y_\iota & \text{otherwise.} \end{cases}$$

and let

$$\begin{aligned} \bar f :\ \textstyle\bigcup_{\iota\in I} A_\iota &\longrightarrow \quad \textstyle\bigtimes_{\iota\in I}\lambda_\iota \\ x \quad &\longmapsto \langle \bar f_\iota(x) : \iota \in I\rangle \end{aligned}$$

Then $\bar f$ is obviously a one-to-one function from $\bigcup_{\iota\in I} A_\iota$ into $\bigtimes_{\iota\in I}\lambda_\iota$, which shows that $\sum_{\iota\in I}\kappa_\iota \le \prod_{\iota\in I}\lambda_\iota$.

To prove that $\sum_{\iota\in I}\kappa_\iota < \prod_{\iota\in I}\lambda_\iota$, it is enough to show that there is no bijection between $\bigcup_{\iota\in I} A_\iota$ and $\bigtimes_{\iota\in I}\lambda_\iota$. For this, take any function $g : \bigcup_{\iota\in I} A_\iota \to \bigtimes_{\iota\in I}\lambda_\iota$, and for every $\iota \in I$, let $P_\iota(g[A_\iota])$ be the projection of $g[A_\iota]$ on λ_ι. Then, since $|A_\iota| < \lambda_\iota$, for each $\iota \in I$ we can choose an element $z_\iota \in \lambda_\iota \setminus P_\iota(g[A_\iota])$. Evidently, the sequence $\langle z_\iota : \iota \in I\rangle$ does not belong to $g[\bigcup_{\iota\in I} A_\iota]$ which shows that g is not surjective, and consequently, g is not bijective. $\dashv$

As an immediate consequence we get the following

COROLLARY 3.30. *For every infinite cardinal κ we have*

$$\kappa < \kappa^{\mathrm{cf}(\kappa)} \quad \textit{and} \quad \mathrm{cf}(2^\kappa) > \kappa .$$

In particular, $\mathrm{cf}(\mathfrak{c}) > \omega$.

Proof. Let $\langle \alpha_\nu : \nu \in \mathrm{cf}(\kappa)\rangle$ be a cofinal sequence of κ. On the one hand we have

$$\kappa = \Big|\bigcup_{\nu\in\mathrm{cf}(\kappa)} \alpha_\nu\Big| \leq \sum_{\nu\in\mathrm{cf}(\kappa)} |\alpha_\nu| \leq \mathrm{cf}(\kappa)\cdot\kappa = \kappa,$$

and hence, $\kappa = \sum_{\nu\in\mathrm{cf}(\kappa)} |\alpha_\nu|$. On the other hand, for each $\nu \in \mathrm{cf}(\kappa)$ we have $|\alpha_\nu| < \kappa$, and therefore, by THEOREM 3.29, we have

$$\sum_{\nu\in\mathrm{cf}(\kappa)} |\alpha_\nu| < \prod_{\nu\in\mathrm{cf}(\kappa)} \kappa = \kappa^{\mathrm{cf}(\kappa)}.$$

Thus, we have $\kappa < \kappa^{\mathrm{cf}(\kappa)}$.

In order to see that $\mathrm{cf}(2^\kappa) > \kappa$, assume towards a contradiction that $\mathrm{cf}(2^\kappa) \leq \kappa$. Now,

$$\mathrm{cf}(2^\kappa) \leq \kappa \quad \text{implies} \quad (2^\kappa)^{\mathrm{cf}(2^\kappa)} \leq (2^\kappa)^\kappa$$

and since $(2^\kappa)^\kappa = 2^{\kappa\cdot\kappa} = 2^\kappa$, we get

$$(2^\kappa)^{\mathrm{cf}(2^\kappa)} \leq 2^\kappa\,,$$

which contradicts the fact that $2^\kappa < (2^\kappa)^{\mathrm{cf}(2^\kappa)}$.

To see that $\mathrm{cf}(\mathfrak{c}) > \omega$, recall that $\mathfrak{c} = 2^\omega$. So, by the previous result we obtain $\mathrm{cf}(2^\omega) > \omega$. ⊣

NOTES

Some of the papers mentioned below, or at least their translation into English, can be found in the collection [96] edited by van Heijenoort (whose biography is written by Feferman [26]).

Let us discuss the development of Set Theory: To some extent, Set Theory is the theory of infinite sets; but, what is the infinite and does it exist?

The Infinite. As mentioned before, there are two different kinds of infinite, namely the *actual infinite* and the *potential infinite*. To illustrate the difference, let us consider the collection of prime numbers. Euclid proved that for any prime number p there is a prime number p' which is larger than p (see [22, Book IX]). This shows that there are arbitrarily many prime numbers, and therefore, the collection of primes is "potentially" infinite. However, he did not claim that the collection of all prime numbers as a whole "actually" exists. (The difference between actual and potential infinite is discussed in greater detail, for example, in Bernays [4, Teil II].)

Two quite similar attempts to prove the objective existence of the (actual) infinite are due to Bolzano [9, 10§13] and Dedekind [19§5, No. 66], and both are similar to the approach suggested in Plato's *Parmenides* [80*132a-b*] (for a philosophical view of the notion of infinity we refer the reader to Mancosu [65]). However, Russell [84, Chapter XIII, p. 139 ff.] (see also [86, Chapter XLIII]) shows that these attempts must fail. Moreover, he demonstrates that the infinite is neither self-contradictory nor demonstrable logically and writes that *we must conclude that nothing can be known a priori as to whether the number of things in the world is finite or infinite. The conclusion is, therefore, to adopt a Leibnizian phraseology, that some of the possible worlds are finite, some infinite, and we have no means of knowing to which of these two kinds our actual world belongs. The axiom of infinity will be true in some possible worlds and false in others; whether it is true or false in this world, we cannot tell* (*cf.* [84, p. 141]).

