Achim Blumensath
blumens@fi.muni.cz

This document was last updated 2024-04-09.
The latest version can be found at

www.fi.muni.cz/~blumens

Contents

Part A.

Set Theory

A1. *Basic set theory*

1. *Sets and classes*

In mathematics there are basically two ways to define the objects under consideration. On the one hand, one can explicitly construct them from already known objects. For instance, the rational numbers and the real numbers are usually introduced in this way. On the other hand, one can take the axiomatic approach, that is, one compiles a list of desired properties and one investigates any object meeting these requirements. Some well known examples are groups, fields, vector spaces, and topological spaces.

Since set theory is meant as foundation of mathematics there are no more basic objects available in terms of which we could define sets. Therefore, we will follow the axiomatic approach. We will present a list of six axioms and any object satisfying all of them will be called a *model of set theory.* Such a model consists of two parts: (1) a collection $\mathbb{S}$ of objects that we will call *sets,* and (2) some method which, given two sets a and b, tells us whether a is *an element of* b.

We will not care what exactly the objects in $\mathbb{S}$ are or how this method looks like. For example, one could imagine a model of set theory consisting of natural numbers. If we define that a natural number a is *an element of* the natural number b if and only if the a-th bit in the binary encoding of b is 1, then all but one of our axioms will be satisfied. It is conceivable that a similar but more involved definition might yield a model that satisfies all of them.

We will introduce our axioms in a stepwise fashion during the following sections. To help readers trying to look up a certain axiom we include a complete list below even if most of the needed definitions are still missing.

Axiom of Extensionality. Two sets a and b are equal if, and only if, we have $x \in a \Leftrightarrow x \in b$, for all sets x.

Axiom of Separation. If a is a set and φ a property then $\{ x \in a \mid \varphi \}$ is a set.

Axiom of Creation. For every set a there is a set S such that S is a stage and $a \in S$.

Axiom of Infinity. There exists a set that is a limit stage.

Axiom of Choice. For every set A there exists a well-order R over A.

Axiom of Replacement. If F is a function and $\operatorname{dom} F$ is a set then so is $\operatorname{rng} F$.

Asking whether these axioms are *true* does make as much sense as the question of whether the field axioms are true, or those of a vector space. Instead, what we are concerned with is their *consistency* and *completeness*. That is, there should *exist* at least one object satisfying these axioms and all such objects should *look alike*. Unfortunately, one can prove that there is no complete axiom system for set theory. Hence, we will have to deal with the fact that there are many different models of set theory and there is no way to choose one of them as the 'canonical one'. In particular, there is no such thing as 'the real model of set theory'.

More seriously, it is even impossible to prove that our axiom system is consistent. That is, it might be the case that there is *no* model of set theory and we have wasted our time studying a nonsensical theory.

The first problem is dealt with rather easily. It does not matter which of these models we are given since any theorem that we can derive from the axioms holds in every model. But the second problem is serious. All we can do is to restrict ourselves to as few axioms as possible and to hope that no one will ever be able to derive a contradiction. Of course, the weaker the axioms the more different models we might get and the fewer theorems we will be able to prove.

In the following we will assume that $\mathbb{S}$ is an arbitrary but fixed model of set theory. That is, $\mathbb{S}$ is a collection of objects that satisfies all the axioms we will introduce below. $\mathbb{S}$ will be called the *universe* and its elements are called *sets*. Note that $\mathbb{S}$ itself is not a set since we will prove below that no set is an element of itself. By convention, if below we say that some set *exists* then we mean that it is contained in $\mathbb{S}$. Similarly, we say that *all* sets have some property if all elements of $\mathbb{S}$ do so.

Intuitively, a set is a collection of objects called its *elements*. If a and b are sets, i.e., elements of $\mathbb{S}$, we write $a \in b$ if a is *an element of* b and we define

$$a \subseteq b \quad :\text{iff} \quad \text{every element } x \in a \text{ is also an element } x \in b\,.$$

If $a \subseteq b$, we call a a *subset* of b, and we say that a is *included* in b, or that b is a *superset* of a. We use the usual abbreviations such as $a \subset b$ for $a \subseteq b$ and $a \neq b$; $a \ni b$ for $b \in a$; and $a \notin b$ if $a \in b$ does not hold.

Since a set is a collection of objects it is natural to require that a set is uniquely determined by its elements. Our first axiom can therefore be regarded as the definition of a set.

Axiom of Extensionality. *Two sets a and b are equal if, and only if,*

$$x \in a \quad \text{iff} \quad x \in b\,, \quad \textit{for all sets } x\,.$$

Lemma 1.1. *Two sets a and b are equal if and only if $a \subseteq b$ and $b \subseteq a$.*

In order to define a set we have to say what its elements are. If the set is finite we can just enumerate them. Otherwise, we have to find some property φ such that an object x is an element of a if, and only if, it has the property φ.

Definition 1.2. (a) Let φ be a property. $\{ x \mid \varphi \}$ denotes the set a such that, for all sets x, we have

$$x \in a \quad \text{iff} \quad x \text{ has property } \varphi\,.$$

If $\mathbb{S}$ does not contain such an object then the expression $\{\, x \mid \varphi \,\}$ is undefined.

(b) Let $b_0, \dots, b_{n-1}$ be sets. We define

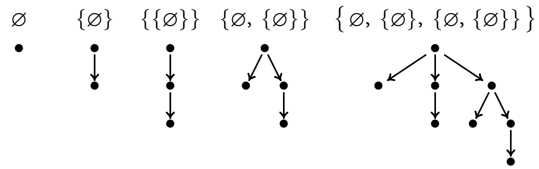
$$\{b_0, \dots, b_{n-1}\} := \{\, x \mid x = b_i \text{ for some } i < n \,\}\,.$$

(c) The *empty set* is $\varnothing := \{\, x \mid x \neq x \,\}$.

Note that, by the Axiom of Extensionality, if the set $\{\, x \mid \varphi \,\}$ exists, it is unique.

In a model of set theory nothing but sets exists. But how can we have sets without some objects that serve as elements? The answer of course is to construct sets of other sets. First of all, there is one set that we can form even if we do not have any suitable elements: the empty set $\varnothing$. So we already have one object and we use it as element of other sets. In the next step we can form the set $\{\varnothing\}$, then we can form the sets $\{\,\{\varnothing\}\,\}$ and $\{\,\varnothing, \{\varnothing\}\,\}$ and so on.

Sometimes it is helpful to imagine such sets as trees. The empty set $\varnothing$ corresponds to a single vertex •. To a nonempty sets a we associate the tree consisting of a root to which we attach, for every element $b \in a$ the tree corresponding to b. For example, we have



To better understand this inductive construction of sets we introduce a toy version of set theory which has the advantage that it can be defined explicitly. It consists of all sets that one can construct from the empty set in finitely many steps.

Definition 1.3. We construct a sequence $\mathrm{HF}_0 \subseteq \mathrm{HF}_1 \subseteq \dots$ of sets as follows. We start with the empty set $\mathrm{HF}_0 := \varnothing$. When the set HF_n has already been defined, the next stage

$$\mathrm{HF}_{n+1} := \big\{\, x \bigm| x \subseteq \mathrm{HF}_n \,\big\}$$

consists of all sets that we can construct from elements of HF_n.

A set is called *hereditary finite* if it is an element of some HF_n. The set of all hereditary finite sets is

$$\mathrm{HF} := \{\, x \mid x \in \mathrm{HF}_n \text{ for some } n \,\}\,.$$

Note that we cannot prove at the moment that HF really is a set. Since the empty universe $\mathbb{S} = \varnothing$ trivially satisfies the Axiom of Extensionality, we even cannot show that the empty set exists without additional axioms. Let us assume for the moment that HF does exists. Its first stages are

$$\begin{aligned}
\mathrm{HF}_0 &= \varnothing \\
\mathrm{HF}_1 &= \{\varnothing\} \\
\mathrm{HF}_2 &= \big\{\varnothing, \{\varnothing\}\big\} \\
\mathrm{HF}_3 &= \big\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\big\} \\
&\dots
\end{aligned}$$

By induction on n, one can prove that $\mathrm{HF}_n \subseteq \mathrm{HF}_{n+1}$ and each set $a \in \mathrm{HF}_{n+1}$ is of the form $a = \{b_0, \dots, b_{k-1}\}$, for finitely many elements $b_0, \dots, b_{k-1} \in \mathrm{HF}_n$. Note that each stage HF_n is hereditary finite since $\mathrm{HF}_n \in \mathrm{HF}_{n+1} \subseteq \mathrm{HF}$, but their union HF is not because $\mathrm{HF} \notin \mathrm{HF}$.

Exercise 1.1. Prove the following statements by induction on n. (Although we have not defined the natural numbers yet, you may assume for this exercise that they are available and that their usual properties hold.)

(a) $\mathrm{HF}_n \subseteq \mathrm{HF}_{n+1}$.

(b) HF_n has finitely many elements.

(c) Every set $a \in \mathrm{HF}_{n+1}$ is of the form $a = \{b_0, \dots, b_{k-1}\}$, for finitely many elements $b_0, \dots, b_{k-1} \in \mathrm{HF}_n$.

HF can be regarded as an approximation to the class of all sets. In fact, all but one of the usual axioms of set theory hold for HF. The only exception is the Axiom of Infinity which states that there exists an infinite set.

We can encode natural numbers by special hereditary finite sets.

Definition 1.4. To each natural number n we associate the set

$$[n] := \{[0], \dots, [n-1]\}.$$

The set of all natural numbers is

$$\mathbb{N} := \{ [n] \mid n \text{ a natural number} \}.$$

Note that $[n] \in \mathrm{HF}_{n+1}$ but $[n] \notin \mathrm{HF}_n$, and $\mathbb{N} \notin \mathrm{HF}$. This construction can be used to define the natural numbers in purely set theoretic terms. In the following by a *natural number* we will always mean a set of the form $[n]$.

It would be nice if there were a universe $\mathbb{S}$ that contains all sets of the form $\{ x \mid \varphi \}$. Unfortunately, such a universe does not exists, that is, if we add the axiom that claims that $\{ x \mid \varphi \}$ is defined for all φ, we obtain a theory that is inconsistent, i.e., it contradicts itself. In fact, we can even show that there are properties φ such that *no* model of set theory contains a set of the form $\{ x \mid \varphi \}$. And we can do so without using a single axiom of set theory.

Theorem 1.5 (Zermelo-Russell Paradox). $\{ x \mid x \notin x \}$ *is not a set.*

Proof. Suppose that the set $a := \{ x \mid x \notin x \}$ exists. Let x be an arbitrary set. By definition, we have $x \in a$ if and only if $x \notin x$. In particular, for $x = a$, we obtain $a \in a$ iff $a \notin a$. A contradiction. □

To better understand what is going on, let us see what happens if we restrict ourselves to hereditary finite sets. The set $\{ x \in \mathrm{HF} \mid x \notin x \}$ equals HF since no hereditary finite set contains itself. But $\mathrm{HF} \notin \mathrm{HF}$ is not hereditary finite. The same happens in real set theory. The condition

$x \notin x$ is satisfied by all sets and we have $\{ x \mid x \notin x \} = \mathbb{S}$, which is not a set.

In general, an expression of the form $\{ x \mid \varphi \}$ denotes a collection $X \subseteq \mathbb{S}$ that may or may not be a set, i.e., an element $X \in \mathbb{S}$. We will call objects of the form $\{ x \mid \varphi \}$ *classes*. Classes that are not sets will be called *proper classes*. If $X = \{ x \mid \varphi \}$ and $Y = \{ x \mid \psi \}$ are classes and a is a set, we write

$$\begin{aligned} a \in X \quad &:\text{iff} \quad a \text{ has property } \varphi, \\ X \subseteq Y \quad &:\text{iff} \quad \text{every set with property } \varphi \text{ also has property } \psi, \\ \text{and} \quad X = Y \quad &:\text{iff} \quad X \subseteq Y \text{ and } Y \subseteq X. \end{aligned}$$

If X is a proper class then we define $X \notin Y$, for every Y. Note that, if X and Y are sets then these definitions coincide with the ones above. Finally, we remark that every set a is a class since we can write a as $\{ x \mid x \in a \}$.

When defining classes we have to be a bit careful about what we call a property. Let us define a property to be a statement that is build up from basic propositions of the form $x \in y$ and $x = y$ by

- logical conjunctions like 'and', 'or', 'not', 'if-then';
- constructs of the form 'there exists a set x such that …' and 'for all sets x it holds that …'.

(Such statements will be defined in a more formal way in Chapter C1 where we will call them 'first-order formulae'.) Things we are not allowed to say include statements of the form 'There exists a property φ such that …' or 'For all classes X it holds that …'.

We have defined a class to be an object of the form $\{ x \mid \varphi \}$ where φ is a statement about sets. What happens if we allow statements about arbitrary classes? Note that, if φ is a property referring to a class $X = \{ x \mid \psi \}$ then we can transform φ into an equivalent statement only talking about sets by replacing all propositions $y \in X$, $X \in y$, $X = y$, etc. by their respective definitions.

Example. Let $X = \{ x \mid \varnothing \notin x \}$. We can write the class

$$\{ y \mid y \neq \varnothing \text{ and } y \subseteq X \}$$

in the form

$$\{ y \mid y \neq \varnothing \text{ and } \varnothing \notin x \text{ for all } x \in y \} .$$

The situation is analogous to the case of the complex numbers which are obtained from the real numbers by adding imaginary elements. We can translate any statement about complex numbers $x + iy$ into one about pairs $\langle x, y \rangle$ of real numbers. Consequently, it does not matter whether we allow classes in the definition of other classes.

Intuitively, the reason for a proper class such as $\mathbb{S}$ not being a set is that it is too 'large'. For instance, when considering HF we see that a set $a \subseteq \text{HF}$ is hereditary finite if, and only if, it has only finitely many elements. Hence, if we can show that a class $X = \{ x \mid \varphi \}$ is 'small', it should form a set. What do we mean by 'small'? Clearly, we would like every set to be small. Furthermore, it is natural to require that, if Y is small and $X \subseteq Y$ then X is also small. Therefore, we define a class X to be small if it is a subclass $X \subseteq a$ of some set a.

Definition 1.6. For a class A and a property φ we define

$$\{ x \in A \mid \varphi \} := \{ x \mid x \in A \text{ and } x \text{ has property } \varphi \} .$$

This definition ensures that every class of the form $X = \{ x \in a \mid \varphi \}$ where a is a set is small. Conversely, if $X = \{ x \mid \varphi \}$ is small then $X \subseteq a$, for some set a, and we have $X = \{ x \in a \mid \varphi \}$. Our second axiom states that every small class is a set.

Axiom of Separation. *If a is a set and φ a property then the class*

$$\{ x \in a \mid \varphi \}$$

is a set.

With this axiom we still cannot prove that there is any set. But if we have at least one set a, we can deduce, for instance, that also the empty set $\varnothing = \{ x \in a \mid x \neq x \}$ exists.

Definition 1.7. Let A and B be classes.

(a) The *intersection* of A is the class

$$\bigcap A := \{ x \mid x \in y \text{ for all } y \in A \} .$$

(b) The *intersection* of A and B is

$$A \cap B := \{ x \mid x \in A \text{ and } x \in B \} .$$

(c) The *difference* between A and B is

$$A \setminus B := \{ x \in A \mid x \notin B \} .$$

Lemma 1.8. *Let a be a set and B a class. Then $a \cap B$ and $a \setminus B$ are sets. If B contains at least one element then $\bigcap B$ is a set.*

Proof. The fact that $a \cap B = \{ x \in a \mid x \in B \}$ and $a \setminus B$ are sets follows immediately from the Axiom of Separation. If B contains at least one element $c \in B$ then we can write

$$\bigcap B = \{ x \in c \mid x \in y \text{ for all } y \in B \} . \qquad \square$$

Note that $\bigcap \varnothing = \mathbb{S}$ is not a set.

2. *Stages and histories*

The construction of HF above can be extended to one of the class $\mathbb{S}$ of all sets. We define $\mathbb{S}$ as the union of an increasing sequence of sets S_α, called the *stages* of $\mathbb{S}$. Again, we start with the empty set $S_0 := \varnothing$. If S_α is defined then the next stage $S_{\alpha+1}$ contains all subsets of S_α. But this time, we do not stop when we have defined S_α for all natural numbers α. Instead, every time we have defined an infinite sequence of stages we continue by taking their union to form the next stage. So our sequence starts with

$$S_0 = \mathrm{HF}_0\,, \quad S_1 = \mathrm{HF}_1\,, \quad S_2 = \mathrm{HF}_2\,, \quad \dots$$

The next stage after all the finite ones is $S_\omega := \mathrm{HF}$ and we continue with

$$S_{\omega+1} = \{\, x \mid x \subseteq \mathrm{HF} \,\}\,, \quad S_{\omega+2} = \{\, x \mid x \subseteq S_{\omega+1} \,\}\,, \quad \dots$$

After we have defined $S_{\omega+n}$ for all natural numbers n we again take the union

$$S_{\omega+\omega} = \{\, x \mid x \in S_{\omega+n} \text{ for some } n \,\}\,,$$

and so on.

Unfortunately, making this construction precise turns out to be quite technical since we cannot define the numbers α yet that we need to index the sequence S_α. This has to wait until Section A3.2. Instead, we start by giving a condition for some set S to be a *stage*, i.e., one of the S_α. If we order all such sets by inclusion then we obtain the desired sequence

$$S_0 \subseteq S_1 \subseteq \cdots \subseteq S_\omega \subseteq S_{\omega+1} \subseteq \cdots,$$

without the need to refer to its indices.

First, we isolate some characteristic properties of the sets HF_n which we would like that our stages S_α share. Note that, at the moment, we cannot prove that any of the sets mentioned below actually exists.

Definition 2.1. Let A be a class.

(a) We call A *transitive* if $x \in y \in A$ implies $x \in A$.

(b) We call A *hereditary* if $x \subseteq y \in A$ implies $x \in A$.

(c) The *accumulation* of A is the class

$$\mathrm{acc}(A) := \{\, x \mid \text{there is some } y \in A \text{ such that } x \in y \text{ or } x \subseteq y \,\}\,.$$

Note that each stage HF_n of HF is hereditary and transitive.