If the infinite exists, the problem still remains how one would recognise infinite sets, or in other words, how one would define the predicate "infinite". Dedekind provided a definition in [19§5, No. 64], which is—as Schröder [88, p. 303 f.] pointed out—equivalent to the definition given three years earlier by Peirce (*cf.* [77, p. 202] or [3, p. 51]). However, the fact that an infinite set can be mapped injectively into a proper subset of itself—which is the key idea of Dedekind's definition of infinite sets—was already discovered and clearly explained about 250 years earlier by Galilei (see [36, First Day]). Another definition of the infinite—which will be compared with Dedekind's definition in Chapter 8—can be found in von Neumann [72, p. 736]. More definitions of finiteness, as well as their dependencies, can be found for example in Lévy [62] and in Spišiak and Vojtáš [93].

Birth of Set Theory. As mentioned above, the birth of Set Theory dates to 1873 when Cantor proved that the set of real numbers is uncountable. One could even argue that the exact birth date is 7 December 1873, the date of Cantor's letter to Dedekind informing him of his discovery.

Cantor's first proof that there is no bijection between the set of real numbers and the set of natural numbers used an argument with nested intervals (*cf.* [13§2] or [18, p. 117]). Later, he improved the result by showing that $2^{\mathfrak{m}} > \mathfrak{m}$ for every cardinal $\mathfrak{m}$ (*cf.* [15] or [18, III.8]), which is nowadays called CANTOR'S THEOREM. The argument used in the proof of PROPOSITION 3.17—which is in fact just a special case of CANTOR'S THEOREM—is sometimes called *Cantor's diagonal argument*. The word "diagonal" comes from the diagonal process used in the proofs of PROPOSITION 3.17 and CANTOR'S THEOREM. The diagonal process is a technique of constructing a new member of a set of lists which is distinct from all members of a given list. This is done by first arranging the list as a matrix, whose *diagonal* gives information about the x^{th} term of the x^{th} row of the matrix. Then, by changing each term of the diagonal, we get a new list which is distinct from every row of the matrix (see also Kleene [58§2]).

For a brief biography of Cantor and for the development of Set Theory see, for example, Fraenkel [33], Schoenflies [87], and Kanamori [53].

Russell's Paradox. The fact that a naïve approach to the notion of "set" leads to contradictions was discovered by Russell in June 1901 while he was working on his *Principles of Mathematics* [86] (see also Grattan-Guinness [41]). When Russell published his discovery, other mathematicians and set-theorists like Zermelo (see [101, footnote p. 118 f.] or Rang and Thomas [81]) had already been aware of this antinomy, which—according to Hilbert—*had a downright catastrophic effect when it became known throughout the world of Mathematics* (*cf.* [46, p. 169] or [47, p. 190]). However, Russell was the first to discuss the contradiction at length in his published works, the first to attempt to formulate solutions and the first to fully appreciate its importance. For example, the entire Chapter X of [86] was dedicated to discussing this paradox (in particular, see [86, Chapter X, §102]). In order to prevent the emergence of antinomies and paradoxes in Set Theory and in Logic in general, Russell developed in [86, Appendix B] (see also [83]) his theory of logical types which rules out self-reference. According to this theory, self-referential statements are neither true nor false, but meaningless.

Russell's Paradox as well as some other antinomies can also be found in Fraenkel, Bar-Hillel, and Lévy [28, Chapter I].

Axiomatisation of Set Theory. In 1908, Zermelo published in [102] his first axiomatic system consisting of seven axioms, which he called:

1. Axiom der Bestimmtheit
   which corresponds to the Axiom of Extensionality
2. Axiom der Elementarmengen
   which includes the Axiom of Empty Set as well as the Axiom of Pairing
3. Axiom der Aussonderung
   which corresponds to the Axiom Schema of Separation
4. Axiom der Potenzmenge
   which corresponds to the Axiom of Power Set
5. Axiom der Vereinigung
   which corresponds to the Axiom of Union
6. Axiom der Auswahl
   which corresponds to the Axiom of Choice
7. Axiom des Unendlichen
   which corresponds to the Axiom of Infinity

In 1930, Zermelo presented in [103] his second axiomatic system, which he called the ZF-system, in which he incorporated ideas of Fraenkel [30], Skolem [91], and von Neumann citevNeumannI,vNeumannII,vNeumannIII (see also Zermelo [99]). In fact, he added the Axiom Schema of Replacement and the Axiom of Foundation to his former system, cancelled the Axiom of Infinity (since he thought that it does not belong to the general theory of sets), and did not explicitly mention the Axiom of

Choice (because of its different character and since he considered it as a general logical principle). For Zermelo's published work in Set Theory, described and analysed in its historical context, see Zermelo [104], Kanamori [56] and Ebbinghaus [21].

The need for the Axiom Schema of Replacement was first noticed by Fraenkel (see [104, p. 23]) who introduced a certain form of it in [30] (another form of it he gave in [29, Definition 2, p. 158]). However, the present form was introduced by von Neumann [73] (see the note below on the TRANSFINITE RECURSION THEOREM). As a matter of fact we would like to mention that the Axiom Schema of Replacement was already used implicitly by Cantor in 1899 (*cf.* [18, p. 444, line 3]). Besides Fraenkel, Skolem also realised that Zermelo's first axiomatic system was not sufficient to provide a complete foundation for the usual theory of sets and introduced—independently of Fraenkel—in 1922 the Axiom Schema of Replacement (see [91] or [92, p. 145 f.]). In [91], he also gave a proper definition of the notion "definite proposition" and, based on a theorem of Löwenheim [64], he discovered the following fact [92, p. 139] (stated in Chapter 16 as the LÖWENHEIM–SKOLEM THEOREM 16.1): *If the axioms are consistent, there exists a domain in which the axioms hold and whose elements can all be enumerated by means of the positive finite integers.* At first glance this looks strange, since we know, for example, that the set of real numbers is uncountable. However, this so-called Skolem Paradox—which we will meet in a slightly different form in Chapter 16—is not a paradox in the sense of an antinomy, it is just a somewhat unexpected feature of formal systems (see also Kleene [58, p. 426 f.] and von Plato [97]).