Exercise 2.1. By induction on n, show that the set $[n]$ is transitive. Give an example of a number n such that $[n]$ is not hereditary.

The next lemmas follow immediately from the definitions.

Lemma 2.2. *Let A be a class, and b, c sets. The following statements are equivalent:*

(a) *$c \in b \in A$ implies $c \in A$, that is, A is transitive.*

(b) *$b \in A$ implies $b \subseteq A$.*

(c) *$b \in A$ implies $b \cap A = b$.*

Lemma 2.3. *Let A and B be classes.*

(a) $A \subseteq \mathrm{acc}(A)$

(b) *If B is hereditary and transitive and if $A \subseteq B$, then* $\mathrm{acc}(A) \subseteq B$.

(c) *A is hereditary and transitive if, and only if,* $\mathrm{acc}(A) = A$.

Lemma 2.4. *If A and B are transitive classes then so is $A \cap B$.*

Exercise 2.2. Prove Lemmas 2.2, 2.3, and 2.4.

Definition 2.5. Let A be a class.

(a) A *minimal element* of A is an element $b \in A$ such that $b \cap A = \varnothing$, that is, there is no element $c \in A$ with $c \in b$.

(b) A set a is *founded* if every set $b \ni a$ has a minimal element.

(c) The *founded part* of A is the set

$$\mathrm{fnd}(A) := \{\, x \in A \mid x \text{ is founded} \,\}\,.$$

Example. The empty set $\varnothing$ and the set $\{\varnothing\}$ are founded. To see that $\{\varnothing\}$ is founded, consider a set $b \ni \{\varnothing\}$. If $\{\varnothing\}$ is not a minimal element of b, then $b \cap \{\varnothing\} \neq \varnothing$. Hence, $\varnothing \in b$ is a minimal element of b.

Exercise 2.3. Prove that every hereditary finite set is founded.

We will introduce an axiom below which implies that every class has a minimal element. Hence, every set is founded and we have $\mathrm{fnd}(A) = A$, for all classes A. Although the notions of a founded set and the founded part of a set will turn out to be trivial, we still need them to define stages and to formulate the axiom.

Lemma 2.6. *If B is a hereditary class and $a \in B$ then* $\mathrm{fnd}(a) \in \mathrm{fnd}(B)$.

Proof. For a contradiction suppose that $\mathrm{fnd}(a) \notin \mathrm{fnd}(B)$. Since B is hereditary and $\mathrm{fnd}(a) \subseteq a \in B$, we have $\mathrm{fnd}(a) \in B$. Consequently, $\mathrm{fnd}(a) \notin \mathrm{fnd}(B)$ implies that there is some set $x \ni \mathrm{fnd}(a)$ without minimal element. In particular, $\mathrm{fnd}(a)$ is not a minimal element of x, that is, there exists some set $y \in x \cap \mathrm{fnd}(a)$. But $y \in \mathrm{fnd}(a)$ implies that y is founded. Therefore, from $y \in x$ it follows that x has a minimal element. A contradiction. □

In the language of Section A3.1 the next theorem states that the membership relation $\in$ is well-founded on every class of transitive, hereditary sets.

Theorem 2.7. *Let A be a nonempty class. If every element $x \in A$ is hereditary and transitive, then A has a minimal element.*

Proof. Choose an arbitrary element $c \in A$ and set

$$b := \{ \mathrm{fnd}(x) \mid x \in c \cap A \} .$$

If $b = \varnothing$ then $c \cap A = \varnothing$ and c is a minimal element of A. Therefore, we may assume that $b \neq \varnothing$. Since $c \in A$ is hereditary, it follows from Lemma 2.6 that $b \subseteq \mathrm{fnd}(c)$. Fix some $x \in b \subseteq \mathrm{fnd}(c)$. Then x is founded and $x \in b$ implies that b has a minimal element y. By definition of b, we have $y = \mathrm{fnd}(z)$, for some $z \in c \cap A$.

We claim that z is a minimal element of A. Suppose otherwise. Then there exists some element $u \in z \cap A$. Since c is transitive we have $u \in c$. Hence, $u \in c \cap A$ implies $\mathrm{fnd}(u) \in b$. On the other hand, since $z \in A$ is hereditary it follows from Lemma 2.6 that $\mathrm{fnd}(u) \in \mathrm{fnd}(z)$. Hence,

$\mathrm{fnd}(u) \in \mathrm{fnd}(z) \cap b \neq \varnothing$ and $y = \mathrm{fnd}(z)$ is not a minimal element of b. A contradiction. □

We would like to define that a set S is a stage if it is hereditary and transitive. Unfortunately, this definition is too weak to show that the stages can be arranged in an increasing sequence $S_0 \subseteq S_1 \subseteq \dots \subseteq S_\alpha \subseteq \cdots$. Therefore, our definition will be slightly more involved. To each stage S_α we will associate its *history*

$$H(S_\alpha) = \{ S_\beta \mid \beta < \alpha \} ,$$

and we will call a set S a *stage* if $S = \mathrm{acc}(H(S))$. Note that, for HF_n, we have

$$H(\mathrm{HF}_n) = \{\mathrm{HF}_0, \dots, \mathrm{HF}_{n-1}\} \quad \text{and} \quad \mathrm{HF}_n = \mathrm{acc}\big(H(\mathrm{HF}_n)\big) .$$

Of course, to avoid a vicious cycle we have to define a history without mentioning stages.

Definition 2.8. (a) A class H is a *history* if every element $a \in H$ is hereditary, transitive, and satisfies

$$a = \mathrm{acc}(H \cap a) .$$

(b) If H is a history, we call the class $S := \mathrm{acc}(H)$ the *stage* with history H.

Let us show that these definitions have the desired effect.

Lemma 2.9. *Let S be a stage with history H.*

(a) $H \subseteq S$.

(b) *Every set $a \in H$ is a stage with history $H \cap a$.*

(c) *S is hereditary and transitive.*

(d) $S = \{ x \mid x \subseteq s$ *for some stage* $s \in S \}$.

(e) $H(S) := \{ s \in S \mid s$ *is a stage* $\}$ *is a history of* S.

Proof. (a) $a \subseteq a \in H$ implies $a \in \mathrm{acc}(H) = S$.

(b) By definition of a history, we have $a = \mathrm{acc}(H \cap a)$. Hence, if we can show that $H \cap a$ is a history then its stage is a. Clearly, every element of $H \cap a \subseteq H$ is hereditary and transitive. Let $b \in H \cap a$. Then $b \subseteq \mathrm{acc}(H \cap a) = a$. It follows that $H \cap b = (H \cap a) \cap b$. Furthermore, since H is a history we have

$$b = \mathrm{acc}(H \cap b) = \mathrm{acc}((H \cap a) \cap b)\,,$$

which shows that $H \cap a$ is a history.

(c) Let $b \in S$. The class

$$a := \{\, s \in H \mid b \in s \text{ or } b \subseteq s \,\}$$

is nonempty because $b \in S = \mathrm{acc}(H)$. By Theorem 2.7, it has a minimal element $s \in a$.

If $b \in s = \mathrm{acc}(H \cap s)$, there is some set $z \in H \cap s$ such that $b \in z$ or $b \subseteq z$. It follows that $z \in a$. But $z \in s \cap a$ implies that s is not a minimal element of a. Contradiction.

Therefore, $b \notin s$ which implies, by definition of a, that $b \subseteq s$. For transitivity, note that $x \in b$ implies

$$x \in b \subseteq s = \mathrm{acc}(H \cap s) \subseteq \mathrm{acc}(H) = S\,.$$

For hereditarity, let $x \subseteq b$. Then $x \subseteq b \subseteq s \in H$, which implies $x \in \mathrm{acc}(H) = S$.

(d) By (c) we know that $x \subseteq s \in S$ implies $x \in S$. For the other direction, suppose that $x \in S = \mathrm{acc}(H)$. There is some set $s \in H$ such that $x \in s$ or $x \subseteq s$. By (a), (b), and (c) it follows that $s \in S$, s is a stage, and s is hereditary and transitive. By transitivity, if $x \in s$ then $x \subseteq s$. Consequently, we have $x \subseteq s \in S$ in both cases and the claim follows.

(e) By (d), we have $S = \mathrm{acc}(H(S))$. It remains to show that $H(S)$ is a history. By (c), every element $s \in H(S)$ is hereditary and transitive. Furthermore, since S is transitive we have $s \subseteq S$ and it follows that

$$H(S) \cap s = \{\, x \in s \mid x \text{ is a stage } \}\,.$$

Since s is a stage we know by (d) that $s = \mathrm{acc}(H(S) \cap s)$. □

Note that, by (a) and (b) above, we have $H \subseteq H(S)$, for all histories H of S. In fact, $H(S)$ is the only history of S but we need some further results before we can prove this.

Exercise 2.4. Prove, by induction on n, that $\{\mathrm{HF}_0, \dots, \mathrm{HF}_{n-1}\}$ is a history with stage HF_n.

Exercise 2.5. Construct a hereditary transitivity set a that is not a stage. *Hint.* It is sufficient to consider sets $\mathrm{HF}_n \subset a \subset \mathrm{HF}_{n+1}$, for a small n.

After we have seen how to define stages we now prove that they form a strictly increasing sequence $S_0 \subseteq S_1 \subseteq \dots$. Together with Theorem 2.7 it follows that the class of all stages is well-ordered by the membership relation $\in$ (see Section A3.1).

Theorem 2.10. *If S and T are stages that are sets then we have*

$$S \in T \quad \textit{or} \quad S = T \quad \textit{or} \quad T \in S\,.$$

Proof. Suppose that there are stages S and T such that

$$(*) \quad S \notin T\,, \quad S \neq T\,, \quad \text{and} \quad T \notin S\,.$$

Define

$$A := \{\, s \mid s \text{ is a stage and there is some stage } t \text{ such that } s \text{ and } t \text{ satisfy } (*) \,\}\,.$$

By Theorem 2.7, the class A has a minimal element S_0. Define

$$B := \{\, t \mid t \text{ is a stage such that } S_0 \text{ and } t \text{ satisfy } (*) \,\}\,.$$

Again there is a minimal element $T_0 \in B$.

If we can show that $H(S_0) = H(T_0)$, it follows that

$$S_0 = \mathrm{acc}(H(S_0)) = \mathrm{acc}(H(T_0)) = T_0$$

in contradiction to our choice of S_0 and T_0.

Let $s \in S_0$ be a stage. Then $s \neq T_0$ since $T_0 \notin S_0$. Furthermore, we have $T_0 \notin s$ since, otherwise, transitivity of S_0 would imply that $T_0 \in S_0$. By minimality of S_0 it follows that s and T_0 do not satisfy $(*)$. Therefore, we have $s \in T_0$.

We have shown that $H(S_0) \subseteq H(T_0)$. A symmetric argument shows that $H(T_0) \subseteq H(S_0)$. Hence, we have $H(S_0) = H(T_0)$ as desired. □

Lemma 2.11. *Let S and T be stages that are sets.*

(a) $S \notin S$

(b) $S \subseteq T$ *if and only if* $S \in T$ *or* $S = T$.

(c) $S \subseteq T$ *or* $T \subseteq S$.

(d) $S \subset T$ *if, and only if,* $S \in T$.

Proof. (a) Suppose otherwise. Let X be the class of all stages s such that $s \in s$. By Theorem 2.7, X has a minimal element s, that is, an element such that $s \cap X = \varnothing$. But $s \in s \cap X$. Contradiction.

(b) If $S = T$ then $S \subseteq T$, and if $S \in T$ then $S \subseteq T$, by transitivity of T. Conversely, if neither $S = T$ nor $S \in T$ then Theorem 2.10 implies that $T \in S$. If $S \subseteq T$ then $T \in S \subseteq T$ would contradict (a).

(c) If $S \nsubseteq T$ then (b) implies that $S \notin T$ and $S \neq T$. By Theorem 2.10, it follows that $T \in S$ which, again by (b), implies $T \subseteq S$.

(d) We have $S \subset T$ iff $S \subseteq T$ and $S \neq T$. By (a) and (b), the latter is equivalent to $S \in T$. □

3. The cumulative hierarchy

In the previous section we have seen that we can arrange all stages in an increasing sequence

$$S_0 \subset S_1 \subset \cdots \subset S_\alpha \subset \cdots,$$

which we will call the *cumulative hierarchy*. If $S \in T$ are stages then we will say that S is *earlier* than T, or that T is *later* than S.

From the axioms we have available we cannot prove that there actually are any stages. We introduce a new axiom which ensures that enough stages are available.

Axiom of Creation. *For every set a there is a set $S \ni a$ which is a stage.*

In particular, this axiom implies that

- for every stage S that is a set, there exists a later stage $T \ni S$ that is also a set.
- the universe $\mathbb{S}$ is the union of all stages.

Of course, even with this new axiom we might still have $\mathbb{S} = \varnothing$. But if at least one set exists, we can now prove that $\mathrm{HF} \subseteq \mathbb{S}$. In particular, $\mathbb{S} = \mathrm{HF}$ satisfies all axioms we have introduced so far.

Exercise 3.1. Prove that $\mathbb{S}$ is a stage with history

$$H(\mathbb{S}) = \{ S \mid S \text{ is a stage} \}.$$

Definition 3.1. (a) We say that a stage T is the *successor* of the stage S if $S \in T$ and there exists no stage T' such that $S \in T' \in T$. A nonempty stage is a *limit* if it is not the successor of some other stage.

(b) Let A be a class. We denote by $S(A)$ the earliest stage such that $A \subseteq S(A)$.

Note that $S(A)$ is well-defined by Theorem 2.7. We have $S(s) = s$, for every stage s, in particular, $S(\varnothing) = \varnothing$. The stages $\mathbb{S}$ and HF are limits and HF_{n+1} is the successor of the stage HF_n.

Lemma 3.2. *$a \in b$ implies $S(a) \in S(b)$.*

Proof. Since $a \in b \subseteq S(b) = \mathrm{acc}(H(S(b)))$ it follows that there is some stage $s \in S(b)$ such that $a \in s$ or $a \subseteq s$. In particular, $S(a)$ is not later than s which implies that $S(a) \subseteq s \in S(b)$. As $S(b)$ is hereditary we therefore have $S(a) \in S(b)$. □

Lemma 3.3. *$\mathbb{S}$ is the only stage that is a proper class.*

Proof. Let S be a stage. If $S \neq \mathbb{S}$, there is some set $a \in \mathbb{S} \setminus S$. Hence, $S(a) \notin S$ which implies that

$$T \notin H(S)\,, \quad \text{for all stages } T \supseteq S(a)\,.$$

By Lemma 2.9 (e) and Theorem 2.10, we have

$$H(S) \subseteq \{\, T \mid T \text{ is a stage with } T \in S(a) \,\} = H(S(a))\,.$$

In particular, $H(S)$ is a set, which implies that so is $S = \mathrm{acc}(H(S))$. □

Lemma 3.4. *Let A be a class. The following statements are equivalent:*

(1) *A is a proper class.*

(2) *$S(A)$ is a proper class.*

(3) *$S(A) = \mathbb{S}$.*

Proof. (3) ⇒ (1) By the Axiom of Creation, if A is a set then so is $S(A)$.
(1) ⇒ (2) If $S(A)$ is a set then $A \subseteq S(A)$ implies that

$$A = \{\, x \in S(A) \mid x \in A \,\}$$

is also a set.
(2) ⇒ (3) follows by Lemma 3.3. □

With the Axiom of Creation we are finally able to prove most 'obvious' properties of sets such that no set is an element of itself or that the union of sets is a set.

Lemma 3.5. *If a is a set then $a \notin a$.*

Proof. Suppose that there exists some set such that $a \in a$. Then $a \in a \subseteq S(a)$ and, by Lemma 2.9 (d), there is some stage $s \in S(a)$ with $a \subseteq s$. This contradicts the minimality of $S(a)$. □

Theorem 3.6. *Every nonempty class A has a minimal element.*

Proof. By Theorem 2.7, we can choose some element $b \in A$ such that $S(b)$ is minimal. We claim that b is a minimal element of A. Suppose otherwise. Then there exists some element $x \in A \cap b$. Since $x \in b \subseteq S(b)$, Lemma 2.9 (d) implies that there is some stage $s \in S(b)$ such that $x \subseteq s$. Hence, x is an element of A with $S(x) \in S(b)$ in contradiction to the choice of b. □

We will see in Section A3.1 that Theorem 3.6 implies that there are no infinite descending sequences $a_0 \ni a_1 \ni \dots$ of sets. (If such a sequence exists then the set $\{a_0, a_1, \dots\}$ has no minimal element.)

Example. By induction on n, it trivially follows that, if $a_0 \ni \dots \ni a_{k-1}$ is a sequence of sets starting with $a_0 \in \mathrm{HF}_n$, then $k < n$. What happens if $a_0 = \mathrm{HF}$? Then $a_1 \in \mathrm{HF}_n$, for some n, and the sequence is of length $k \leq n$. But note that, for every n, we can find a sequence of length n starting with $a_0 = \mathrm{HF}$. So there is no one bound that works for all sequences.

Definition 3.7. Let A and B be classes.

(a) The *union* of A is the class

$$\bigcup A := \{\, x \mid x \in b \text{ for some } b \in A \,\}\,.$$

(b) The *union* of A and B is

$$A \cup B := \{\, x \mid x \in A \text{ or } x \in B \,\}\,.$$

(c) The *power set* of A is the class

$$\wp(A) := \{\, x \mid x \subseteq A \,\}\,.$$

Remark. Note that, by definition, a class contains only sets. In particular, the power set $\wp(A)$ of a proper class contains only the *subsets* of A, not all subclasses. For instance, we have $\wp(\mathbb{S}) = \mathbb{S}$.

Lemma 3.8. *If a and b are sets then so are $\bigcup a$, $a \cup b$, $\{a\}$, and $\wp(a)$.*

Proof. Let S_0 and S_1 be stages such that $a \in S_0$ and $b \in S_1$. We know that $S_0 \subseteq S_1$ or $S_1 \subseteq S_0$. By choosing either S_0 or S_1 we can find a stage S such that $S_0 \subseteq S$ and $S_1 \subseteq S$. By transitivity of S it follows that

$$\begin{aligned} \bigcup a &= \{\, x \in S \mid x \in b \text{ for some } b \in a \,\}\,, \\ a \cup b &= \{\, x \in S \mid x \in a \text{ or } x \in b \,\}\,, \\ \{a\} &= \{\, x \in S \mid x = a \,\}\,, \\ \text{and}\quad \wp(a) &= \{\, b \in S \mid b \subseteq a \,\}\,. \end{aligned}$$ □

Corollary 3.9. *If $a_0, \dots, a_{n-1}$ are sets then so is*

$$\{a_0, \dots, a_{n-1}\} = \{a_0\} \cup \dots \cup \{a_{n-1}\}\,.$$

In particular, every finite class is a set.