Concerning the terminology we would like to mention that the definition of ordered pairs given above was introduced by Kuratowski [61, Définition V, p. 171] (compare with Hausdorff [44, p. 32] and see also Kanamori [55§5]), and that the infinite set which corresponds to $\omega = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots\}$ was introduced by von Neumann [70]. For more historical background, see Bachmann [2] or Fraenkel [6, Part I], and for a brief discussion of the axiom systems of von Neumann, Bernays, and Gödel, see Fraenkel [6, Part I, Section 7].

The Axiom of Foundation. As mentioned above, Zermelo introduced this axiom in his second axiomatisation of Set Theory in 1930, but it goes back to von Neumann (*cf.* [71, p. 239] and [74, p. 231]), and in fact, the idea can already be found in Mirimanoff [67, 68]: For example in [67, p. 211] he calls a set x *regular* (French "ordinaire") if every descending sequence $x \ni x_1 \ni x_2 \ni \ldots$ is finite. However, he did not postulate the regularity of sets as an axiom, but if one would do so, one would get the Axiom of Regularity saying that every set is regular. Now, as a consequence of the Axiom of Foundation we got the fact that there are no infinite descending sequences of the form $x_1 \ni x_2 \ni \ldots \ni x_i \ldots$, which just tells us that every set is regular. Thus, the Axiom of Foundation implies the Axiom of Regularity. The converse is not true, unless we assume some non-trivial form of the Axiom of Choice (see Mendelson [66]). As a matter of fact, we would like to mention that Zermelo, when he formulated the Axiom of Foundation in [103], gave both definitions and just mentioned (without proof) that they are equivalent.

Ordinal Numbers. The theory of ordinals was first developed in an axiomatic way by von Neumann in [70] (see also [71, 72, 73]). For an alternative axiomatic approach to ordinals, independently of ordered sets and types, see Tarski [95] or Lindenbaum and Tarski [63]. For some more definitions of ordinals, see Bachmann [2, p. 24].

The Transfinite Recursion Theorem. The Transfinite Recursion Theorem was first formulated and proved by von Neumann [73], who also pointed out that, in addition to Zermelo's axioms, the Axiom Schema of Replacement also has to be used. Even though a certain form of the Axiom Schema of Replacement was already given by Fraenkel (see above), von Neumann showed that Fraenkel's notion of function is not sufficient to prove the TRANSFINITE RECURSION THEOREM. Moreover, he showed (*cf.* [73, I.3]) that Fraenkel's version of the Axiom Schema of Replacement given in [31§1] follows from the other axioms given there (see also Fraenkel's note [32]).

The Cantor–Bernstein Theorem. This theorem, unfortunately also known as the SCHRÖDER–BERNSTEIN THEOREM, was first stated and proved by Cantor (*cf.*[14, VIII.4] or [18, p. 413], and [16§2, Satz B] or [18, p. 285]). In order to prove this theorem, Cantor used the Trichotomy of Cardinals, which is—as we will see in Chapter 6—equivalent to the Axiom of Choice (see also [18, p. 351, Anm. 2]). An alternative proof, avoiding any form of the Axiom of Choice, was found by Bernstein, who was initially a student of Cantor's. Bernstein presented his proof around Easter 1897 in one of Cantor's seminars in Halle, and the result was published in 1898 in Borel [11, p. 103–106] About the same time, Schröder gave a similar proof in [88] (submitted May 1896), but unfortunately, Schröder's proof was flawed by an irreparable error. While other mathematicians regarded his proof as correct, Korselt wrote to Schröder about the error in 1902. In his reply, Schröder admitted his mistake which he had already found some time ago but did not have the opportunity to make public. A few weeks later, Korselt submitted the paper [60]—which appeared almost a decade later—with a proof of the CANTOR–BERNSTEIN THEOREM which is quite different from the one given by Bernstein. A proof of the CANTOR–BERNSTEIN THEOREM, similar to Korselt's proof, was found in 1906 independently by Peano [76] and Zermelo (see [102, footnote p. 272 f.]). However, they could not know that they had just rediscovered the proof that had already been obtained twice by Dedekind in 1887 and 1897, since Dedekind's proof—in our terminology given above—was not published until 1932 (see [20, LXII & Erl. p. 448] and [18, p. 449]). More about the CANTOR–BERNSTEIN THEOREM can be found in Hinkis [48].

Cantor products. Motivated by a result due to Euler on partition numbers (*cf.* [23, Caput XVI]), Cantor showed in [12] (see also [18, pp. 43–50]) that every real number $r > 1$ can be written in a unique way as a product of the form $\prod_{n\in\omega}(1+\frac{1}{q_n})$, where all q_n's are positive integers and $q_{n+1} \geq q_n^2$. He also showed that $r = \prod_{n\in\omega}(1+\frac{1}{q_n})$ is rational if and only if there is an $m \in \omega$ such that for all $n \geq m$ we have $q_{n+1} = q_n^2$, and further he gave the representation of the square roots of some small natural numbers. For example, the q_n's in the representation of $\sqrt{2}$ are $q_0 = 3$ and $q_{n+1} = 2q_n^2 - 1$. More about Cantor products can be found, for example, in Perron [78§35].