The next definition provides a useful tool which sometimes allows us to replace a proper class A by a set a. Instead of taking every element $x \in A$ we only consider those such that $S(x)$ is minimal.

Definition 3.10. The *cut* of a class A is the set

$$\operatorname{cut} A := \{\, x \in A \mid S(x) \subseteq S(y) \text{ for all } y \in A \,\}\,.$$

Exercise 3.2. What are $\operatorname{cut} \mathbb{S}$ and $\operatorname{cut} \{\, x \mid a \in x \,\}$?

Lemma 3.11. *Every class of the form* $\operatorname{cut} A$ *is a set.*

Proof. If $A = \varnothing$ then $\operatorname{cut} A = \varnothing$. Otherwise, choose an arbitrary set $a \in A$. Then $\operatorname{cut} A \subseteq S(a)$ which implies that $\operatorname{cut} A$ is a set. □

The following lemmas clarify the structure of the cumulative hierarchy.

Lemma 3.12. *The successor of a stage S is $\wp(S)$.*

Proof. By Theorem 2.7, there exists a minimal stage T with $S \in T$. We have to prove that $T = \wp(S)$. $a \subseteq S \in T$ implies $a \in T$ since T is hereditary. Hence, $\wp(S) \subseteq T$.

Conversely, if $s \in T$ is a stage then $S \notin s$ because T is the successor of S. By Theorem 2.10, it follows that $s \in S$ or $s = S$. This implies $s \subseteq S$.

We have shown that $s \in T$ iff $s \subseteq S$, for all stages s. It follows by Lemma 2.9 (d) that

$$\begin{aligned} T &= \{\, x \mid x \subseteq s \text{ for some stage } s \in T \,\} \\ &= \{\, x \mid x \subseteq s \text{ for some stage } s \subseteq S \,\} = \{\, x \mid x \subseteq S \,\} = \wp(S)\,. \end{aligned}$$ □

Lemma 3.13. *Let S be a nonempty stage. The following statements are equivalent:*

(1) *S is a limit stage.*

(2) *$S = \bigcup H(S)$.*

(3) *For every set $a \in S$, there exists some stage $s \in S$ with $a \in s$.*

(4) *If $a \in S$ then $\wp(a) \in S$.*

(5) *If $a \in S$ then $\{a\} \in S$.*

(6) *If $a \subseteq S$ then* $\operatorname{cut} a \in S$.

Proof. (2) ⇒ (1) Suppose that S is the successor of a stage T. Then we have

$$H(S) = \{T\} \cup H(T)\,.$$

Since $s \subseteq T$, for all $s \in H(T)$, it follows that

$$\bigcup H(S) = T \neq S.$$

(1) ⇒ (2) Suppose that S is a limit stage. By Lemma 2.9 (d), we have

$$\begin{aligned} S &= \bigcup \{\, \wp(s) \mid s \in H(S) \,\} \\ &= \bigcup \{\, t \mid t \text{ is the successor of some stage } s \in H(S) \,\} \\ &= \bigcup \{\, t \mid t \in H(S) \,\} \\ &= \bigcup H(S)\,. \end{aligned}$$

(1) ⇒ (3) Suppose that S is a limit and let $a \in S$. By Lemma 2.9 (d), there is some stage $s \in S$ with $a \subseteq s$. Hence, $a \in \wp(s)$. Since $T := \wp(s)$ is the successor of s we have $T \in S$.

(3) ⇒ (4) For each $a \in S$, there is some stage $s \in S$ with $a \in s$. Since s is transitive it follows that $x \subseteq a$ implies $x \in s$. Hence, $\wp(a) \subseteq s$. By transitivity of S, we obtain $\wp(a) \in S$.

(4) ⇒ (5) If $a \in S$ then $\{a\} \subseteq \wp(a) \in S$. Since S is hereditary, it follows that $\{a\} \in S$.

(5) ⇒ (1) If S is no limit, there is some stage $T \in S$ such that $S = \wp(T)$. By assumption, $\{T\} \in S = \wp(T)$. Hence, $\{T\} \subseteq T$ which implies that $T \in T$. A contradiction.

(3) ⇒ (6) Let $b := \operatorname{cut} a$. If $a = \varnothing$ then $b = \varnothing$ and we are done. If there is some element $x \in a$ then, by assumption, we can find a stage $s \in S$ with $x \in s$. By definition, $b \subseteq s$, and it follows that $b \in S$.

(6) ⇒ (5) Let $a \in S$ and set $b := \{ x \in S \mid a \subseteq x \}$. Clearly, $b \subseteq S$. By assumption, we therefore have $c := \operatorname{cut} b \in S$. Hence, $\{a\} \subseteq c$ implies $\{a\} \in S$. □

So far, we still might have $\mathbb{S} = \varnothing$ or $\mathbb{S} = \text{HF}$. To exclude these cases we introduce a new axiom which states that $\text{HF} \in \mathbb{S}$.

Axiom of Infinity. *There exists a set that is a limit stage.*

We call the theory consisting of the four axioms

- Axiom of Extensionality
- Axiom of Creation
- Axiom of Separation
- Axiom of Infinity

basic set theory. Every model of this theory consist of a hierarchy of stages

$$S_0 \subset S_1 \subset \cdots \subset S_\omega \subset S_{\omega+1} \subset \dots$$

where $S_n = \text{HF}_n$, for finite n. The differences between two such models can be classified according to two axes: the length of the hierarchy and the size of each stage.

Let $\mathbb{S}$ and $\mathbb{S}'$ be two models with stages $(S_\alpha)_{\alpha<\kappa}$ and $(S'_\alpha)_{\alpha<\lambda}$, respectively. We know that their lengths κ and λ are at least what we will call $\omega + \omega$ in Section A3.2. But our current axioms do not tell us whether the process of creation stops there or whether we again take the union of all stages and continue taking power sets until we reach $\omega + \omega + \omega$. At this point we again have to decide whether to stop or to continue, and so on.

The second possible difference stems from the fact that the power-set operation is ambiguous. We know that $S_n = \text{HF}_n = S'_n$, for all finite n. But we might have $S_\alpha \neq S'_\alpha$, for infinite α. The reason is that there is no way to express that *all* subsets of S_α are contained in $S_{\alpha+1}$. We have the Axiom of Separation which states that all subsets exist that we can explicitly define. But there are much more possible subsets than there are definitions.

A2. Relations

1. Relations and functions

With basic set theory available we can define most of the concepts used in mathematics. The simplest one is the notion of an ordered pair. The characteristic property of such pairs is that $\langle a, b\rangle = \langle c, d\rangle$ implies $a = c$ and $b = d$.

Definition 1.1. (a) Let a and b be sets. The *ordered pair* $\langle a, b\rangle$ is the set

$$\langle a, b\rangle := \big\{\{a\}, \{a, b\}\big\} .$$

(b) Let A and B be classes. The *cartesian product* of A and B is the class

$$A \times B := \{ c \mid c = \langle a, b\rangle \text{ for some } a \in A \text{ and } b \in B \} .$$

Let us show that ordered pairs have the desired property.

Lemma 1.2. *If $\{a, b\} = \{a, c\}$ then $b = c$.*

Proof. We have $b \in \{a, b\} = \{a, c\}$. Hence, $b = a$ or $b = c$. In the latter case we are done. Otherwise, we have $c \in \{a, c\} = \{a, b\} = \{b\}$ which implies that $c = b$. □

Lemma 1.3. *If $\langle a, b\rangle = \langle c, d\rangle$ then $a = c$ and $b = d$.*

Proof. Suppose that $\langle a, b\rangle = \langle c, d\rangle$.

$$\{a\} \in \big\{\{a\}, \{a, b\}\big\} = \big\{\{c\}, \{c, d\}\big\}$$

implies $\{a\} = \{c\}$ or $\{a\} = \{c, d\}$. In the latter case, we have $a = c = d$. In both cases, we therefore have $\{a\} = \{c\}$. By the preceding lemma, it follows that $\{a, b\} = \{c, d\}$ and, applying the lemma again, we obtain $b = d$. □

Remark. The above definition of an ordered pair $\langle a, b\rangle$ does only work for sets. Nevertheless we will use also pairs $\langle A, B\rangle$ where A or B are proper classes. There are several ways to make such an expression well-defined. A simple one is to define

$$\langle A, B\rangle := (\{[0]\} \times A) \cup (\{[1]\} \times B) \quad (= A \uplus B)$$

whenever at least one of A and B is a proper class. (The operation $\uplus$ will be defined more generally in the next section.) It is easy to check that with this definition the term $\langle A, B\rangle$ has the properties of an ordered pair, that is, $A \uplus B = C \uplus D$ implies $A = C$ and $B = D$.

Definition 1.4. (a) For sets $a_0, \dots, a_n$ we define inductively

$$\langle\rangle := \varnothing\,, \qquad \langle a_0\rangle := a_0\,,$$

and $\quad \langle a_0, \dots, a_n\rangle := \langle\langle a_0, \dots, a_{n-1}\rangle, a_n\rangle\,.$

We call $\langle a_0, \dots, a_{n-1}\rangle$ a *tuple of length n.* $\langle\rangle$ is the *empty tuple.*

(b) For a class A, we define its *n-th power* by

$$A^0 := \{\langle\rangle\}\,, \quad A^1 := A\,, \quad \text{and} \quad A^{n+1} := A^n \times A\,, \quad \text{for } n > 1\,.$$

Definition 1.5. A *relation,* or a *predicate,* of *arity n* is a subclass $R \subseteq \mathbb{S}^n$. If $R \subseteq A^n$, for some class A, we say that R is *over A*.

Note that $\varnothing$ and $\{\langle\rangle\}$ are the only relations of arity 0. In logic they are usually interpreted as *false* and *true.* A relation of arity 1 over A is just a subclass $R \subseteq A$.

Definition 1.6. Let R be a binary relation. The *domain* of R is the class

$$\operatorname{dom} R := \{\, a \mid \langle a, b\rangle \in R \text{ for some } b \,\}\,,$$

and its *range* is

$$\operatorname{rng} R := \{\, b \mid \langle a, b\rangle \in R \text{ for some } a \,\}\,.$$

The *field* of R is $\operatorname{dom} R \cup \operatorname{rng} R$.

In particular, $\operatorname{dom} R$ and $\operatorname{rng} R$ are the smallest classes such that

$$R \subseteq \operatorname{dom} R \times \operatorname{rng} R\,.$$

Definition 1.7. (a) A binary relation R is called *functional* if, for every $a \in \operatorname{dom} R$, there exists exactly one set b such that $\langle a, b\rangle \in R$. We denote this unique element b by $R(a)$. Hence, we can write R as

$$R = \{\, \langle a, R(a)\rangle \mid a \in \operatorname{dom} R \,\}\,.$$

A functional relation $R \subseteq A \times B$ is also called a *partial function* from A to B.

(b) A *function* from A to B is a functional relation $f \subseteq A \times B$ with $\operatorname{dom} f = A$ and $\operatorname{rng} f \subseteq B$. Functions are also called *maps* or *mappings.* We write $f : A \to B$ to denote the fact that f is a function from A to B.

A function of *arity n* is a function of the form

$$f : A_0 \times \dots \times A_{n-1} \to B\,.$$

We will write $x \mapsto y$ to denote the function f such that $f(x) = y$. (Usually, y will be an expression depending on x.)

(c) For a set a and a class B, we denote by B^a the class of all functions $f : a \to B$.

Remark. A 0-ary function $f : A^0 \to B$ is uniquely determined by the value $f(\langle\rangle)$. We will call such functions *constants* and identify them with their only value.

Sometimes we write binary relations and functions in infix notation, that is, for a relation $R \in A \times A$, we write $a\ R\ b$ instead of $\langle a, b\rangle \in R$ and, for $f : A \times A \to A$, we write $a\ f\ b$ instead of $f(a, b)$.

Definition 1.8. (a) For every class A, we define the *identity function* $\mathrm{id}_A : A \to A$ by $\mathrm{id}_A(a) := a$.

(b) If $R \subseteq A \times B$ and $S \subseteq B \times C$ are relations, we can define their *composition* $S \circ R : A \times C$ by

$$S \circ R := \{ \langle a, c \rangle \mid \text{there is some } b \in B \text{ such that } \langle a, b \rangle \in R \text{ and } \langle b, c \rangle \in S \} .$$

(Note the reversal of the ordering.) In particular, if $f : A \to B$ and $g : B \to C$ are functions then

$$(g \circ f)(x) := g(f(x)) .$$

(c) The *inverse* of a relation $R \subseteq A \times B$ is the relation

$$R^{-1} := \{ \langle b, a \rangle \mid \langle a, b \rangle \in R \} .$$

In particular, a function $g : B \to A$ is the inverse of the function $f : A \to B$ if

$$g(f(a)) = a \quad \text{and} \quad f(g(b)) = b , \quad \text{for all } a \in A \text{ and } b \in B ,$$

that is, if $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

For $b \in B$, we will write

$$R^{-1}(b) := \{ a \mid \langle a, b \rangle \in R \} .$$

Note that, if R^{-1} is a function, we have already defined

$$R^{-1}(b) := a \quad \text{for the unique } a \text{ such that } \langle a, b \rangle \in R .$$

It should always be clear from the context which of these two definitions we have in mind when we write $R^{-1}(b)$.

(d) The *restriction* of a relation $R \subseteq A \times B$ to a class C is the relation

$$R|_C := R \cap (C \times C) .$$

Its *left restriction* is

$$R \upharpoonright C := R \cap (C \times B) .$$

(e) The *image* of a class C under a binary relation $R \subseteq A \times B$ is the class

$$R[C] := \mathrm{rng}\,(R \upharpoonright C) .$$

Remark. The set A^A together with the operation $\circ$ forms a *monoid*, that is, $\circ$ is *associative*

$$f \circ (g \circ h) = (f \circ g) \circ h , \qquad \text{for all } f, g, h \in A^A ,$$

and there exists a *neutral element*

$$\mathrm{id}_A \circ f = f \quad \text{and} \quad f \circ \mathrm{id}_A = f \qquad \text{for all } f \in A^A .$$

Exercise 1.1. Is it true that $R^{-1} \circ R = \mathrm{id}_A$, for all relations $R \subseteq A \times B$?

Exercise 1.2. Prove that $\circ$ is associative and that id_A is a neutral element.

Definition 1.9. Let $f : A \to B$ be a function.

(a) f is *injective* if there is no pair $a, a' \in A$ of distinct elements such that $f(a) = f(a')$.

(b) f is *surjective* if $\mathrm{rng}\, f = B$.

(c) f is called *bijective* if it is both injective and surjective.

Lemma 1.10. *Let $f : A \to B$ be a function.*

(a) *The following statements are equivalent:*

  (1) *f is bijective.*

  (2) *f^{-1} is a function $B \to A$.*

  (3) *There exists a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.*

(b) *The following statements are equivalent:*

(1) f *is injective.*

(2) $f \circ g = f \circ h$ *implies* $g = h$, *for all functions* $g, h : C \to A$.

(3) $A = \varnothing$ *or there exists some function* $g : B \to A$ *such that* $g \circ f = \mathrm{id}_A$.

(4) $f^{-1}[f[X]] = X$, *for all* $X \subseteq A$.

(c) *The following statements are equivalent:*

(1) f *is surjective.*

(2) $g \circ f = h \circ f$ *implies* $g = h$, *for all functions* $g, h : B \to C$.

(3) $f[f^{-1}[Y]] = Y$, *for all* $Y \subseteq B$.

(d) *If there exists some function* $g : B \to A$ *such that* $f \circ g = \mathrm{id}_B$ *then* f *is surjective.*

Proof. (a) (1) ⇒ (2) Let $b \in B$. Since f is surjective there exists some $a \in A$ such that $f(a) = b$. If $a' \in A$ is some element with $f(a') = b$ then the injectivity of f implies that $a' = a$. We have shown that, for every element $b \in B$, there is a unique $a \in A$ such that $f^{-1}(b) = a$. Hence, f^{-1} is functional and $\mathrm{dom}\, f^{-1} = B$.

(2) ⇒ (3) $f^{-1} : B \to A$ is a function and we have $f^{-1} \circ f = \mathrm{id}_A$ and $f \circ f^{-1} = \mathrm{id}_B$.

(3) ⇒ (1) If $f(a) = f(b)$, for $a, b \in A$, then

$$a = \mathrm{id}_A(a) = (g \circ f)(a) = (g \circ f)(b) = \mathrm{id}_A(b) = b\,.$$

Consequently, f is injective. To show that it is also surjective let $b \in B$. Setting $a := g(b)$ we have

$$f(a) = (f \circ g)(b) = \mathrm{id}_B(b) = b\,.$$

Hence, $b \in \mathrm{rng}\, f$.

(b) (1) ⇒ (4) Let $X \subseteq A$. For every $a \in X$, we have $f(a) \in f[X]$ and, therefore, $a \in f^{-1}[f[X]]$. Consequently, $X \subseteq f^{-1}[f[X]]$. Conversely, suppose that $a \in f^{-1}[f[X]]$ and set $b := f(a)$. Since $b \in f[X]$ there is some $c \in X$ with $f(c) = b$. As f is injective this implies that $a = c \in X$.

(4) ⇒ (3) If $A = \varnothing$ then there is nothing to do. Hence, assume that $A \neq \varnothing$. We define g as follows. For every $b \in \mathrm{rng}\, f$, there is some element $a \in A$ with $f(a) = b$. Since $f^{-1}(b) = f^{-1}[f[\{a\}]] = \{a\}$ it follows that this element a is unique. Hence, fixing $a_0 \in A$ we can define g by

$$g(b) := \begin{cases} a & \text{if } f^{-1}(b) = \{a\}\,, \\ a_0 & \text{if } b \notin \mathrm{rng}\, f\,. \end{cases}$$

(3) ⇒ (2) If $A = \varnothing$, there are no functions $C \to A$ and the claim holds trivially. Hence, assume that $A \neq \varnothing$ and let k be a function such that $k \circ f = \mathrm{id}_A$. Then $f \circ g = f \circ h$ implies

$$g = \mathrm{id}_A \circ g = k \circ f \circ g = k \circ f \circ h = \mathrm{id}_A \circ h = h\,.$$

(2) ⇒ (1) Suppose that f is not injective. Then there are two elements $a, b \in A$ with $a \neq b$ such that $f(a) = f(b)$. Let $C := [1] = \{0\}$ be a set with a single element and define $g, h : C \to A$ by $g(0) := a$ and $h(0) := b$. Then $g \neq h$ but $f \circ g = f \circ h$.

(c) (1) ⇒ (2) Suppose that $g \neq h$. There is some element $b \in B$ with $g(b) \neq h(b)$. Since f is surjective we can find an element $a \in A$ with $f(a) = b$. Hence, $(g \circ f)(a) = g(b) \neq h(b) = (h \circ f)(a)$.

(2) ⇒ (1) Suppose that f is not surjective. Then there is some element $b \in B \setminus \mathrm{rng}\, f$. Let $C := [2] = \{0, 1\}$ be a set with two elements and define $g, h : B \to C$ by

$$g(x) := \begin{cases} 1 & \text{if } x = b\,, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad h(x) := 0\,, \quad \text{for all } x \in B\,.$$

Then we have $g \neq h$ but $g \circ f = h \circ f$.

(3) ⇒ (1) $f[f^{-1}[B]] = B$ implies that $\mathrm{rng}\, f = B$.

(1) ⇒ (3) Let $Y \subseteq B$. If $b \in f[f^{-1}[Y]]$ then there is some $a \in f^{-1}[Y]$ with $f(a) = b$. Hence, $a \in f^{-1}[Y]$ implies that $b = f(a) \in Y$. Consequently, we have $f[f^{-1}[Y]] \subseteq Y$.

For the converse, let $b \in Y$. Since f is surjective there is some $a \in A$ with $f(a) = b$. Hence, $a \in f^{-1}[Y]$ and it follows that $b = f(a) \in f[f^{-1}[Y]]$.

(d) Let k be a function such that $f \circ k = \mathrm{id}_B$. Then $g \circ f = h \circ f$ implies

$$g = g \circ \mathrm{id}_B = g \circ f \circ k = h \circ f \circ k = h \circ \mathrm{id}_B = h\,.$$

By (c), it follows that f is surjective. □

Remark. The converse of (d) also holds but we cannot prove it without the Axiom of Choice, which we will introduce in Section A4.1 below. Actually one can prove that the Axiom of Choice is equivalent to the claim that, for every surjective function f, there exists some function g with $f \circ g = \mathrm{id}$.

Remark. The subset of all bijective functions $f : A \to A$ forms a *group* since, by the preceding lemma, every element f has an *inverse* f^{-1}.

Exercise 1.3. Let $f : A \to B$ and $g : B \to C$ be functions. Prove that, if f and g are (a) injective, (b) surjective, or (c) bijective then so is $g \circ f$.

We conclude this section with two important results about the existence of functions. The first one can be used to prove that there exists a bijection between two given sets without constructing this function explicitly.

Lemma 1.11. *Let $A \subseteq B \subseteq C$ be sets. If there exists a bijective function $f : C \to A$, there is also a bijection $g : C \to B$.*

Proof. Let

$$Z := \bigcap \{\, X \subseteq C \mid C \setminus B \subseteq X \text{ and } f[X] \subseteq X \,\}\,.$$

Then $C \setminus B \subseteq Z$ and $f[Z] \subseteq Z$. We claim that

$$g(x) := \begin{cases} f(x) & \text{if } x \in Z\,, \\ x & \text{otherwise,} \end{cases}$$

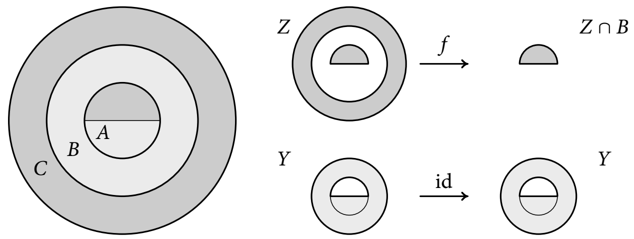


Figure 1.. The proof of Lemma 1.11.

is the desired bijection $g : C \to B$.

Let $Y := C \setminus Z$ be the complement of Z. Since $g[Y] \subseteq Y$ and $g[Z] \subseteq Z$ it is sufficient to show that the restrictions $g \restriction Y : Y \to Y$ and $g \restriction Z : Z \to Z \cap B$ are bijections. Clearly, $g \restriction Y = \mathrm{id}_Y$ is bijective and $g \restriction Z = f \restriction Z$ is injective. Therefore, we only need to prove that $f[Z] = Z \cap B$.

By definition of Z, we have $f[Z] \subseteq Z \cap \operatorname{rng} f \subseteq Z \cap B$. For the other inclusion, suppose that there exists some element $a \in (Z \cap B) \setminus f[Z]$. Since $a \in B$ the set $X := Z \setminus \{a\}$ satisfies $C \setminus B \subseteq X$ and $f[X] \subseteq X$. By definition of Z, it follows that $Z \subseteq X$. Contradiction. □

Theorem 1.12 (Bernstein). *If there are injective functions $f : A \to B$ and $g : B \to A$ then there exists a bijective function $h : A \to B$.*

Proof. We have $g[f[A]] \subseteq g[B] \subseteq A$. Since f and g are injective so is their composition $g \circ f$. When regarded as function $g \circ f : A \to g[f[A]]$ it is also surjective. Hence, by the preceding lemma, there exists a bijective mapping $h : A \to g[B]$. Since $k := g^{-1} \restriction g[B] : g[B] \to B$ is bijective it follows that so is $k \circ h : A \to B$. □

The second result deals with functions between a set and its power set.

Theorem 1.13 (Cantor). *For every set a, there exists an injective function $a \to \wp(a)$ but no surjective one.*

Proof. The function $f : a \to \wp(a)$ with $f(x) := \{x\}$ is injective.

For a contradiction, suppose that there is also a surjective function $f : a \to \wp(a)$. We define the set

$$z := \{ x \in a \mid x \notin f(x) \} \subseteq a .$$

Since f is surjective there is some element $b \in a$ with $f(b) = z$. By definition of z, we have

$$b \in z \quad \text{iff} \quad b \notin f(b) = z .$$

A contradiction. □

Corollary 1.14. *For all sets a, there are no injective functions $\wp(a) \to a$.*

Proof. Suppose that $f : \wp(a) \to a$ is injective. We define a function $g : a \to \wp(a)$ by

$$g(x) := \begin{cases} f^{-1}(x) & \text{if } x \in \operatorname{rng} f , \\ \varnothing & \text{otherwise} . \end{cases}$$

Note that g is well-defined since f is injective. Furthermore, we have $g \circ f = \mathrm{id}_{\wp(A)}$. Hence, Lemma 1.10 (d) implies that g is surjective. This contradicts the Theorem of Cantor. □

2. Products and unions

So far, we have defined cartesian products of finitely many sets and tuples of finite length. In this section we will show how to generalise these definitions to infinitely many factors.

Remark. (a) There is a canonical bijection $\pi : A^{[n]} \to A^n$ between the set $A^{[n]}$ of all functions $[n] \to A$ and the n-th power A^n of A. π maps a function $f : [n] \to A$ to the tuple

$$\pi(f) := \langle f(0), \dots, f(n-1) \rangle ,$$

and its inverse π^{-1} maps a tuple $\langle a_0, \dots, a_{n-1} \rangle$ to the function $f : [n] \to A$ with $f(i) = a_i$.

(b) There is also a canonical bijection $\pi : (A \times B) \times C \to A \times (B \times C)$ defined by

$$\pi(\langle a, b \rangle, c \rangle := \langle a, \langle b, c \rangle \rangle .$$

(c) Finally, let us define a canonical bijection $\pi : A^{B \times C} \to (A^C)^B$ that maps a function $f : B \times C \to A$ to the function $g : B \to A^C$ with

$$g(b) := h_b \quad \text{where} \quad h_b(c) := f(b, c) , \qquad \text{for } b \in B,\ c \in C .$$

In the theory of programming languages this transformation of a function $B \times C \to A$ into a function $B \to A^C$ is called *currying*.

Part (a) of the above remark gives a hint on how to generalise finite tuples. A tuple of length n corresponds to a map $[n] \to A$. Therefore, we define an infinite tuple as map $\mathbb{N} \to A$.

Definition 2.1. (a) Let A be a class and I a set. A function $f : I \to A$ is called a *sequence*, or *family*, over I. If $f(i) = a_i$ then we also write f in the form $(a_i)_{i \in I}$.

(b) Let I be a set, $(A_i)_{i \in I}$ a sequence of sets, and $A := \bigcup \{ A_i \mid i \in I \}$ their union. The *product* of $(A_i)_{i \in I}$ is the class

$$\prod_{i \in I} A_i := \{ f \in A^I \mid f(i) \in A_i \text{ for all } i \} .$$

(c) Let $(A_i)_{i \in I}$ be a sequence of sets and $k \in I$. The *projection* to the k-th coordinate is the map

$$\mathrm{pr}_k : \prod_{i \in I} A_i \to A_k \quad \text{with} \quad \mathrm{pr}_k(f) := f(k) .$$

Remark. (a) If $A_i = A$, for all $i \in I$, then $\prod_{i \in I} A_i = A^I$.

(b) As we have seen above there is a canonical bijection between $A_0 \times A_1$ and $\prod_{i \in [2]} A_i$. In the following we will not distinguish between these sets.

Let us introduce some notation and conventions regarding sequences. To indicate that a certain variable refers to a sequence we will write it with a bar $\bar{a}$. If the sequence is over I, the components of $\bar{a}$ will always be $(a_i)_{i\in I}$. Sometimes we will not distinguish between a sequence $\bar{a} = (a_i)_{i\in I}$ and its range $\operatorname{rng}\bar{a} = \{\, a_i \mid i \in I \,\}$. In particular, we write $\bar{a} \cup \bar{b}$ instead of $\operatorname{rng}\bar{a} \cup \operatorname{rng}\bar{b}$ and, if we do not want to specify the index set I, we will write $\bar{a} \subseteq A$ instead of $\bar{a} \in A^I$. Finally, for a function $f : A \to B$, we write $f(\bar{a})$ to denote the sequence $(f(a_i))_{i\in I}$.

Lemma 2.2. *Let A be a set and $(B_i)_{i\in I}$ a sequence of sets. For every sequence $(f_i)_{i\in I}$ of functions $f_i : A \to B_i$ there exists a unique function $g : A \to \prod_i B_i$ such that*

$$\mathrm{pr}_i \circ g = f_i\,, \qquad \textit{for all } i \in I\,.$$

Proof. The function

$$g(a) := (f_i(a))_{i\in I}$$

has obviously the desired properties. We have to show that it is unique. Let $h : A \to \prod_i B_i$ be another such function. If $g \neq h$, there is some element $a \in A$ such that $g(a) \neq h(a)$. Let $(b_i)_{i\in I} := h(a)$. For every $i \in I$, we have

$$b_i = (\mathrm{pr}_i \circ h)(a) = f_i(a)\,.$$

Hence $g(a) = (f_i(a))_i = (b_i)_i = h(a)$. A contradiction. □

Definition 2.3. The *disjoint union* of a sequence $(A_i)_{i\in I}$ of sets is the class

$$\dot{\bigcup_{i\in I}} A_i := \{\, \langle i, a\rangle \mid i \in I,\ a \in A_i \,\}\,.$$

Similarly, if A and B are classes then we can define their disjoint union as

$$A \mathbin{\dot\cup} B := (\{[0]\} \times A) \cup (\{[1]\} \times B)\,.$$

The k-th *insertion* is the canonical map

$$\mathrm{in}_k : A_k \to \dot{\bigcup_{i\in I}} A_i \quad \text{with} \quad \mathrm{in}_k(a) := \langle k, a\rangle\,.$$

Remark. If $A_i = A$, for all $i \in I$, then $\dot\bigcup_{i\in I} A_i = I \times A$.

Lemma 2.4. *Let B be a set and $(A_i)_{i\in I}$ a sequence of sets. For every sequence $(f_i)_{i\in I}$ of functions $f_i : A_i \to B$ there exists a unique function $g : \dot\bigcup_i A_i \to B$ such that*

$$g \circ \mathrm{in}_i = f_i\,, \qquad \textit{for all } i \in I\,.$$

Proof. The function

$$g\langle i, a\rangle := f_i(a)$$

has obviously the desired properties. We have to show that it is unique. Let $h : \dot\bigcup_i A_i \to B$ be another such function. If $g \neq h$ then there is some element $\langle k, a\rangle \in \dot\bigcup_i A_i$ such that $g\langle k, a\rangle \neq h\langle k, a\rangle$. We have

$$h\langle k, a\rangle = (h \circ \mathrm{in}_k)(a) = f_k(a) = g\langle k, a\rangle\,.$$

A contradiction. □

3. *Graphs and partial orders*

When considering relations it is frequently necessary to specify the sets they are over.

Definition 3.1. A *graph* is a pair $\langle A, R\rangle$ where $R \subseteq A \times A$ is a binary relation on A.

More generally one can consider sets together with several relations and functions. This will lead to the notion of a structure in Chapter B1.

Definition 3.2. Let $\langle A, R\rangle$ be a graph.

(a) R is *reflexive* if $\langle a, a\rangle \in R$, for all $a \in A$.

(b) R is *irreflexive* if $\langle a, a\rangle \notin R$, for all $a \in A$.

(c) R is *symmetric* if we have $\langle a, b\rangle \in R$ if, and only if, $\langle b, a\rangle \in R$, for all $a, b \in A$.

(d) R is *antisymmetric* if $\langle a, b\rangle \in R$ and $\langle b, a\rangle \in R$ implies $a = b$.

(e) R is *transitive* if $\langle a, b\rangle \in R$ and $\langle b, c\rangle \in R$ implies $\langle a, c\rangle \in R$, for all $a, b, c \in A$.

Note that, for the definition of reflexivity, it is important to specify the set A. If $\langle A, R\rangle$ is reflexive and $A \subset B$ then $\langle B, R\rangle$ is not reflexive.

Example. (a) The relation $A \times A$ is reflexive, symmetric, and transitive. It is irreflexive if, and only if, $A = \varnothing$, and it is antisymmetric if, and only if, A contains at most one element.

(b) The diagonal $\mathrm{id}_A = \{\langle a, a\rangle \mid a \in A\}$ is reflexive, symmetric, antisymmetric, and transitive. It is irreflexive if, and only if, $A = \varnothing$.

(c) The empty relation $\varnothing \subseteq A \times A$ is irreflexive, symmetric, antisymmetric, and transitive. It is reflexive if, and only if, $A = \varnothing$.

Definition 3.3. (a) A (non-strict) *partial order* is a graph $\langle A, \leq\rangle$ where $\leq$ is reflexive, transitive, and antisymmetric.

(b) A *strict partial order* is a graph $\langle A, <\rangle$ where $<$ is irreflexive and transitive.

(c) A partial order $\langle A, \leq\rangle$ is *linear,* or *total,* if

$$a \leq b \text{ or } b \leq a, \quad \text{for all } a, b \in A.$$

(d) Instead of saying that $\langle A, R\rangle$ is a partial or linear order we also say that R is a partial/linear order on A, or that R *orders* A partially/linearly.

(e) If $\mathfrak{A} = \langle A, \leq\rangle$ is a partial order, we denote by $\mathfrak{A}^{\mathrm{op}} := \langle A, \leq^{-1}\rangle$ the graph where the order relation is reversed. $\mathfrak{A}^{\mathrm{op}}$ is called the *opposite order.*

Remark. (a) To each non-strict partial order $\leq$ on A we can associate the strict partial order

$$a < b \quad :\text{iff} \quad a \leq b \text{ and } a \neq b.$$

Similarly, if $<$ is a strict partial order on A, we can define a non-strict version by

$$a \leq b \quad :\text{iff} \quad a < b \text{ or } a = b.$$

(b) If $\mathfrak{A}$ is a partial order then so is $\mathfrak{A}^{\mathrm{op}}$.

Example. (a) The subset relation $\subseteq$ is a partial order on $\mathbb{S}$.

(b) The usual ordering $\leq$ is a linear order on the rational numbers $\mathbb{Q}$.

(c) The divisibility relation

$$a \mid b \quad :\text{iff} \quad b = ac \text{ for some } c$$

is a partial order on the natural numbers $\mathbb{N}$.

Definition 3.4. Let $\mathfrak{A} = \langle A, \leq\rangle$ be a partial order.

(a) An *initial segment* of A is a subset $I \subseteq A$ such that $a \in I$ and $b \leq a$ implies $b \in I$. Similarly, a *final segment* of A is a subset $F \subseteq A$ such that $a \in F$ and $b \geq a$ implies $b \in F$.

(b) A set $X \subseteq A$ *generates* the segments

$$\Downarrow_{\mathfrak{A}} X := \{ a \in A \mid a \leq b \text{ for some } b \in X \},$$
and $$\Uparrow_{\mathfrak{A}} X := \{ a \in A \mid a \geq b \text{ for some } b \in X \}.$$

For $X = \{x\}$, we also write $\Downarrow_{\mathfrak{A}} x$ and $\Uparrow_{\mathfrak{A}} x$. Similarly, we define

$$\downarrow_{\mathfrak{A}} X := \{ a \in A \mid a < b \text{ for some } b \in X \},$$
and $$\uparrow_{\mathfrak{A}} X := \{ a \in A \mid a > b \text{ for some } b \in X \}.$$

Finally, we set

$$[a, b]_{\mathfrak{A}} := \Uparrow_{\mathfrak{A}} a \cap \Downarrow_{\mathfrak{A}} b \quad \text{and} \quad (a, b)_{\mathfrak{A}} := \uparrow_{\mathfrak{A}} a \cap \downarrow_{\mathfrak{A}} b.$$

(c) Let $X \subseteq A$ and $a \in X$. We call a the *greatest element* of X if $x \leq a$, for all $x \in X$. And we say that a is *maximal* if there is no $x \in X$ with $a < x$. *Least* and *minimal* elements are defined analogously. We denote the greatest element of X by $\max_{\mathfrak{A}} X$ and the least element by $\min_{\mathfrak{A}} X$, provided these elements exist.