Cardinal Numbers. The concept of cardinal number is one of the most fundamental concepts in Set Theory. Cantor describes cardinal numbers as follows (*cf.*[16§1] or [18, p. 282 f.]): *The general concept which with the aid of our active intelligence results from a set M, when we abstract from the nature of its various elements and from the order of their being given, we call the "power" or "cardinal number" of M.* This double abstraction suggests his notation $\overline{\overline{M}}$ for the cardinality of M. As mentioned above, one can define the cardinal number of a set M as an object $\overline{\overline{M}}$ which consists of all those sets (including M itself) which have the same cardinality as M. This approach, which was for example taken by Frege (*cf.* [34, 35]), and Russell (*cf.* [82, p. 378] or [83, Section IX, p. 256]), has the advantage that it can be carried out in naïve Set Theory (see also Kleene [58, p. 9]). However, it has the disadvantage that for every non-empty set M, the object $\overline{\overline{M}}$ is a proper class and therefore does not belong to the set-theoretic universe.

Hartogs' Theorem. The proof of HARTOGS' THEOREM is taken from Hartogs [43]. In that paper, Hartogs' main motivation was to find a proof for Zermelo's Well-Ordering Principle which does not make use of the Axiom of Choice. However, since the Well-Ordering Principle and the Axiom of Choice are equivalent, he had to assume something similar, which he had done assuming explicitly Trichotomy of Cardinals. These principles will be discussed in greater detail in Chapter 6.

In 1935, Hartogs was forced to retire from his position in Munich, where he committed suicide in August 1943 because he could no longer bear the continuous humiliations by the Nazis.

The Axiom of Choice. Fraenkel writes in [28, p. 56 f.] that *the* Axiom of Choice *is probably the most interesting and, in spite of its late appearance, the most discussed axiom of Mathematics, second only to Euclid's axiom of parallels which was introduced more than two thousand years ago.* We would also like to mention a different view of choice functions, namely the view of Peano. In 1890, Peano published a proof in which he was constrained to choose a single element from each set in a certain infinite sequence $A_1, A_2, \ldots$ of infinite subsets of $\mathbb{R}$. In that proof, he remarked carefully (*cf.* [75, p. 210]): *But as one cannot apply infinitely many times an arbitrary rule by which one assigns to a class A an individual of this class, a determinate rule is stated here, by which, under suitable hypotheses, one assigns to each class A an individual of this class.* To obtain his rule, he employed least upper bounds. According to Moore [69, p. 76], Peano was the first mathematician who—while accepting infinite collections—categorically rejected the use of infinitely many arbitrary choices.

The difficulty is well illustrated by a Russellian anecdote (*cf.* Sierpiński [90, p. 125]): *A millionaire possesses an infinite number of pairs of shoes, and an in-*

finite number of pairs of socks. One day, in a fit of eccentricity, he summons his valet and asks him to select one shoe from each pair. When the valet, accustomed to receiving precise instructions, asks for details as to how to perform the selection, the millionaire suggests that the left shoe be chosen from each pair. Next day the millionaire proposes to the valet that he select one sock from each pair. When asked as to how this operation is to be carried out, the millionaire is at a loss for a reply, since, unlike shoes, there is no intrinsic way of distinguishing one sock of a pair from the other. In other words, the selection of the socks cannot be carried out without the aid of some choice function.

As long as the implicit and unconscious use of the Axiom of Choice by Cantor and others involved only generalised arithmetical concepts and properties well-known from finite numbers, nobody took offence. However, the situation changed drastically after Zermelo [100] published his first proof (essentially the proof of PROPOSITION 3.20) that every set can be well-ordered—which was one of the earliest assertions of Cantor. It is worth mentioning that, according to Zermelo [100, p. 514] and [101, footnote p. 118], it was in fact Erhard Schmidt's idea to use the Axiom of Choice in order to build the f-sets. Zermelo considered the Axiom of Choice as a *logical principle*, that *cannot be reduced to a still simpler one, but is used everywhere in mathematical deductions without hesitation* (see [100, p. 516]). Even though in Zermelo's view the Axiom of Choice was "self-evident", which is not the same as "obvious" (see Shapiro [89§5] for a detailed discussion of the meaning of "self-evidence"), not all mathematicians at that time shared Zermelo's opinion. Moreover, after the first proof of the Well-Ordering Principle was published in 1904, the mathematical journals (especially volume 60 of *Mathematische Annalen*) were flooded with critical notes rejecting the proof (see, for example, Moore [69, Chapter 2]), mostly arguing that the Axiom of Choice was either illegitimate or meaningless (*cf.* Fraenkel, Bar-Hillel, and Lévy [28, p. 82]). The reason for this was not only due to the non-constructive character of the Axiom of Choice, but also because it was not yet clear what a "set" should be. So, Zermelo decided to publish a more detailed proof, essentially the proof of THEOREM 3.23, and at the same time taking the opportunity to reply to his critics. This resulted in [101], his second proof of the Well-Ordering Principle, which was published in 1908, the same year as he presented his first axiomatisation of Set Theory in [101]. It seems that this was not a coincidence. Moore [69, p. 159] writes that *Zermelo's axiomatisation was primarily motivated by a desire to secure his demonstration of the* Well-Ordering Principle *and, in particular, to save his* Axiom of Choice. Moreover, Hallett [42, p. xvi] goes even further by trying to show *that the selection of the axioms themselves was guided by the demands of Zermelo's reconstructed [second] proof.* Hallett's statement is motivated by a remark on page 124 in Zermelo [101], where he emphasises that the proof is just based on certain fixed principles to *build initial sets* and to *derive new sets from given ones*—exactly what we would require for principles to form an axiomatic system of Set Theory.

We would like to mention that because of its different character (*cf.* Bernays [5]) and since he considered the Axiom of Choice as a general logical principle, he did not include the Axiom of Choice in his second axiomatic system of Set Theory.