(d) Let $X \subseteq A$. We say that a is an *upper bound* of X if $x \leq a$, for all $x \in X$. If a is an upper bound of X and $a \leq b$, for every other upper bound b of X, then a is the *least upper bound*, or *supremum*, of X. If the least upper bound of X exists, we denote it by $\sup_{\mathfrak{A}} X$.

The notion of a (*greatest*) *lower bound* is defined analogously. The greatest lower bound is also called the *infimum* of X. We denote it by $\inf_{\mathfrak{A}} X$. If the order $\mathfrak{A}$ is understood we will omit the subscript $\mathfrak{A}$ and we just write $\sup X$ and $\inf X$.

(e) A linearly ordered subset $C \subseteq A$ is called a *chain*.

Example. (a) Let $\mathfrak{Q} := \langle \mathbb{Q}, \leq \rangle$. The set $I := \{ x \in \mathbb{Q} \mid x < \sqrt{2} \}$ is an initial segment of $\mathfrak{Q}$. Every rational number $y > \sqrt{2}$ is an upper bound of I but I has no least upper bound.

(b) Consider $\langle \mathbb{N}, \mid \rangle$. Its least element is the number 1 and its greatest element is 0. The least upper bound of two elements $[k], [m] \in \mathbb{N}$ is their least common multiple $\operatorname{lcm}(k, m)$, and their greatest lower bound is their greatest common divisor $\gcd(k, m)$. The set $P \subseteq \mathbb{N}$ of all prime numbers has the least upper bound 0 and the greatest lower bound 1. The set $\{ 2^n \mid n \in \mathbb{N} \}$ of all powers of two forms a chain.

Exercise 3.1. Consider $\langle B, \subseteq \rangle$ where

$$B := \{ X \subseteq \mathbb{N} \mid X \text{ is finite or } \mathbb{N} \setminus X \text{ is finite} \} .$$

(a) Construct a set $X \subseteq B$ that has no minimal element.

(b) Construct a set $X \subseteq B$ with lower bounds but without infimum.

Lemma 3.5. *Let $\langle A, \leq \rangle$ be a partial order. If A is a set, the following statements are equivalent:*

(1) *Every subset $X \subseteq A$ has a supremum.*

(2) *Every subset $X \subseteq A$ has an infimum.*

Proof. We only prove (1) ⇒ (2). The other direction follows in exactly the same way. Let $X \subseteq A$ and set

$$C := \{ a \in A \mid a \text{ is a lower bound of } X \} .$$

By assumption, $c := \sup C$ exists. We claim that $\inf X = c$. Let $b \in X$. By definition, we have $a \leq b$, for all $a \in C$. Hence, b is an upper bound of C and we have $b \geq \sup C = c$. As b was arbitrary it follows that c is a lower bound of X. If a is an arbitrary lower bound of X, we have $a \in C$, which implies that $a \leq c$. Consequently, c is the greatest lower bound of X. □

Definition 3.6. A partial order $\langle A, \leq \rangle$ is *complete* if every subset $X \subseteq A$ has an infimum and a supremum.

Remark. Every complete partial order has a least element $\bot := \sup \emptyset$ and a greatest element $\top := \inf \emptyset$.

Example. (a) Let A be a set. The partial order $\langle \wp(A), \subseteq \rangle$ is complete. If $X \subseteq \wp(A)$ then

$$\sup X = \bigcup X \in \wp(A) \quad \text{and} \quad \inf X = \bigcap X \in \wp(A) .$$

(b) The order $\langle \mathbb{R}, \leq \rangle$ is complete. $\langle \mathbb{Q}, \leq \rangle$ is not since the set

$$\{ x \in \mathbb{Q} \mid x \leq \pi \}$$

has no least upper bound in $\mathbb{Q}$.

(c) The order $\langle \mathbb{N}, \leq \rangle$ is not complete since $\inf \emptyset$ and $\sup \mathbb{N}$ do not exist.

(d) Let $\mathfrak{A} = \langle A, \leq \rangle$ be an arbitrary partial order. We can construct a complete partial order $\mathfrak{C} = \langle C, \subseteq \rangle$ containing $\mathfrak{A}$ as follows. Let $C \subseteq \wp(A)$ be the set of all initial segments of A ordered by inclusion. The desired embedding $f : A \to C$ is given by $f(a) := \Downarrow_A a$.

Next we turn to the study of functions between partial orders. In particular, we will consider functions $f : A \to A$ mapping one partial order into itself. To simplify notation, we will write

$$f : \mathfrak{A} \to \mathfrak{B},$$

for partial orders $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$, to denote that f is a function $f : A \to B$.

Definition 3.7. Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be partial orders.
(a) A function $f : A \to B$ is *increasing* if

$$a \leq_A b \quad \text{implies} \quad f(a) \leq_B f(b), \qquad \text{for all } a, b \in A,$$

and f is *strictly increasing* if

$$a <_A b \quad \text{implies} \quad f(a) <_B f(b), \qquad \text{for all } a, b \in A.$$

(b) A function $f : A \to B$ is an *embedding* if we have

$$a \leq_A b \quad \text{iff} \quad f(a) \leq_B f(b), \qquad \text{for all } a, b \in A.$$

A bijective embedding is called an *isomorphism*. If there exists an isomorphism $f : A \to B$ then we say that $\mathfrak{A}$ and $\mathfrak{B}$ are *isomorphic* and we write $\mathfrak{A} \cong \mathfrak{B}$.

Remark. Every isomorphism is strictly increasing.

Exercise 3.2. Define a function that is

(a) increasing but not strictly increasing;

(b) strictly increasing but not an embedding;

(c) an embedding but not an isomorphism.

Exercise 3.3. Construct a strictly increasing function

$$f : \langle \mathbb{N}, \mid \rangle \to \langle \wp(\mathbb{N}), \subseteq \rangle.$$

Lemma 3.8. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be partial orders and $h : A \to B$ an increasing function. Let $C \subseteq A$ and $a \in A$.*

(a) *If a is an upper bound of C then $h(a)$ is an upper bound of $h[C]$.*

(b) *If a is a lower bound of C then $h(a)$ is a lower bound of $h[C]$.*

Lemma 3.9. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be partial orders and $h : A \to B$ an embedding. Let $C \subseteq A$ and $a \in A$.*

(a) $h(a) = \sup h[C]$ *implies* $a = \sup C$.

(b) $h(a) = \inf h[C]$ *implies* $a = \inf C$.

Proof. (a) Since h is an embedding it follows that $h(c) \leq_B h(a)$ implies $c \leq_A a$, for $c \in C$. Hence, a is an upper bound of C. To show that it is the least one, suppose that b is another upper bound of C. Then $c \leq_A b$, for $c \in C$, implies $h(c) \leq_B h(b)$. Hence, $h(b)$ is an upper bound of $h[C]$. Since $h(a)$ is the least such bound it follows that $h(a) \leq_B h(b)$. Consequently, we have $a \leq_A b$, as desired.

(b) h is also an embedding of $\langle A, \geq_A \rangle$ into $\langle B, \geq_B \rangle$. Hence, (b) follows from (a) by reversing the orders. □

Corollary 3.10. *Let $\langle F, \subseteq \rangle$ be a partial order with $F \subseteq \wp(A)$ and $C \subseteq F$.*

(a) $\bigcup C \in F$ *implies* $\sup C = \bigcup C$.

(b) $\bigcap C \in F$ *implies* $\inf C = \bigcap C$.

Proof. We can apply Lemma 3.9 to the inclusion map $F \to \wp(A)$. □

Corollary 3.11. *Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order. If $B \subseteq A$ is a nonempty set such that*

$$\inf_{\mathfrak{A}} X \in B \quad \textit{and} \quad \sup_{\mathfrak{A}} X \in B, \quad \textit{for every nonempty } X \subseteq B,$$

then $\mathfrak{B} := \langle B, \leq \rangle$ is a complete partial order where, for every nonempty subset $X \subseteq B$, we have

$$\inf_{\mathfrak{B}} X = \inf_{\mathfrak{A}} X \quad \textit{and} \quad \sup_{\mathfrak{B}} X = \sup_{\mathfrak{A}} X.$$

Proof. If $X \subseteq B$ is nonempty then, applying Lemma 3.9 to the inclusion map $\mathfrak{B} \to \mathfrak{A}$, it follows that

$$\inf_{\mathfrak{B}} X = \inf_{\mathfrak{A}} X \quad \text{and} \quad \sup_{\mathfrak{B}} X = \sup_{\mathfrak{A}} X\,.$$

In particular, $\inf_{\mathfrak{B}} X$ and $\sup_{\mathfrak{B}} X$ exist. For the empty set, it follows similarly that

$$\inf_{\mathfrak{B}} \varnothing = \sup_{\mathfrak{B}} B = \sup_{\mathfrak{A}} B \in B\,,$$
$$\text{and} \quad \sup_{\mathfrak{B}} \varnothing = \inf_{\mathfrak{B}} B = \inf_{\mathfrak{A}} B \in B\,.$$

Consequently, $\mathfrak{B}$ is complete. □

We have seen that although increasing functions preserve the ordering of elements they do not necessarily preserve supremums and infimums. Let us take a look at functions that do.

Definition 3.12. Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be partial orders. A function $f : A \to B$ is *continuous* if, whenever a nonempty chain $C \subseteq A$ has a supremum then $f[C]$ also has a supremum and we have

$$\sup f[C] = f(\sup C)\,.$$

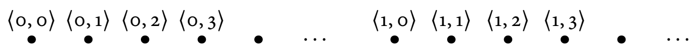
f is called *strictly continuous* if it is continuous and injective.

Remark. Every (strictly) continuous function is (strictly) increasing.

Exercise 3.4. Prove that continuous functions are increasing.

Example. (a) Let $\langle A, \leq \rangle$ be the linear order where $A = \mathbb{N} \cup \mathbb{N}$ and

$$\langle i, a \rangle \leq \langle k, b \rangle \quad :\text{iff} \quad i < k\,, \text{ or } i = k \text{ and } a \leq b\,.$$



The function $f : A \to A : \langle i, a \rangle \mapsto \langle i, a + 1 \rangle$ is not continuous. Consider the initial segment $X := \{0\} \times \mathbb{N} = {\downarrow}\langle 1, 0 \rangle \subseteq A$. We have $\sup X = \langle 1, 0 \rangle$ but

$$\sup f[X] = \langle 1, 0 \rangle < \langle 1, 1 \rangle = f(\langle 1, 0 \rangle)\,.$$

(b) Let A be a set and $\langle F, \subseteq \rangle$ the partial order with

$$F := \{\, X \subseteq A \mid A \setminus X \text{ is finite} \,\}\,.$$

For every bijective function $\sigma : A \to A$ we obtain a continuous mapping $f : F \to F$ by setting

$$f(X) := \{\, \sigma(x) \mid x \in X \,\}\,.$$

Lemma 3.13. *Every isomorphism $f : \mathfrak{A} \to \mathfrak{B}$ is strictly continuous.*

Proof. Let $C \subseteq A$ be a nonempty chain with supremum. For every $a \in C$, we have $a \leq \sup C$, which implies that $f(a) \leq f(\sup C)$. Hence,

$$\sup f[C] \leq f(\sup C)\,.$$

For the converse, let $b := \sup f[C]$. By Lemma 3.9, it follows that $\sup C = f^{-1}(b)$. □

4. *Fixed points and closure operators*

Many objects can be defined as solution to an equation of the form $x = f(x)$. Such solutions are called *fixed points* of the function f. For example, the solutions of a system of linear equations $Ax = b$ are exactly the fixed points of the function

$$f(x) := Ax + x - b\,.$$

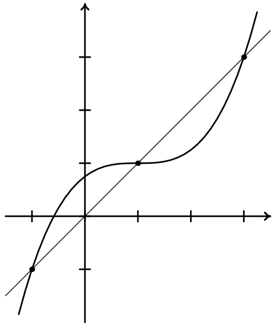


Figure 2.. Fixed points of $f(x) = \frac{1}{4}x^3 - \frac{3}{4}x^2 + \frac{3}{4}x + \frac{3}{4}$

Definition 4.1. Let $f : A \to A$ be a function. An element $a \in A$ with $f(a) = a$ is called a *fixed point* of f. The class of all fixed points of f is denoted by

$$\operatorname{fix} f := \{\, a \in A \mid f(a) = a \,\}\,.$$

We denote the *least* and *greatest* fixed point of f, if it exists, by

$$\operatorname{lfp} f := \min \operatorname{fix} f \quad \text{and} \quad \operatorname{gfp} f := \max \operatorname{fix} f\,.$$

Example. (a) Let $\langle \mathbb{R}, < \rangle$ be the order of the real numbers. The function

$$f(x) := \tfrac{1}{4}x^3 - \tfrac{3}{4}x^2 + \tfrac{3}{4}x + \tfrac{3}{4}$$

has 3 fixed points: $\operatorname{fix} f = \{-1, 1, 3\}$.

(b) Consider $\langle \mathbb{N}, \leq \rangle$. The function $f : \mathbb{N} \to \mathbb{N}$ with $f(n) := n + 1$ has no fixed points.

(c) Consider $\langle \wp[2], \subseteq \rangle$. The function $f : \wp[2] \to \wp[2]$ with

$$f(x) := \begin{cases} \{0\} & \text{if } x = \varnothing\,, \\ x & \text{otherwise}\,, \end{cases}$$

has the fixed points $\{0\}$, $\{1\}$, $\{0, 1\}$. It has no least fixed point.

(d) Consider $\langle F, \subseteq \rangle$ where

$$F := \{\, X \subseteq \mathbb{N} \mid X \text{ or } \mathbb{N} \setminus X \text{ is finite} \,\}\,.$$

The function $f : F \to F$ defined by

$$f(X) := \begin{cases} X \cup \{1 + \max X\} & \text{if } X \text{ is finite,} \\ X & \text{otherwise,} \end{cases}$$

has fixed points

$$\operatorname{fix} f = \{\, X \subseteq \mathbb{N} \mid \mathbb{N} \setminus X \text{ is finite} \,\}\,,$$

but no least one.

Exercise 4.1. Let $\mathfrak{A} = \langle \wp(\mathbb{N}), \subseteq \rangle$. Construct a function $f : \mathfrak{A} \to \mathfrak{A}$ that has a least fixed point but no greatest one.

Not every function has fixed points. The next theorem presents an important special case where we always have a least fixed point. In Section A3.3 we will collect further results about the existence of fixed points and methods to compute them.

Theorem 4.2 (Knaster, Tarski). *Let $\langle A, \leq \rangle$ be a complete partial order where A is a set. Every increasing function $f : A \to A$ has a least fixed point and we have*

$$\operatorname{lfp} f = \inf \{\, a \in A \mid f(a) \leq a \,\}\,.$$

Proof. Set $B := \{\, a \in A \mid f(a) \leq a \,\}$ and $b := \inf B$. For every $a \in B$, $b \leq a$ implies $f(b) \leq f(a) \leq a$, since f is increasing. Hence, $f(b)$ is a lower bound of B and it follows that $f(b) \leq \inf B = b$. This implies that $f(f(b)) \leq f(b)$ and, by definition of B, it follows that $f(b) \in B$. Hence, $f(b) \geq \inf B = b$. Consequently, we have $f(b) = b$ and b is a fixed point of f.

Let a be another fixed point of f. Then $f(a) = a$ implies $a \in B$ and we have $b = \inf B \leq a$. Hence, b is the least fixed point of f. □

Theorem 4.3. *Let $\langle A, \leq \rangle$ be a complete partial order where A is a set and let $f : A \to A$ be increasing. The set $F := \operatorname{fix} f$ is nonempty and $\mathfrak{F} := \langle F, \leq \rangle$ forms a complete partial order where, for $X \subseteq F$,*

$$\inf_{\mathfrak{F}} X = \sup_{\mathfrak{A}} \{ a \in A \mid a \leq \inf_{\mathfrak{A}} X \text{ and } f(a) \geq a \} ,$$
$$\sup_{\mathfrak{F}} X = \inf_{\mathfrak{A}} \{ a \in A \mid a \geq \sup_{\mathfrak{A}} X \text{ and } f(a) \leq a \} .$$

Proof. We have already shown in the preceding theorem that $F \neq \varnothing$. It remains to prove that $\mathfrak{F}$ is complete. For $X \subseteq A$, let $U := \Uparrow \sup_{\mathfrak{A}} X \subseteq A$ be the set of all upper bounds of X. If $Z \subseteq U$ then

$$\sup_{\mathfrak{A}} Z \geq \sup_{\mathfrak{A}} X \quad \text{and} \quad \inf_{\mathfrak{A}} Z \geq \sup_{\mathfrak{A}} X .$$

It follows that the partial order $\langle U, \leq \rangle$ is complete. Furthermore, if $a \in U$ and $x \in X$ then $a \geq x$ implies $f(a) \geq f(x)$. Hence, $f \upharpoonright U$ is an increasing function $U \to U$. By Theorem 4.2, it follows that

$$\sup_{\mathfrak{F}} X = \operatorname{lfp}(f \upharpoonright U) = \inf_{\mathfrak{A}} \{ a \in U \mid f(a) \leq a \} ,$$

as desired. The claim for $\inf_{\mathfrak{F}} X$ follows by applying the equation for $\sup_{\mathfrak{F}} X$ to the opposite order $\mathfrak{A}^{\mathrm{op}}$. □

Example. Consider a closed interval $[a, b] \subseteq \mathbb{R}$ of the real line.

(a) Since the order $\langle [a, b], < \rangle$ is complete, it follows by the Theorem of Knaster and Tarski that every increasing function $f : [a, b] \to [a, b]$ has a fixed point.

(b) Let $f : [0, 2] \to [0, 2]$ be the polynomial function

$$f(x) := \tfrac{1}{4}x^3 - \tfrac{3}{4}x^2 + \tfrac{3}{4}x + \tfrac{3}{4}$$

from Figure 2. We have $\{ x \mid f(x) \leq x \} = [1, 2]$ and $\operatorname{lfp} f = 1$.