For a comprehensive survey of Zermelo's Axiom of Choice, its origins, development, and influence, we refer the reader to Moore [69] (see also Kanamori [54], Jech [49], and Fraenkel, Bar-Hillel, and Lévy [28, Chapter II, §4]); and for a biography of Zermelo (including the history of AC and axiomatic Set Theory) we refer the reader to Ebbinghaus [21].

Gödel's Constructible Universe. According to Kanamori [52, p. 28 ff.], in October of 1935 Gödel informed von Neumann at the Institute for Advanced Study in Princeton that he had established the relative consistency of the Axiom of Choice. Gödel established consistency by devising his constructible hierarchy **L** (originally denoted by the old German script "$\mathcal{L}$") and verifying the Axiom of Choice and ZF axioms there. Furthermore, Gödel conjectured that the Continuum Hypothesis would also hold in **L**, but he soon fell ill and only gave a proof of this and the Generalised Continuum Hypothesis (*i.e.*, for all $\alpha \in \Omega$, $2^{\omega_\alpha} = \omega_{\alpha+1}$) two years later. The crucial idea apparently came to him during the night of June 14/15, 1937 (see also [40, pp. 1–8]).

Gödel's article [37] was the first announcement of these results, in which he describes the model **L** as the class *of all "mathematically constructible" sets, where the term "constructible" is to be understood in the semi-intuitionistic sense which excludes impredicative procedures. This means "constructible" sets are defined to be those sets which can be obtained by Russell's ramified hierarchy of types, if extended to include transfinite orders.* In the succeeding article [38], Gödel provided more details in the context of ZF, and in his monograph [39]—based on lectures given at the Institute for Advanced Study during the winter of 1938/39—Gödel gave another presentation of **L**. This time he generated **L** set by set with a transfinite recursion in terms of eight elementary set generators, a sort of Gödel numbering into the transfinite (*cf.* Kanamori [52, p. 30], and for Gödel's work in Set Theory, see Kanamori [57]).

***Cardinal Arithmetic in the Presence of* AC.** The definition of cardinals given above can also be found, for example, in von Neumann [72, VII.2, p. 731].

The first proof of THEOREM 3.25 appeared in Hessenberg [45, p. 593] (see also Jourdain [51]).

Regularity of cardinals was investigated by Hausdorff, who also raised the question of the existence of regular limit cardinals (*cf.* [44, p. 131]).

The INEQUALITY OF KÖNIG–JOURDAIN–ZERMELO 3.29—also known as König's Theorem—was proven by König [59] (but only for countable sums and products), and independently by Jourdain [50] and by Zermelo [102] (for historical facts, see Moore [69, p. 154] and Fraenkel [27, p. 98]). Obviously, the INEQUALITY OF KÖNIG–JOURDAIN–ZERMELO implies the Axiom of Choice (since it guarantees that every Cartesian product of non-empty sets is non-empty), and consequently we get that the INEQUALITY OF KÖNIG–JOURDAIN–ZERMELO is *equivalent* to the Axiom of Choice.

REFERENCES

1. ARISTOTLE, ***Physics***, Athens, *published by Andronikos of Rhodos around* 40 B.C.
2. HEINZ BACHMANN, ***Transfinite Zahlen***, Springer-Verlag, Berlin · Heidelberg, 1967.
3. KAREL BERKA AND LOTHAR KREISER, ***Logik-Texte.*** *Kommentierte Auswahl zur Geschichte der modernen Logik*, Akademie-Verlag, Berlin, 1971.
4. PAUL BERNAYS, *Die Philosophie der Mathematik und die Hilbertsche Beweistheorie*, ***Blätter für deutsche Philosophie***, vol. 4 (1930), 326–367 (also published in [7]).
5. ______, *Sur le platonisme dans les mathématiques*, ***L'Enseignement Mathématique***, vol. 34 (1935), 52–69 (see [8]/[7] for a translation into English/German).
6. ______, ***Axiomatic Set Theory.*** *With a historical introduction by Abraham A. Fraenkel*, 2nd ed., [Studies in Logic and the Foundations of Mathematics], North-Holland, Amsterdam, 1968 [reprint: Dover Publications, New York, 1991].
7. ______, ***Abhandlungen zur Philosophie der Mathematik***, Wissenschaftliche Buchgesellschaft, Darmstadt, 1976.
8. ______, *On platonism in mathematics*, *in* ***Philosophy of Mathematics. Selected Readings*** (Paul Benacerraf and Hilary Putnam, eds.), Cambridge University Press, Cambridge, 2nd ed., 1983, pp. 258–271.
9. BERNARD BOLZANO, ***Paradoxien des Unendlichen***, Hrsg. aus dem schriftlichen Nachlasse des Verfassers von Fr. Přihonský, C. H. Reclam Sen., Leipzig, 1851.
10. ______, ***Paradoxes of the Infinite***, [translated by D. A. Steele], Routledge and Kegan Paul, London, 1950.
11. ÉMILE BOREL, ***Leçons sur la Théorie des Fonctions***, Gauthier-Villars et Fils, Paris, 1898.
12. GEORG CANTOR, *Zwei Sätze über eine gewisse Zerlegung der Zahlen in unendliche Produkte*, ***Zeitschrift für Mathematik und Physik***, vol. 14 (1869), 152–158.
13. ______, *Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen*, ***Journal für die reine und angewandte Mathematik (Crelle)***, vol. 77 (1874), 258–262.