(c) The order $\langle \mathbb{R}, < \rangle$ is not complete. Again, let $f : \mathbb{R} \to \mathbb{R}$ by the function from Figure 2. We have already seen that its fixed points are -1, 1, and 3. But the set

$$\{ x \mid f(x) \leq x \} = (-\infty, -1] \cup [1, 3]$$

has no minimal element.

As a special case of Theorem 4.3 we consider complete partial orders obtained via closure operators.

Definition 4.4. Let A be a class.

(a) A *closure operator* on A is a function $c : \wp(A) \to \wp(A)$ such that, for all $x, y \in \wp(A)$,

- $x \subseteq c(x)$,
- $c(c(x)) = c(x)$, and
- $x \subseteq y$ implies $c(x) \subseteq c(y)$.

(b) A set $x \subseteq A$ is *c-closed* if $c(x) = x$.

(c) A closure operator c has *finite character* if, for all sets $x \subseteq A$, we have

$$c(x) = \bigcup \{ c(x_0) \mid x_0 \subseteq x \text{ is finite} \} .$$

If c has finite character we also say that c is *algebraic*.

(d) A closure operator c is *topological* if we have

- $c(\varnothing) = \varnothing$ and
- $c(x \cup y) = c(x) \cup c(y)$, for all $x, y \in \wp(A)$.

Remark. Let c be a closure operator on A.

(a) The class of c-closed sets is $\operatorname{fix} c = \operatorname{rng} c$.

(b) If the class A is a set then it is c-closed.

Example. (a) Let V be a vector space. For $X \subseteq V$, let $\langle\!\langle X \rangle\!\rangle$ be the subspace of V spanned by X. The function $X \mapsto \langle\!\langle X \rangle\!\rangle$ is a closure operator with finite character.

(b) Let X be a topological space. For $A \subseteq X$, let $c(A)$ be the topological closure of A in X. Then c is a topological closure operator.

(c) Let A be a set and $a \in A$. The functions $c, d : \wp(A) \to \wp(A)$ with

$$c(X) := X \quad \text{and} \quad d(X) := X \cup \{a\}$$

are closure operators on A.

Exercise 4.2. Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order. For $X \subseteq A$, we define

$$c(X) := \{ \sup C \mid C \subseteq X \text{ is a nonempty chain with supremum} \} .$$

(a) Prove that the function c is a topological closure operator on A.
(b) Let $\mathfrak{B}$ be a second partial order and d the corresponding closure operator. Prove that a function $f : \mathfrak{A} \to \mathfrak{B}$ is continuous if, and only if, every d-closed set $X \in \operatorname{fix} d$ has a c-closed preimage $f^{-1}[X] \in \operatorname{fix} c$.

Exercise 4.3. Let $\langle A, \leq \rangle$ be a partial order. For sets $X \subseteq A$, we define

$$U(X) := \{ a \in A \mid a \text{ is an upper bound of } X \} ,$$
$$L(X) := \{ a \in A \mid a \text{ is a lower bound of } X \} .$$

Prove that the function $c : X \mapsto L(U(X))$ is a closure operator on A.

Lemma 4.5. *Let c be a closure operator on A and $x, y \subseteq A$ sets.*
(a) $c(x) \cup c(y) \subseteq c(x \cup y)$.
(b) $c(x \cup y) = c(c(x) \cup c(y))$.

Proof. (a) By monotonicity of c, we have $c(x) \subseteq c(x \cup y)$ and $c(y) \subseteq c(x \cup y)$.
(b) It follows from $x \cup y \subseteq c(x) \cup c(y)$ and (a) that

$$c(x \cup y) \subseteq c(c(x) \cup c(y)) \subseteq c(c(x \cup y)) = c(x \cup y) . \qquad \square$$

Lemma 4.6. *Let c be a closure operator on A with finite character. For every chain $C \subseteq \operatorname{fix} c$, we have*

$$c(\bigcup C) = \bigcup C .$$

Proof. By definition, we have $\bigcup C \subseteq c(\bigcup C)$. For the converse, let $x_0 \subseteq \bigcup C$ be finite. Since C is linearly ordered by $\subseteq$ there exists some element $x \in C$ with $x_0 \subseteq x$. Hence, we have $c(x_0) \subseteq c(x) = x \subseteq \bigcup C$. It follows that

$$c(\bigcup C) = \bigcup \{ c(x_0) \mid x_0 \subseteq \bigcup C \text{ finite} \} \subseteq \bigcup C . \qquad \square$$

If c is a closure operator, the set $\mathcal{C} := \operatorname{fix} c$ of c-closed sets has the following properties.

Definition 4.7. A set $\mathcal{C} \subseteq \wp(A)$ is called a *system of closed sets* if we have

- $A \in \mathcal{C}$ and
- $\bigcap Z \in \mathcal{C}$, for every $Z \subseteq \mathcal{C}$.

A pair $\langle A, \mathcal{C} \rangle$ where $\mathcal{C} \subseteq \wp(A)$ is a system of closed sets is called a *closure space.*

Lemma 4.8. (a) *If c is a closure operator on A then* $\operatorname{fix} c$ *forms a system of closed sets.*
(b) *If $\mathcal{C} \subseteq \wp(A)$ is a system of closed sets then the mapping*

$$c : X \mapsto \bigcap \{ C \in \mathcal{C} \mid X \subseteq C \}$$

defines a closure operator on A with $\operatorname{fix} c = \mathcal{C}$.

The following theorem states that the family of c-closed sets forms a complete partial order. We can use this result to prove that a given partial order $\mathfrak{A}$ is complete by defining a closure operator whose closed sets are exactly the elements of $\mathfrak{A}$. An example of such a proof is provided in Corollary 4.17.

Theorem 4.9. *Let A be a set and c a closure operator on A. The graph $\langle F, \subseteq \rangle$ with $F := \operatorname{fix} c$ forms a complete partial order with*

$$\inf X = \bigcap X \quad \text{and} \quad \sup X = c(\bigcup X) , \quad \text{for all } X \subseteq F .$$

Proof. Since closure operators are increasing we can apply Theorem 4.3. By Lemma 4.8 (b), it follows that

$$\begin{aligned} \sup X &= \bigcap \{ Z \subseteq A \mid Z \supseteq \bigcup X \text{ and } c(Z) \subseteq Z \} \\ &= \bigcap \{ Z \subseteq A \mid Z \supseteq \bigcup X \text{ and } c(Z) = Z \} \\ &= c(\bigcup X) , \end{aligned}$$

and $\inf X = \bigcup \{ Z \subseteq A \mid Z \subseteq \bigcap X \text{ and } c(Z) \supseteq Z \}$

$$= \bigcup \{ Z \subseteq A \mid Z \subseteq \bigcap X \}$$
$$= \bigcap X . \qquad \square$$

Corollary 4.10. *Let c be a closure operator on A and set $F := \operatorname{fix} c$. The operator c is continuous if we consider it as a function*

$$c : \langle \wp(A), \subseteq \rangle \to \langle F, \subseteq \rangle .$$

Proof. For a nonempty chain $X \subseteq \wp(A)$, we have

$$\begin{aligned} c(\sup X) = c(\bigcup X) &\subseteq c(\bigcup c[X]) = \sup c[X] \\ &\subseteq \sup \{ c(\sup X) \} = c(\sup X) . \end{aligned} \qquad \square$$

As an application of closure operators we consider equivalence relations.

Definition 4.11. (a) A binary relation $\sim \subseteq A \times A$ is an *equivalence relation* on A if it is reflexive, symmetric, and transitive.

(b) Let $\sim \subseteq A \times A$ be an equivalence relation. If A is a set, we define the $\sim$*-class* of an element $a \in A$ by

$$[a]_\sim := \{ b \in A \mid b \sim a \} .$$

For proper classes A, we set

$$[a]_\sim := \operatorname{cut} \{ b \in A \mid b \sim a \} .$$

Note that, despite the name, a $\sim$-class is always a set. We denote the class of all $\sim$-classes by

$$A/{\sim} := \{ [a]_\sim \mid a \in A \} .$$

Example. (a) The diagonal id_A is the smallest equivalence relation on A. The largest one is the full relation $A \times A$.

(b) The isomorphism relation $\cong$ is an equivalence relation on the class of all partial orders.

Lemma 4.12. *Let $\sim$ be an equivalence relation on A and $a, b \in A$. Then*

$$a \sim b \quad \text{iff} \quad [a]_\sim = [b]_\sim \quad \text{iff} \quad [a]_\sim \cap [b]_\sim \neq \varnothing .$$

Remark. Let A be a set. A *partition* of A is a set $P \subseteq \wp(A)$ of nonempty subsets of A such that $A = \bigcup P$ and $p \cap q = \varnothing$, for all $p, q \in P$ with $p \neq q$.

If $\sim$ is an equivalence relation on A then $A/{\sim}$ forms a partition on A. Conversely, given a partition P of A, we can define an equivalence relation $\sim_P$ on A with $A/{\sim_P} = P$ by setting

$$a \sim_P b \quad :\text{iff} \quad \text{there is some } p \in P \text{ with } a, b \in p .$$

Definition 4.13. Let A be a set and $R \subseteq A \times A$ a binary relation on A. The *transitive closure* of R is the relation

$$\mathrm{TC}(R) := \bigcap \{ S \subseteq A \times A \mid S \supseteq R \text{ is transitive} \} .$$

Since the family of transitive relations is closed under intersections we can use Lemma 4.8 (b) to prove that TC is a closure operator.

Lemma 4.14. *Let A be a class.* TC *is a closure operator on $A \times A$.*

Exercise 4.4. Prove Lemma 4.14.

Lemma 4.15. *If $R \subseteq A \times A$ is a symmetric relation then so is* $\mathrm{TC}(R)$.

Proof. Let $S := \mathrm{TC}(R) \cap (\mathrm{TC}(R))^{-1}$. Since R is symmetric we have $R \subseteq S$. We claim that S is transitive.

Let $\langle a, b \rangle, \langle b, c \rangle \in S$. Then $\langle a, b \rangle, \langle b, c \rangle \in \mathrm{TC}(R)$ and $\langle b, a \rangle, \langle c, b \rangle \in \mathrm{TC}(R)$. Therefore, we have $\langle a, c \rangle \in \mathrm{TC}(R)$ and $\langle c, a \rangle \in \mathrm{TC}(R)$. This implies that $\langle a, c \rangle \in S$, as desired.

We have shown that S is a transitive relation containing R. By the definition of TC it follows that $\mathrm{TC}(R) \subseteq S = \mathrm{TC}(R) \cap \mathrm{TC}(R)^{-1}$. This implies that $\mathrm{TC}(R)^{-1} = \mathrm{TC}(R)$. Hence, $\mathrm{TC}(R)$ is symmetric. $\square$

Lemma 4.16. *Let $R \subseteq A \times A$ be a binary relation.*

(a) *The smallest reflexive relation containing R is $R \cup \mathrm{id}_A$.*

(b) *The smallest symmetric relation containing* R *is* $R \cup R^{-1}$.

(c) *The smallest transitive relation containing* R *is* $\mathrm{TC}(R)$.

(d) *The smallest equivalence relation containing* R *is* $\mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)$.

Proof. (a) $R \cup \mathrm{id}_A$ is obviously reflexive and it contains R. Conversely, suppose that $S \supseteq R$ is reflexive. Then $\mathrm{id}_A \subseteq S$ implies that $R \cup \mathrm{id}_A \subseteq S$.

(b) is proved analogously.

(c) Let $S \supseteq R$ be transitive. Then the intersection in the definition of TC contains S. Hence, $\mathrm{TC}(R) \subseteq S$. Furthermore, we have $R \subseteq \mathrm{TC}(R)$ by definition. It remains to prove that $\mathrm{TC}(R)$ is transitive.

Let $\langle a, b\rangle, \langle b, c\rangle \in \mathrm{TC}(R)$. Then we have $\langle a, b\rangle, \langle b, c\rangle \in S$, for every transitive relation $S \supseteq R$. Hence, we have $\langle a, c\rangle \in S$, for each such relation S. This implies that $\langle a, c\rangle \in \mathrm{TC}(R)$.

(d) Set $E := \mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)$. Clearly, we have $R \subseteq E$ and, if $S \supseteq R$ is an equivalence relation then $E \subseteq S$. Hence, it is remains to prove that E is an equivalence relation. It is transitive by (c), symmetric by Lemma 4.15, and E is reflexive since $\mathrm{id}_A \subseteq \mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)$. □

Corollary 4.17. *Let* A *be a set and* $F \subseteq \wp(A \times A)$ *the set of all equivalence relations on* A. *Then* $\langle F, \subseteq\rangle$ *forms a complete partial order. If* $X \subseteq F$ *is nonempty then we have*

$$\inf X = \bigcap X \quad \textit{and} \quad \sup X = \mathrm{TC}(\bigcup X)\,.$$

Proof. By Lemma 4.16, we have $F = \mathrm{fix}\, c$ where c is the closure operator with

$$c(R) := \mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)\,.$$

The relation $E := \bigcup X$ is reflexive and symmetric since X is nonempty. Hence, we have $\mathrm{TC}(E \cup E^{-1} \cup \mathrm{id}_A) = \mathrm{TC}(E)$. Consequently, the claim follows from Theorem 4.9. □

A3. Ordinals

1. Well-orders

When defining stages we frequently used the fact that any class of stages has a minimal element. In this section we study arbitrary orders with this property.

Definition 1.1. Let $\langle A, R\rangle$ be a graph.

(a) An element $a \in A$ is R-*minimal* if $\langle b, a\rangle \in R$ implies $b = a$.

(b) A relation R is *left-narrow* if $R^{-1}(a)$ is a set, for every set $a \in \mathrm{rng}\, R$.

(c) R is *well-founded* if every nonempty subset $B \subseteq A$ contains an R-minimal element. A left-narrow, well-founded linear order is called a *well-order.*

Example. (a) $\langle \mathbb{N}, \leq\rangle$ is a well-order.

(b) $\langle \mathbb{N}, \mid\,\rangle$ is a well-founded partial order.

(c) The membership relation $\in$ is a well-founded partial order on $\mathbb{S}$. It is a well-order on the class of all stages.

(d) $\langle \wp(\mathbb{N}), \subseteq\rangle$ is not well-founded.

(e) A partial order $\langle A, \leq\rangle$ is left-narrow if, and only if, $\Downarrow a$ is a set, for all $a \in A$.

Exercise 1.1. Prove that $\langle \wp(\mathbb{N}), \subseteq\rangle$ is not well-founded.

Lemma 1.2. *If* $\langle A, R\rangle$ *is a well-founded graph and* $B \subseteq A$ *then* $\langle B, R|_B\rangle$ *is also well-founded.*

Proof. Every nonempty subset $C \subseteq B$ is also a nonempty subset of A and has an R-minimal element. □

Lemma 1.3. *If $\langle A, \leq \rangle$ is a well-founded and left-narrow partial order, there exists no infinite sequence $(a_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ such that $a_n \neq a_{n+1}$ and $a_{n+1} \leq a_n$, for all n.*

Proof. If there exists such an infinite sequence then the class $\operatorname{rng} \bar{a} = \{ a_n \mid n \in \mathbb{N} \}$ is nonempty and has no $\leq$-minimal element. Furthermore, $\operatorname{rng} \bar{a} \subseteq \Downarrow a_0$ is a set since the order is left-narrow. □

The reason why well-founded relations are of interest is that these are exactly those relations that admit proofs by induction. As the theorem below shows we can prove that every element of a well-founded partial order $\langle A, \leq \rangle$ satisfies a given property φ by showing that, if every element $b < a$ satisfies φ then a also satisfies φ.

Lemma 1.4. *Let $\langle A, \leq \rangle$ be a well-founded, left-narrow partial order. Every nonempty subclass $X \subseteq A$ has a minimal element.*

Proof. Let $X \subseteq A$ be nonempty and fix some element $a \in X$. $\Downarrow a$ is a set since $\leq$ is left-narrow. Hence, $Y := X \cap \Downarrow a$ is a nonempty subset of A and has a minimal element b. Note that $b \in Y \subseteq X$ and, if $c \in X$ is some element with $c \leq b \leq a$, then $c \in Y$. Therefore, it follows that b is also a minimal element of X. □

Theorem 1.5. *Let $\langle A, \leq \rangle$ be a well-founded, left-narrow partial order. If $X \subseteq A$ is a subclass such that*

$$\downarrow a \subseteq X \quad \textit{implies} \quad a \in X\,, \qquad \textit{for all } a \in A\,,$$

then $X = A$.

Proof. Let $X \subseteq A$ be a class as above. For a contradiction, suppose that $X \neq A$. Fix some element $a \in A \setminus X$. Since $\leq$ is left-narrow $B := \Downarrow a \setminus X$ is a set. Hence, B has a $\leq$-minimal element b. It follows that $\downarrow b \subseteq A \setminus B \subseteq X$, which implies that $b \in X$. Contradiction. □

Example. Consider the well-order $\langle \mathbb{N}, < \rangle$ of the natural numbers. Suppose that $X \subseteq \mathbb{N}$ is a subset such that we can show that

$$b \in X\,, \text{ for all } b < a\,, \quad \text{implies} \quad a \in X\,,$$

then we have $X = \mathbb{N}$. Proofs based on this fact are called 'proofs by induction'. The above corollary states that such proofs work not only for the natural numbers but for all well-orders.

Let $\langle A, \leq \rangle$ be a well-order. The minimal element of a given subclass $X \subseteq A$ is unique since A is linearly ordered. Therefore, if A is not empty, it has a least element $\perp$. The *successor* a^+ of an element $a \in A$ is the least element of the class $\uparrow a$. a^+ is defined for every element of A except for the greatest one. An element that is neither the least one nor a successor of some other element is called a *limit*.

It turns out that we can define a canonical well-founded order on the class Wo of all well-orders.

Remark. Note that speaking of 'the class of all well-orders' is sloppy language since, by definition, a class contains only sets. Instead, we should call Wo 'the class of all well-orders that are sets'.

Definition 1.6. Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be well-orders. We define

$$\mathfrak{A} < \mathfrak{B} \quad :\text{iff} \quad A \text{ is a set and, for some } b \in B\,, \text{ there exists an isomorphism } f : A \to \downarrow_B b\,.$$

(Note that, if f exists, it is necessarily a set because A and $\downarrow_B b$ are both sets.)

To prove that this defines an order on Wo we need some technical lemmas.

Lemma 1.7. *Let $\langle A, \leq \rangle$ be a well-order. If $f : A \to A$ is a strictly increasing function then $a \leq f(a)$, for all $a \in A$.*

Proof. Suppose that there exists some $a \in A$ with $a > f(a)$. Let a_0 be the minimal such element. By minimality of a_0 we have

$$f(a_0) \leq f(f(a_0)) .$$

On the other hand, since f is strictly increasing we have

$$f(f(a_0)) < f(a_0) .$$

Contradiction. □

Lemma 1.8. *Let $\langle A, \leq \rangle$ be a well-order and $I \subseteq A$. The following statements are equivalent:*

(1) *I is a proper initial segment of A.*

(2) *$I = {\downarrow_A} a$, for some $a \in A$.*

(3) *I is an initial segment of A and I is non-isomorphic to A.*

Proof. (1) ⇒ (2) If I is a proper subclass of A then $A \setminus I$ is nonempty and has a least element a. Consequently, we have $I = {\downarrow} a$.

(2) ⇒ (3) Let $I = {\downarrow} a$. Suppose there exists an isomorphism $f : A \to I$. By Lemma 1.7, we have $f(a) \geq a$. Hence, $f(a) \notin I = \operatorname{rng} f$. Contradiction.

(3) ⇒ (1) is trivial. □

Corollary 1.9. *$<$ is a strict partial order on* Wo.

Proof. We can see immediately from the definition that $<$ is transitive. Suppose that $\mathfrak{A} < \mathfrak{A}$, for some well-order $\mathfrak{A} = \langle A, \leq \rangle$. By definition there exists an element $a \in A$ and an isomorphism $f : A \to {\downarrow_A} a$. This contradicts the preceding lemma. □

Lemma 1.10. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be well-orders. There exists at most one isomorphism $f : A \to B$.*

Proof. Let $f, g : A \to B$ be isomorphisms. Then so is $g \circ f^{-1} : B \to B$. In particular, $g \circ f^{-1}$ is strictly increasing. By Lemma 1.7, we obtain

$$f(a) \leq (g \circ f^{-1})(f(a)) = g(a) , \qquad \text{for all } a \in A .$$

Similarly, we derive $g(a) \leq f(a)$, for all a. It follows that $f = g$. □

We still have to prove that $<$ is linear. Unfortunately, this is not true. The following theorem states that, for all well-orders $\mathfrak{A}$ and $\mathfrak{B}$, exactly one of the following conditions holds $\mathfrak{A} < \mathfrak{B}$ or $\mathfrak{A} \cong \mathfrak{B}$ or $\mathfrak{A} > \mathfrak{B}$. In order for $<$ to be linear, the second condition should read $\mathfrak{A} = \mathfrak{B}$. We will see how to deal with this problem in the next section.

Theorem 1.11. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be well-orders. Exactly one of the following statements holds:*

(1) *There exists an isomorphism $f : A \to J$ where $J \subset B$ is a proper initial segment of B.*

(2) *There exists an isomorphism $f : A \to B$.*

(3) *There exists an isomorphism $f : I \to B$ where $I \subset A$ is a proper initial segment of A.*

(*f might be a proper class.*)

Proof. We claim that

$$f := \{ \langle a, b \rangle \in A \times B \mid \text{there is an isomorphism } {\downarrow} a \to {\downarrow} b \} .$$

is the desired isomorphism.

First, we show that $\langle a_0, b_0 \rangle, \langle a_1, b_1 \rangle \in f$ implies

$$a_0 < a_1 \quad \text{iff} \quad b_0 < b_1 .$$

For a contradiction, suppose that $a_0 < a_1$ and $b_0 \geq b_1$. We have isomorphisms

$$h_0 : {\downarrow} a_0 \to {\downarrow} b_0 \quad \text{and} \quad h_1 : {\downarrow} a_1 \to {\downarrow} b_1 .$$

The restriction of h_1 to $\downarrow a_0$ is an isomorphism

$$h_1 \upharpoonright \downarrow a_0 : \downarrow a_0 \to \downarrow h_1(a_0) .$$

Composing it with h_0^{-1} yields an isomorphism

$$(h_1 \upharpoonright \downarrow a_0) \circ h_0^{-1} : \downarrow b_0 \to \downarrow h_1(a_0) .$$

But this contradicts $h_1(a_0) < b_1 \leq b_0$, by Lemma 1.8.

Therefore, f is the graph of a strictly increasing function. We claim that $\operatorname{dom} f$ and $\operatorname{rng} f$ are initial segments of, respectively, A and B. Suppose, for a contradiction, that there are elements $a < b$ such that $a \notin \operatorname{dom} f$ and $b \in \operatorname{dom} f$. By definition, there is an isomorphism $h : \downarrow b \to \downarrow f(b)$. Its restriction to $\downarrow a$ yields an isomorphism $h \upharpoonright \downarrow a : \downarrow a \to \downarrow h(a)$ which shows that $a \in \operatorname{dom} f$. Contradiction. Analogously, it follows that $\operatorname{rng} f$ is an initial subclass of B.

It remains to show that $\operatorname{dom} f = A$ or $\operatorname{rng} f = B$. Suppose, otherwise. Let a be the minimal element of $A \setminus \operatorname{dom} f$ and b the minimal one of $B \setminus \operatorname{rng} f$. Then $\operatorname{dom} f = \downarrow a$ and $\operatorname{rng} f = \downarrow b$ and f is an isomorphism from $\downarrow a$ to $\downarrow b$. By definition, we therefore have $\langle a, b \rangle \in f$. Contradiction. □

Corollary 1.12. *For all $\mathfrak{A}, \mathfrak{B} \in \mathrm{Wo}$, we have either*

$$\mathfrak{A} < \mathfrak{B} \quad \textit{or} \quad \mathfrak{A} \cong \mathfrak{B} \quad \textit{or} \quad \mathfrak{A} > \mathfrak{B} .$$

We conclude this section with two remarks about continuous mappings between well-orders. The following lemma provides a simple criterion to check whether a mapping between well-orders is continuous.

Lemma 1.13. *Let $\langle A, \leq \rangle$ be a well-order and $\langle B, \leq \rangle$ an arbitrary partial order. A function $f : A \to B$ is continuous if, and only if, it satisfies the following conditions:*

(1) $f(a^+) \geq f(a)$, *for all $a \in A$,*

(2) $f(a) = \sup \{ f(b) \mid b < a \}$, *for every limit $a \in A$.*

Proof. ($\Rightarrow$) By definition, every continuous function satisfies (2). Furthermore, $a^+ = \sup \{a, a^+\}$ implies that $f(a^+) = \sup \{f(a), f(a^+)\}$.

($\Leftarrow$) For the other direction, suppose that f satisfies (1) and (2). First, we show that f is increasing. Suppose otherwise and let $a \in A$ be the minimal element such that $f(b) > f(a)$, for some $b < a$. Note that a is not the minimal element of A since $b < a$. If a were a limit then (2) would imply that

$$f(a) = \sup \{ f(x) \mid x < a \} \geq f(b) .$$

Contradiction. Hence, a must be a successor and we have $a = c^+$, for some $c \in A$. By choice of a, we have $f(x) \leq f(c)$, for all $x \leq c$. In particular, $f(c) \geq f(b) > f(a)$. But (1) implies $f(a) = f(c^+) \geq f(c)$. Again a contradiction.

We have shown that f is increasing. But what we really want to prove is that it is continuous. Let $X \subseteq A$ be a nonempty subset of A with supremum $a := \sup X$. If $b \in X$ then $b \leq a$ implies $f(b) \leq f(a)$. Hence, $f(a)$ is an upper bound of $f[X]$. To prove that $f(a)$ is its least upper bound we distinguish two cases.

If $a \in X$ then $f(a) \in f[X]$, which implies $f(a) = \sup f[X]$.

If $a \notin X$ then $a = \sup X$ is a limit and, for every $b < a$, there is some $x \in X$ with $b \leq x$. If c is another upper bound of $f[X]$ then $f(b) \leq f(x) \leq c$. By (2), it follows that

$$f(a) = \sup \{ f(b) \mid b < a \} \leq \sup \{ f(x) \mid x \in X \} \leq c .$$

Hence, $f(a)$ is the least upper bound of $f[X]$. □

Lemma 1.14. *Let $\langle A, \leq \rangle$ be a well-order and $f : A \to A$ strictly continuous. If $a \geq f(\bot)$ then*

$$\max \{ b \in A \mid f(b) \leq a \}$$

exists.

Proof. If a is the greatest element of A, we can set $b := a$. Otherwise, we have $f(a^+) > f(a) \geq a$, by Lemma 1.7. Hence, there are elements $x \in A$ with $f(x) > a$. Let c be the least such element. We have $c > \perp$ since $f(c) > a \geq f(\perp)$. If c were a limit then, by choice of c, we would have

$$f(c) = \sup\{ f(x) \mid x < c \} \leq a < f(c) .$$

A contradiction. Hence, c is a successor and there exists some $b \in A$ with $c = b^+$. By choice of c, we have $f(b) \leq a$. Furthermore, if $x > b$ then $x \geq c$, which implies that $f(x) \geq f(c) > a$. Therefore, b is the desired element. □

2. Ordinals

We have seen that there exists a well-order on Wo if one does not distinguish between isomorphic orders. We would like to define a subclass $\mathrm{On} \subseteq \mathrm{Wo}$ of *ordinals* such that, for each well-order $\mathfrak{A}$, there exists a unique element $\mathfrak{B} \in \mathrm{On}$ that is isomorphic to $\mathfrak{A}$.

We will present two approaches to do so. The usual one – due to von Neumann – has the disadvantage that it requires the Axiom of Replacement. Without it we cannot prove that, for every well-order α, there exists an isomorphic von Neumann ordinal. Therefore, we will adopt a different approach. The relation $\cong$ forms a congruence (see Section B1.4 below) on the class of all well-orders. A first try might thus consist in representing a well-ordering by its congruence class. Unfortunately, the class of all well-orders isomorphic to a given one is not a set. Hence, with this definition one could not form sets of ordinals. Instead of considering *all* isomorphic well-orders we will therefore only take some of them.

Definition 2.1. The *order type* of a well-order $\mathfrak{A}$ is the set

$$\mathrm{ord}(\mathfrak{A}) := [\mathfrak{A}]_{\cong} = \mathrm{cut}\{ \mathfrak{B} \mid \mathfrak{B} \text{ is a well-order isomorphic to } \mathfrak{A} \} .$$

The elements of $\mathrm{On} := \mathrm{rng}(\mathrm{ord})$ are called *ordinals.*

Instead of a subclass $\mathrm{On} \subseteq \mathrm{Wo}$ the above definition results in a function $\mathrm{ord} : \mathrm{Wo} \to \mathrm{On}$. Below we will see that there exists a canonical way to associate with every ordinal $\alpha \in \mathrm{On}$ a well-order $f(\alpha) \in \mathrm{Wo}$. Using this injection $f : \mathrm{On} \to \mathrm{Wo}$ we can identify the class On with its image $f[\mathrm{On}] \subseteq \mathrm{Wo}$.

First, let us show that the mapping $\mathrm{ord} : \mathrm{Wo} \to \mathrm{On}$ has the desired property of characterising a well-order up to isomorphism.

Lemma 2.2. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be well-orders that are sets. There exists an isomorphism $f : \mathfrak{A} \to \mathfrak{B}$ if, and only if,* $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{B})$.

Proof. If $f : \mathfrak{A} \to \mathfrak{B}$ is an isomorphism then a well-order $\mathfrak{C}$ is isomorphic to $\mathfrak{A}$ if, and only if, it is isomorphic to $\mathfrak{B}$. Therefore $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{B})$. Conversely, suppose $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{B})$. Since $\mathfrak{A}$ is a well-order isomorphic to $\mathfrak{A}$, we have $\mathrm{ord}(\mathfrak{A}) \neq \emptyset$. Fix an arbitrary element $\mathfrak{C} \in \mathrm{ord}(\mathfrak{A})$. By definition, $\mathfrak{C}$ is isomorphic to $\mathfrak{A}$ and to $\mathfrak{B}$. Consequently, $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic. □

Remark. We will prove in Lemma A4.5.3 with the help of the Axiom of Replacement that any two well-ordered proper classes are isomorphic. In particular, it follows that in the above lemma we can drop the requirement of $\mathfrak{A}$ and $\mathfrak{B}$ being sets.

Definition 2.3. Let $\mathfrak{On} := \langle \mathrm{On}, < \rangle$ where the ordering $<$ is defined by

$$\mathrm{ord}(\mathfrak{A}) < \mathrm{ord}(\mathfrak{B}) \quad :\text{iff} \quad \mathfrak{A} < \mathfrak{B} .$$

For $\alpha \in \mathrm{On}$, recall that $\downarrow\alpha = \{ \beta \in \mathrm{On} \mid \beta < \alpha \}$.

Remark. (a) The ordering $<$ is well-defined since $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{A}')$ and $\mathrm{ord}(\mathfrak{B}) = \mathrm{ord}(\mathfrak{B}')$ implies that $\mathfrak{A} < \mathfrak{B}$ iff $\mathfrak{A}' < \mathfrak{B}'$.

(b) In the chapters on set theory we will strictly distinguish between an ordinal α and the set $\downarrow\alpha$. But in the remainder of the book we will usually drop the arrow and write α in both cases.

Combining Corollaries 1.9 and 1.12 and Lemma 2.2 it follows that On is well-ordered.

Theorem 2.4. $\mathfrak{On}$ *is a well-order.*

The notions of a *successor ordinal* and a *limit ordinal* are defined in the same way as for arbitrary well-orders. Recall that we denote the successor of α by α^+. Furthermore, we define

$$0 := \operatorname{ord}\langle\varnothing,\varnothing\rangle\,, \quad 1 := 0^+\,, \quad 2 := 1^+\,, \ldots$$

The first limit ordinal is $\omega := \operatorname{ord}\langle\mathbb{N},\leq\rangle$.

Lemma 2.5. *Let* $\alpha,\beta \in \mathrm{On}$. *If* $\alpha \leq \beta$ *then* $S(\alpha) \subseteq S(\beta)$.

Proof. If $\alpha = \beta$, the claim is trivial. Therefore, we assume that $\alpha < \beta$. Let $\mathfrak{A} = \langle A, \leq_A\rangle \in \alpha$ and $\mathfrak{B} = \langle B, \leq_B\rangle \in \beta$. Since $\alpha < \beta$ there exists an isomorphism $f : A \to {\downarrow_B} b$, for some $b \in B$. Set $\mathfrak{B}_0 := \langle {\downarrow_B} b, \leq_B\rangle$. Then $\operatorname{ord}\mathfrak{B}_0 = \alpha$ and $\mathfrak{A} \in \operatorname{ord}\mathfrak{B}_0$ implies that $S(\mathfrak{A}) \subseteq S(\mathfrak{B}_0)$. Since $S(\mathfrak{B}_0) \subseteq S(\mathfrak{B})$ it follows that $S(\mathfrak{A}) \subseteq S(\mathfrak{B})$. We have shown that $S(x) \subseteq S(y)$, for all $x \in \alpha$ and $y \in \beta$. Consequently, we have $S(\alpha) \subseteq S(\beta)$. □

To every ordinal α we can associate a canonical well-order of type α.

Lemma 2.6. $\langle {\downarrow}\alpha, \leq\rangle$ *is a well-order of type* $\operatorname{ord}\langle {\downarrow}\alpha, \leq\rangle = \alpha$.

Proof. Let $\langle A, \leq\rangle$ be a well-order of type $\operatorname{ord}\langle A, \leq\rangle = \alpha$. We claim that the function $f : A \to \mathrm{On}$ with

$$f(a) := \operatorname{ord}\langle {\downarrow_A} a, \leq\rangle$$

is an isomorphism $f : A \to {\downarrow}\alpha$.

f is strictly increasing since, if $a < b$ then ${\downarrow_A} a$ is a proper initial segment of ${\downarrow_A} b$. By Lemma 1.8 and Lemma 2.2, it follows that

$$f(a) = \operatorname{ord}\langle {\downarrow_A} a, \leq\rangle < \operatorname{ord}\langle {\downarrow_A} b, \leq\rangle = f(b)\,.$$

Furthermore, f is surjective since, for every $\beta < \alpha$, there exists some $a \in A$ with

$$\beta = \operatorname{ord}\langle {\downarrow_A} a, \leq\rangle = f(a)\,. \qquad \square$$

Lemma 2.7. On *is not a set.*

Proof. Suppose that On is a set. Since On is well-ordered there exists some ordinal $\alpha \in \mathrm{On}$ with $\alpha = \operatorname{ord}\langle \mathrm{On}, \leq\rangle$. We have just seen that $\operatorname{ord}\langle {\downarrow}\alpha, \leq\rangle = \alpha$. Therefore, there exists an isomorphism $f : {\downarrow}\alpha \to \mathrm{On}$. But ${\downarrow}\alpha$ is a proper initial segment of On. This contradicts Lemma 1.8. □

Lemma 2.8. *A subclass* $X \subseteq \mathrm{On}$ *is a set if, and only if, it has an upper bound.*

Proof. ($\Leftarrow$) If $X \subseteq \mathrm{On}$ has an upper bound α then $X \subseteq {\Downarrow}\alpha$. Since ${\Downarrow}\alpha$ is a set the claim follows.

($\Rightarrow$) Suppose that X is a set. Since On is a proper class there exists some ordinal $\alpha \in \mathrm{On} \setminus S(X)$. We claim that α is an upper bound of X. Suppose there exists some $\beta \in X$ with $\beta \nleq \alpha$. Then $\alpha < \beta$ and we have $\alpha \subseteq S(\alpha) \subseteq S(\beta) \in S(X)$, which implies that $\alpha \in S(X)$. This contradicts our choice of α. □

Corollary 2.9. *Every set of ordinals has a supremum.*

Another consequence is the following special case of the Axiom of Replacement which we will introduce in Section A4.5.

Corollary 2.10. *If* $F : \mathrm{On} \to \mathrm{On}$ *is increasing then* $F[{\downarrow}\alpha]$ *is a set, for all* $\alpha \in \mathrm{On}$.