14. ______, *Mitteilungen zur Lehre vom Transfiniten. I. & II.*, ***Zeitschrift für Philosophie und philosophische Kritik***, vol. 91/92 (1887/1888), 81–125 / 240–265.
15. ______, *Über eine elementare Frage der Mannigfaltigkeitslehre*, ***Jahresbericht der Deutschen Mathematiker-Vereinigung***, vol. 1 (1891), 75–78.
16. ______, *Beiträge zur Begründung der transfiniten Mengenlehre. I./II.*, ***Mathematische Annalen***, vol. 46/49 (1895/1897), 481–512/207–246 (see [17] for a translation into English).
17. ______, ***Contributions to the Founding of the Theory of Transfinite Numbers,*** (translation into English of [16]), [translated, and provided with an introduction and notes, by Philip E. B. Jourdain], Open Court Publishing Company, Chicago and London, 1915 [reprint: Dover Publications, New York, 1952].
18. ______, ***Gesammelte Abhandlungen mathematischen und philosophischen Inhalts***, *Mit Erläuternden Anmerkungen sowie mit Ergänzungen aus dem Briefwechsel Cantor-Dedekind*, edited by E. Zermelo, Julius Springer, Berlin, 1932.
19. RICHARD DEDEKIND, ***Was sind und was sollen die Zahlen***, Friedrich Vieweg & Sohn, Braunschweig, 1888 (see also [20, pp. 335–390]).
20. ______, ***Gesammelte mathematische Werke III***, edited by R. Fricke, E. Noether, and Ö. Ore, Friedrich Vieweg & Sohn, Braunschweig, 1932.
21. HEINZ-DIETER EBBINGHAUS (in cooperation with VOLKER PECKHAUS), ***Ernst Zermelo, An Approach to His Life and Work***, Springer-Verlag, Berlin · Heidelberg, 2007.
22. EUCLID, ***"Elements"***, Alexandria, *written around* 300 B.C.
23. LEONHARD EULER, ***Introductio in analysin infinitorum (Tomus Primus)***, Marcum-Michaelem Bousquet & Socios, Lausanne, 1748 (see [25]/[24] for a translation into English/German).
24. ______, ***Einleitung in die Analysis des Unendlichen***, [translated by H. Maser], Julius Springer, Berlin, 1885 [reprint with an introduction by W. Walter: Springer-Verlag, Berlin, 1983].
25. ______, ***Introduction to Analysis of the Infinite***, [translated by John D. Blanton], Springer-Verlag, Berlin · Heidelberg, 1988.
26. ANITA BURDMAN FEFERMAN, ***From Trotsky to Gödel: The Life of Jean Van Heijenoort***, A K Peters, Natick, Massachusetts, 1993.
27. ABRAHAM A. FRAENKEL, ***Abstract Set Theory***, [Studies in Logic and the Foundations of Mathematics], North-Holland, Amsterdam, 1961.
28. ABRAHAM A. FRAENKEL, YEHOSHUA BAR-HILLEL, AND AZRIEL LEVY, ***Foundations of Set Theory***, [Studies in Logic and the Foundations of Mathematics 67], North-Holland, Amsterdam, 1973.
29. ADOLF FRAENKEL, *Axiomatische Begründung der transfiniten Kardinalzahlen. I.*, ***Mathematische Zeitschrift***, vol. 13 (1922), 153–188.
30. ______, *Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre*, ***Mathematische Annalen***, vol. 86 (1922), 230–237.

31. ______, *Axiomatische Theorie der geordneten Mengen. (Untersuchungen über die Grundlagen der Mengenlehre. II.)*, ***Journal für die reine und angewandte Mathematik (Crelle)***, vol. 155 (1926), 129–158.
32. ______, *Zusatz zu vorstehendem Aufsatz Herrn v. Neumanns* (refers to [73]), ***Mathematische Annalen***, vol. 99 (1928), 392–393.
33. ______, *Das Leben Georg Cantors*, *in* ***Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*** (E. Zermelo, ed.), Julius Springer, Berlin, 1932, pp. 452–483.
34. GOTTLOB FREGE, ***Die Grundlagen der Arithmetik, eine logisch-mathematische Untersuchung über den Begriff der Zahl***, Verlag von Wilhelm Koebner, Breslau, 1884. [English translation (with German original): *The foundations of arithmetic. A logo-mathematical enquiry into the concept of number*, translated by J. L. Austin, Basil Blackwell, Oxford, 1950].
35. ______, *The concept of number,* (translation into English of parts of [34]), *in* ***Philosophy of Mathematics. Selected Readings*** (Paul Benacerraf and Hilary Putnam, eds.), Cambridge University Press, Cambridge, 2nd ed., 1983, pp. 130–159.
36. GALILEO GALILEI, ***Discorsi e Dimonstrazioni Matematiche, intorno à due nuove scienze attenenti alle Mecanica & i Moviemnti Locali***, Elzevirs, Leyden, 1638. [English translation: *Two New Sciences. Including Centers of Gravity and Force of Percussion*, translated by S. Drake, The University of Wisconsin Press, Madison, 1974; German translation: *Unterredungen und mathematische Demonstrationen über zwei neue Wissenszweige, die Mechanik und die Fallgesetze betreffend*, translated and edited by A. J. von Oettingen, Verlag Harri Deutsch, Frankfurt/M., 2004].
37. KURT GÖDEL, *The consistency of the axiom of choice and of the generalized continuum-hypothesis*, ***Proceedings of the National Academy of Sciences (U.S.A.)***, vol. 24 (1938), 556–557 (reprinted in [40]).
38. ______, *Consistency-proof for the generalized continuum-hypothesis*, ***Proceedings of the National Academy of Sciences (U.S.A.)***, vol. 25 (1939), 220–224 (reprinted in [40]).
39. ______, ***The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory***, [Annals of Mathematics Studies 3], Princeton University Press, Princeton (New Jersey), 1940 (reprinted in [40]).
40. ______, ***Collected Works, Volume II: Publications 1938–1974***, edited by S. Feferman (Editor-in-chief), J. W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, J. van Heijenoort, Oxford University Press, New York, 1990.
41. IVOR GRATTAN-GUINNESS, *How Bertrand Russell discovered his paradox*, ***Historia Mathematica***, vol. 5 (1978), 127–137.
42. MICHAEL HALLETT, ***Cantorian Set Theory and Limitation of Size***, Oxford Logic Guides 10, The Clarendon Press Oxford University Press, New York, 1984.
43. FRIEDRICH HARTOGS, *Über das Problem der Wohlordnung*, ***Mathematische Annalen***, vol. 76 (1915), 438–443.

44. FELIX HAUSDORFF, ***Grundzüge der Mengenlehre***, de Gruyter, Leipzig, 1914 [reprint: Chelsea, New York, 1965].
45. GERHARD HESSENBERG, ***Grundbegriffe der Mengenlehre***, Vandenhoeck & Ruprecht, Göttingen, 1906.
46. DAVID HILBERT, *Über das Unendliche*, ***Mathematische Annalen***, vol. 95 (1925), 161–190 (see [47] or [96] for a translation into English).
47. ______, *On the infinite*, in ***Philosophy of Mathematics. Selected Readings*** (Paul Benacerraf and Hilary Putnam, eds.), Cambridge University Press, Cambridge, 2nd ed., 1983, pp. 183–201.
48. ARIE HINKIS, ***Proofs of the Cantor-Bernstein theorem: a mathematical excursion***, [Science Networks. Historical Studies, vol. 45], Birkhäuser, Basel, 2013.
49. THOMAS JECH, *About the axiom of choice*, in ***Handbook of Mathematical Logic*** (J. Barwise, ed.), North-Holland, Amsterdam, 1977, pp. 345–370.
50. PHILIP E. B. JOURDAIN, *On infinite sums and products of cardinal numbers*, ***The Quarterly Journal of Pure and Applied Mathematics***, vol. 39 (1907), 375–384.
51. ______, *The multiplication of Alephs*, ***Mathematische Annalen***, vol. 65 (1908), 506–512.
52. AKIHIRO KANAMORI, ***The Higher Infinite, Large Cardinals in Set Theory from Their Beginnings***, [Perspectives in Mathematical Logic], Springer-Verlag, Heidelberg, 1994.
53. ______, *The mathematical development of set theory from Cantor to Cohen*, ***The Bulletin of Symbolic Logic***, vol. 2 (1996), 1–71.
54. ______, *The mathematical import of Zermelo's well-ordering theorem*, ***The Bulletin of Symbolic Logic***, vol. 3 (1997), 281–311.
55. ______, *The empty set, the singleton, and the ordered pair*, ***The Bulletin of Symbolic Logic***, vol. 9 (2003), 273–298.
56. ______, *Zermelo and set theory*, ***The Bulletin of Symbolic Logic***, vol. 10 (2004), 487–553.
57. ______, *Gödel and set theory*, ***The Bulletin of Symbolic Logic***, vol. 13 (2007), 153–188.
58. STEPHEN COLE KLEENE, ***Introduction to Metamathematics***, [Bibliotheca Mathematica, A Series of Monographs on Pure and Applied Mathematics, Volume I], North-Holland, Amsterdam, 1959.
59. JULIUS KÖNIG, *Zum Kontinuum-Problem*, ***Mathematische Annalen***, vol. 60 (1905), 177–180.
60. ALWIN KORSELT, *Über einen Beweis des Äquvialenzsatzes*, ***Mathematische Annalen***, vol. 70 (1911), 294–296.
61. CASIMIR KURATOWSKI, *Sur la notion de l'ordre dans la théorie des ensembles*, ***Fundamenta Mathematicae***, vol. 2 (1921), 161–171.
62. AZRIEL LÉVY, *The independence of various definitions of finiteness*, ***Fundamenta Mathematicae***, vol. 46 (1958), 1–13.
63. ADOLF LINDENBAUM AND ALFRED TARSKI, *Communication sur les recherches de la théorie des ensembles*, ***Comptes Rendus des Séances de la***

Société des Sciences et des Lettres de Varsovie, Classe III, vol. 19 (1926), 299–330.

64. LEOPOLD LÖWENHEIM, *Über Möglichkeiten im Relativkalkül*, ***Mathematische Annalen***, vol. 76 (1915), 447–470.
65. PAOLO MANCOSU, *Measuring the size of infinite collections of natural numbers: was Cantor's theory of infinite number inevitable?*, ***The Review of Symbolic Logic***, vol. 4 (2009), 612–646.
66. ELLIOTT MENDELSON, *The axiom of fundierung and the axiom of choice*, ***Archiv für mathematische Logik und Grundlagenforschung***, vol. 4 (1958), 67–70.
67. DIMITRY MIRIMANOFF, *Remarques sur la théorie des ensembles et les antinomies cantoriennes. I.*, ***L'Enseignement Mathématique***, vol. 19 (1917), 209–217.
68. ———, *Remarques sur la théorie des ensembles et les antinomies cantoriennes. II.*, ***L'Enseignement Mathématique***, vol. 21 (1920), 29–52.
69. GREGORY H. MOORE, ***Zermelo's Axiom of Choice. Its origins, development, and influence***, [Studies in the History of Mathematics and Physical Sciences 8], Springer-Verlag, New York, 1982.
70. JOHN VON NEUMANN, *Zur Einführung der transfiniten Zahlen*, ***Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco·Josephinae. Sectio Scientiarum Mathematicarum***, vol. 1 (1923), 199–208 (see [96] for a translation into English).
71. ———, *Eine Axiomatisierung der Mengenlehre*, ***Journal für die Reine und Angewandte Mathematik***, vol. 154 (1925), 219–240 (see [96] for a translation into English).
72. ———, *Die Axiomatisierung der Mengenlehre*, ***Mathematische Zeitschrift***, vol. 27 (1928), 669–752.
73. ———, *Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre*, ***Mathematische Annalen***, vol. 99 (1928), 373–391.
74. ———, *Über eine Widerspruchfreiheitsfrage in der axiomatischen Mengenlehre*, ***Journal für die Reine und Angewandte Mathematik***, vol. 160 (1929), 227–241.
75. GIUSEPPE PEANO, *Démonstration de l'intégrabilité des équations différentielles ordinaires*, ***Mathematische Annalen***, vol. 37 (1890), 182–228.
76. ———, *Super theorema de Cantor-Bernstein*, ***Rendiconti del Circolo Matematico di Palermo***, vol. 21 (1906), 360–366.
77. CHARLES S. PEIRCE, *On the algebra of logic*, ***American Journal of Mathematics***, vol. 7 (1885), 180–202.
78. OSKAR PERRON, ***Irrationalzahlen***, 2nd ed., Chelsea Publishing Company, New York, 1951.
79. PLATO, ***Politeia***, Athens, Book I *written around* 390 B.C., Books II–X *written around* 375 B.C.
80. ———, ***Parmenides***, Athens, *written around* 370 B.C.

81. BERNHARD RANG AND WOLFGANG THOMAS, *Zermelo's discovery of the "Russell paradox"*, ***Historia Mathematica***, vol. 8 (1981), 15–22.
82. BERTRAND RUSSELL, *On finite and infinite cardinal numbers,* (Section III of Whitehead [98]), ***American Journal of Mathematics***, vol. 24 (1902), 378–383.
83. ______, *Mathematical logic as based on the theory of types*, ***American Journal of Mathematics***, vol. 30 (1908), 222–262 (also published in [96]).
84. ______, ***Introduction to Mathematical Philosophy***, 2nd ed., Georg Allen & Unwin, London, 1920 (see [85] for a translation into German).
85. ______, ***Einführung in die mathematische Philosophie***, [translated into German by E. J. Gumbel and W. Gordon], Drei Masken Verlag, München, 1923.
86. ______, ***The Principles of Mathematics***, 2nd ed., Georg Allen & Unwin, London, 1937.
87. ARTHUR M. SCHOENFLIES, *Die Entwickelung der Lehre von den Punktmannigfaltigkeiten*, ***Jahresbericht der Deutschen Mathematiker-Vereinigung***, vol. 8_2 (1900), 1–250.
88. ERNST SCHRÖDER, *Über zwei Definitionen der Endlichkeit und G. Cantor'sche Sätze*, ***Nova Acta, Abhandlungen der Kaiserlich Leopoldinisch-Carolinisch Deutschen Akademie der Naturforscher***, vol. 71 (1898), 301–362.
89. STEWART SHAPIRO, *We hold these truths to be self-evident: but what do we mean by that?*, ***The Review of Symbolic Logic***, vol. 2 (2009), 175–207.
90. WACŁAW SIERPIŃSKI, ***Leçons sur les nombres transfinis***, Gauthier-Villars et Fils, Paris, 1928.
91. THORALF SKOLEM, *Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre*, Matematikerkongressen i Helsingfors den 4–7 Juli 1922, Den femte skandinaviska matematikerkongressen (Helsingfors), Akademiska Bokhandeln, 1923, pp. 217–232 (see [96] for a translation into English).
92. ______, ***Selected Works in Logic***, Universitetsforlaget, Oslo-Bergen-Tromsö, 1970.
93. LADISLAV SPIŠIAK AND PETER VOJTÁŠ, *Dependences between definitions of finiteness*, ***Czechoslovak Mathematical Journal***, vol. 38(113) (1988), 389–397.
94. FABIAN STEDMAN, ***Campanalogia: or the Art of Ringing Improved***, W. Godbid, London, 1677 [reprint: Christopher Groome 1990].
95. ALFRED TARSKI, *Sur les principes de l'arithmétique des nombres ordinaux (transfinis)*, ***Annales de la Société Polonaise de Mathématique***, vol. 3 (1925), 148–149.
96. JEAN VAN HEIJENOORT, ***From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931***, [Source Books in the History of Science], Harvard University Press, Cambridge, Massachusetts, 1967.
97. JAN VON PLATO, *In the shadows of the Löwenheim-Skolem theorem: Early combinatorial analyses of mathematical proofs*, ***The Bulletin of Symbolic Logic***, vol. 13 (2007), 189–225.

98. ALFRED NORTH WHITEHEAD, *On cardinal numbers*, ***American Journal of Mathematics***, vol. 24 (1902), 367–394.
99. ERNST ZERMELO, *Bericht an die Notgemeinschaft der Deutschen Wissenschaft über meine Forschungen betreffend die Grundlagen der Mathematik*, typescript, 5 pp., with appendices, 2 pp., dated 3 December 1930, Universitätsarchiv Freiburg, *Zermelo Nachlass, part of C 129/140* (see [104] for a translation into English).
100. ———, *Beweis, dass jede Menge wohlgeordnet werden kann*, ***Mathematische Annalen***, vol. 59 (1904), 514–516 (see [96, 104] for a translation into English).
101. ———, *Neuer Beweis für die Möglichkeit einer Wohlordnung*, ***Mathematische Annalen***, vol. 65 (1908), 107–128 (see [96, 104] for a translation into English).
102. ———, *Untersuchungen über die Grundlagen der Mengenlehre. I.*, ***Mathematische Annalen***, vol. 65 (1908), 261–281 (see [96, 104] for a translation into English).
103. ———, *Über Grenzzahlen und Mengenbereiche. Neue Untersuchungen über die Grundlagen der Mengelehre*, ***Fundamenta Mathematicae***, vol. 16 (1930), 29–47 (see [104] for a translation into English).
104. ———, ***Collected Works / Gesammelte Werke***, *Volume I: Set Theory, Miscellanea / Band I: Mengenlehre, Varia*, [Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften, Nr. 21 (2010)], edited by Heinz-Dieter Ebbinghaus, Craig G. Fraser, and Akihiro Kanamori, Springer-Verlag, Berlin · Heidelberg, 2010.