Proof. Suppose that F is increasing. Then we have $F(\beta) \leq F(\alpha)$, for all $\beta < \alpha$. Consequently, $F(\alpha)$ is an upper bound of $F[{\downarrow}\alpha]$ and, by Lemma 2.8, it follows that $F[{\downarrow}\alpha]$ is a set. □

Let us give a simpler characterisation of the relation $\leq$ on well-orders.

Lemma 2.11. *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be well-orders. Then* $\mathfrak{A} \leq \mathfrak{B}$ *if, and only if, there exists a strictly increasing function* $f : A \to B$.

Proof. ($\Rightarrow$) If $\mathfrak{A} \leq \mathfrak{B}$ then, by definition, there exists an isomorphism $f : A \to I$ between A and an initial segment I of B. In particular, $f : A \to B$ is a strictly increasing function.

(⇐) Suppose that $f : A \to B$ is a strictly increasing function and let $C := \operatorname{rng} f$. Since $C \subseteq B$ is well-ordered there exists an isomorphism $g : C \to I \subseteq \mathrm{On}$ between C and an initial segment of On. Similarly, there is some isomorphism $h : B \to J \subseteq \mathrm{On}$. We claim that

$$k := h^{-1} \circ g \circ f : A \to B$$

is the desired isomorphism between A and an initial segment of B. Since f, g, and h^{-1} are isomorphisms so is k. What remains to be shown is that k is in fact well-defined, that is, $I = \operatorname{rng} g \subseteq \operatorname{rng} h = J$.

We claim that $g(c) \leq h(c)$, for all $c \in C$. Since I and J are initial segments this implies that $I \subseteq J$. For a contradiction, suppose that there is some $c \in C$ with $g(c) > h(c)$ and let c be the minimal such element. Note that, since g and h are strictly increasing and $\operatorname{rng} g$ and $\operatorname{rng} h$ are initial segments we must have

$$g(c) = \min\left(I \setminus \operatorname{rng}(g \upharpoonright {\downarrow_C} c)\right)$$
and $$h(c) = \min\left(J \setminus \operatorname{rng}(h \upharpoonright {\downarrow_B} c)\right).$$

By choice of c, we have $\operatorname{rng}(g \upharpoonright {\downarrow_C} c) \subseteq \operatorname{rng}(h \upharpoonright {\downarrow_B} c)$. But, by the above equations, this implies that $g(c) \leq h(c)$. A contradiction. □

In order to use the theory of ordinals for proofs about arbitrary sets one usually needs to define a well-order on a given set. In general this is only possible if one assumes the Axiom of Choice. Until we introduce this axiom the following theorem will serve as a stopgap. Once we have defined the cardinality of a set in Section A4.2 it will turn out that the ordinal the theorem talks about is $\alpha = |A|^+$.

Theorem 2.12 (Hartogs). *For every set A there exists an ordinal α such that there are no injective functions ${\downarrow}\alpha \to A$.*

Proof. For a contradiction, suppose that there exists a set A such that, for every ordinal α, there is an injective function $f_\alpha : {\downarrow}\alpha \to A$. Let $A_\alpha := \operatorname{rng} f_\alpha \subseteq A$ and set

$$R_\alpha := \{\, \langle a, b\rangle \in A_\alpha \times A_\alpha \mid f_\alpha^{-1}(a) \leq f_\alpha^{-1}(b) \,\}.$$

By construction, $f_\alpha : \langle {\downarrow}\alpha, \leq\rangle \to \langle A_\alpha, R_\alpha\rangle$ is an isomorphism. Hence, by the definition of an ordinal, we have

$$S(\alpha) \subseteq S(\langle A_\alpha, R_\alpha\rangle).$$

Since $R_\alpha \subseteq A \times A \in \wp^3(A) \subseteq \wp^3(S(A))$ it follows that

$$\langle A_\alpha, R_\alpha\rangle = \big\{\{A_\alpha\},\ \{A_\alpha, R_\alpha\}\big\} \subseteq \wp^4(S(A)).$$

We have shown that

$$\alpha \subseteq S(\alpha) \subseteq S(\langle A_\alpha, R_\alpha\rangle) \subseteq \wp^4(S(A)), \quad \text{for all } \alpha \in \mathrm{On}.$$

Consequently, $\mathrm{On} \subseteq \wp^5(S(A))$, which implies that On is a set. This contradicts Lemma 2.7. □

Von Neumann ordinals

We conclude this section with an alternative definition of ordinals. This definition is simpler and the resulting ordinals have many nice properties such that $\alpha = {\downarrow}\alpha$ and $\sup X = \bigcup X$. The only disadvantage is that one needs an additional axiom in order to prove that every well-order is isomorphic to some ordinal. Intuitively, we define a *von Neumann ordinal* to be the set of all smaller ordinals, that is, $\alpha := {\downarrow}\alpha$. As usual, the actual definition is more technical and we have to verify afterwards that it has the desired effect.

Definition 2.13. A set α is a *von Neumann ordinal* if it is transitive and linearly ordered by the membership relation $\in$. We denote the class of all von Neumann ordinals by On_0 and we set $\mathfrak{On}_0 := \langle \mathrm{On}_0, \in\rangle$.

Example. The set $[n] = \{[0], \dots, [n-1]\}$ is a von Neumann ordinal, for each $n \in \mathbb{N}$.

Lemma 2.14. *If $\alpha \in \mathrm{On}_0$ and $\beta \in \alpha$ then $\beta \in \mathrm{On}_0$.*

Proof. First, note that $\beta \in \alpha$ implies $\beta \subseteq \alpha$. As α is linearly ordered by $\in$ it therefore follows that so is $\beta \subseteq \alpha$.

It remains to prove that β is transitive. Suppose that $\eta \in \gamma \in \beta$. By transitivity of α, we have $\eta, \gamma, \beta \in \alpha$. Since α is linearly ordered by $\in$ we know that the relation $\in$, restricted to α, is transitive. Hence, $\eta \in \gamma$ and $\gamma \in \beta$ implies that $\eta \in \beta$. □

Remark. Note that, for $\alpha \in \mathrm{On}_0$, we have

$$\downarrow\alpha = \{ \beta \in \mathrm{On}_0 \mid \beta \in \alpha \} .$$

Hence, $\alpha = \downarrow\alpha$ and our definition of a von Neumann ordinal coincides with the intuitive one.

Exercise 2.1. Suppose that $\alpha = \{\beta_0, \dots, \beta_{n-1}\}$ is a von Neumann ordinal with $n < \omega$ elements. Prove, by induction on n, that $\alpha = [n]$.

Theorem 2.15. $\mathfrak{On}_0$ *is a well-order.*

Proof. $\in$ is irreflexive since we have $a \notin a$, for all sets. Furthermore, $\in$ is transitive on On_0 since, $\alpha \in \beta \in \gamma$ implies $\alpha \in \gamma$, by transitivity of γ. Consequently, $\in$ is a strict partial order on On_0. Since $\in$ is well-founded on any class it remains to prove that it is linear.

Let $\alpha, \beta \in \mathrm{On}_0$. The set $\gamma := \alpha \cap \beta$ is transitive by Lemma A1.2.4. As α is linearly ordered by $\in$ so is $\gamma \subseteq \alpha$. Therefore, $\gamma \in \mathrm{On}_0$. Furthermore, γ is an initial segment of α since $\delta \in \eta \in \gamma$ implies $\delta \in \gamma$, by transitivity. By Lemma 1.8, it follows that $\gamma = \alpha$ or $\gamma = \downarrow\delta = \delta$, for some $\delta \in \alpha$. Hence, we either have $\gamma = \alpha$ or $\gamma \in \alpha$. Similarly, we can prove that either $\gamma = \beta$ or $\gamma \in \beta$. Since $\gamma \notin \gamma = \alpha \cap \beta$ it follows that either $\gamma \notin \alpha$ or $\gamma \notin \beta$. Consequently, we either have $\beta = \gamma \in \alpha$ or $\alpha = \gamma \in \beta$ or $\alpha = \gamma = \beta$. □

Exercise 2.2. Show that $\alpha^+ = \alpha \cup \{\alpha\}$, for every $\alpha \in \mathrm{On}_0$.

Lemma 2.16. On_0 *is not a set.*

Proof. On_0 is transitive and well-ordered by $\in$. If it were a set, it would be an element of itself. □

On_0 is linearly ordered by $\in$. The following sequence of lemmas contains several characterisations of this ordering. In particular, we show that the mapping

$$\mathrm{ord} : \langle \mathrm{On}_0, \in \rangle \to \langle \mathrm{On}, < \rangle$$

is strictly increasing. After we have introduced the Axiom of Replacement in Section A4.5 we will prove that it is actually an isomorphism.

Lemma 2.17. *Let $\alpha, \beta \in \mathrm{On}_0$. We have $\alpha \in \beta$ if, and only if, $\alpha \subset \beta$.*

Proof. ($\Rightarrow$) was already proved in Lemma A1.2.2. For ($\Leftarrow$), suppose that $\alpha \notin \beta$. By Lemma 2.15, it follows that $\alpha = \beta$ or $\beta \in \alpha$. Since $\alpha \subset \beta$ we therefore have $\beta \subset \beta$ or $\beta \in \beta$. Contradiction. □

Lemma 2.18. *Let $\alpha, \beta \in \mathrm{On}_0$. If $f : \alpha \to \beta$ is an isomorphism between α and an initial segment of β then $f = \mathrm{id}_\alpha$.*

Proof. Suppose that $f \neq \mathrm{id}_\alpha$ and let $\gamma \in \alpha$ be the minimal element of α such that $\delta := f(\gamma) \neq \gamma$. Since f is an isomorphism we have $\xi = f(\xi) \in f(\gamma) = \delta$, for all $\xi \in \gamma$. Hence, $\gamma \subseteq \delta$. Since $\delta \neq \gamma$ it follows that $\gamma \subset \delta$, which implies, by Lemma 2.17, that $\gamma \in \delta$. But $\gamma \notin \mathrm{rng}\, f$, since $f(\xi) = \xi \in \gamma$, for $\xi \in \gamma$, and $f(\xi) \ni f(\gamma) = \delta$, for $\xi \ni \gamma$. Hence, $\mathrm{rng}\, f$ is not an initial segment of β. Contradiction. □

Lemma 2.19. *Let $\alpha, \beta \in \mathrm{On}_0$. The following statements are equivalent:*

(1) $\alpha \in \beta$.

(2) $\alpha \subset \beta$.

(3) $S(\alpha) \in S(\beta)$.

(4) $\langle \alpha, \in \rangle < \langle \beta, \in \rangle$.

Proof. (1) $\Leftrightarrow$ (2) was already shown in Lemma 2.17.

(1) $\Rightarrow$ (3) $a \in b$ implies $S(a) \in S(b)$, for arbitrary sets a and b.

(3) $\Rightarrow$ (1) If $\alpha \notin \beta$ then, by Lemma 2.15, we either have $\alpha = \beta$ or $\beta \in \alpha$. Consequently, either $S(\alpha) = S(\beta)$ or $S(\beta) \in S(\alpha)$. It follows that $S(\alpha) \notin S(\beta)$.

(2) ⇒ (4) If $\alpha \subseteq \beta$, the identity $\mathrm{id}_\alpha : \alpha \to \alpha \subseteq \beta$ is an isomorphism from α to an initial segment of β. Hence, $\alpha < \beta$.

(4) ⇒ (2) If $\alpha < \beta$, there exists an isomorphism $f : \alpha \to I \subset \beta$ between α and a proper initial subset of β. By the preceding lemma, it follows that $f = \mathrm{id}_\alpha$ and $\alpha = I \subset \beta$. □

It follows that, similarly to On, the von Neumann ordinals are linearly ordered by the relation <. If we could prove that every well-order is isomorphic to some von Neumann ordinal, we could use On_0 as representatives instead of On.

Corollary 2.20. *For all $\alpha, \beta \in \mathrm{On}_0$, we have*

$$\alpha < \beta \quad \text{or} \quad \alpha = \beta \quad \text{or} \quad \alpha > \beta\,.$$

Infimum and supremum of sets of von Neumann ordinals can be computed especially easily.

Lemma 2.21. *Let $X \subseteq \mathrm{On}_0$.*

(a) *If X is nonempty then* $\inf X = \bigcap X$.

(b) *If X has an upper bound then* $\sup X = \bigcup X$.

Proof. (a) Since X is nonempty it has a minimal element α, which is also the infimum of X. Clearly, $\bigcap X \subseteq \alpha$. Conversely, if $\beta \in \alpha$ then $\beta \in \gamma$, for all $\gamma \in X$, which implies $\beta \in \bigcap X$. It follows that $\inf X = \alpha = \bigcap X$.

(b) Note that we have $\alpha \leq \beta$ iff $\alpha \subseteq \beta$, for all von Neumann ordinals $\alpha, \beta \in \mathrm{On}_0$.

Clearly, we have $\alpha \subseteq \bigcup X$, for all $\alpha \in X$. Hence, $\bigcup X$ is an upper bound of X. Conversely, let β be an upper bound of X. Then $\alpha \subseteq \beta$, for all $\alpha \in X$, which implies that $\bigcup X \subseteq \beta$. □

The reason why there might be less von Neumann ordinals than elements of On is that each von Neumann ordinal is contained in a new stage. That is, we have exactly one von Neumann ordinal for every stage.

Lemma 2.22. *The function $f : \mathrm{On}_0 \to H(\mathbb{S})$ defined by $f(\alpha) := S(\alpha)$ is an isomorphism between On_0 and the class of all stages.*

Proof. By Lemma 2.19 it follows that f is injective and increasing. Suppose that it is not surjective. Let S be the minimal stage such that $S \notin \mathrm{rng}\, f$, and set

$$X := \{\, \alpha \in \mathrm{On}_0 \mid S(\alpha) \in S \,\}\,.$$

Since $X \subseteq S$, X is a set and, hence, a proper initial segment of On_0. Therefore, there is some $\alpha \in \mathrm{On}_0$ such that $X = {\downarrow}\alpha$. Since $S(\beta) \in S$, for all $\beta \in \alpha$, it follows that $S(\alpha) \subseteq S$. By choice of S, we have $S(\alpha) \neq S$. Hence, $S(\alpha) \in S$, which implies that $\alpha \in X = {\downarrow}\alpha$. Contradiction. □

Definition 2.23. For $\alpha \in \mathrm{On}_0$, we set $S_\alpha := S(\alpha)$.

Remark. In On_0 we have finally found the indices to enumerate the cumulative hierarchy

$$S_0 \subset S_1 \subset \dots \subset S_\alpha \subset S_{\alpha+1} \subset \cdots$$

The class of all stages can be written in the form

$$H(\mathbb{S}) = \{\, S_\alpha \mid \alpha \in \mathrm{On}_0 \,\}\,,$$

and we have $\mathbb{S} = \bigcup \{\, S_\alpha \mid \alpha \in \mathrm{On}_0 \,\}$.

Definition 2.24. The *rank* $\rho(a)$ of a set a is the von Neumann ordinal α such that $S(a) = S_\alpha$.

Remark. (a) For $\alpha \in \mathrm{On}_0$, we have $\rho(\alpha) = \alpha$.

(b) Note that

$$\mathrm{cut}\, A = \{\, x \in A \mid \rho(x) \leq \rho(y) \text{ for all } y \in A \,\}\,.$$

Lemma 2.25. *A class X is a set if, and only if, $\{\, \rho(x) \mid x \in X \,\}$ is bounded.*

Exercise 2.3. Prove the preceding lemma.

3. Induction and fixed points

The importance of ordinals stems from the fact that they allow proofs and constructions by *induction*. The next theorem follows immediately from Theorem 1.5.

Theorem 3.1 (Principle of Transfinite Induction). *Let $I \subseteq \mathrm{On}$ be an initial segment of* On. *If $X \subseteq I$ is a class such that, for every $\alpha \in I$,*

$$\Downarrow\alpha \subseteq X \quad \textit{implies} \quad \alpha \in X$$

then $X = I$.

Usually one applies this theorem in the following way. If one wants to prove that all ordinals satisfy a certain property φ, it is sufficient to prove that

- 0 satisfies φ;
- if α satisfies φ then so does α^+;
- if δ is a limit ordinal and every $\alpha < \delta$ satisfies φ then so does δ.

Transfinite induction is not only useful for proofs but also to define sequences. For a class A, we set

$$A^{<\infty} := \{\, f \mid f : \Downarrow\beta \to a \text{ for some } \beta \in \mathrm{On} \text{ and } a \subseteq A \,\}\,.$$

Lemma 3.2. *Let H be a partial function $H : \mathbb{S}^{<\infty} \to \mathbb{S}$. For each ordinal $\alpha \in \mathrm{On}$, there exists at most one function $f : \Downarrow\alpha \to \mathbb{S}$ such that f is a set and*

$$f(\beta) = H(f \restriction \Downarrow\beta)\,, \qquad \textit{for all } \beta < \alpha\,.$$

Proof. Suppose that f and g both satisfy the above condition. We apply the Principle of Transfinite Induction to prove that $f = g$. Let

$$X := \{\, \beta \in \Downarrow\alpha \mid f(\beta) = g(\beta) \,\}\,.$$

If $\beta < \alpha$ is an ordinal such that $\Downarrow\beta \subseteq X$ then $f \restriction \Downarrow\beta = g \restriction \Downarrow\beta$, which implies that

$$f(\beta) = H(f \restriction \Downarrow\beta) = H(g \restriction \Downarrow\beta) = g(\beta)\,.$$

Consequently, $\beta \in X$. By the Principle of Transfinite Induction, it follows that $X = \Downarrow\alpha$, that is, $f = g$. □

Remark. If a function f satisfies the conditions of the preceding lemma then so does $f \restriction I$, for every initial segment $I \subseteq \operatorname{dom} f$. In particular, if $f : \Downarrow\alpha \to \mathbb{S}$ and $g : \Downarrow\beta \to \mathbb{S}$ are two such functions with $\alpha \leq \beta$ then $f = g \restriction \Downarrow\alpha$.

Definition 3.3. Let H be a partial function $H : \mathbb{S}^{<\infty} \to \mathbb{S}$ and let f_α be the unique function $f_\alpha : \Downarrow\alpha \to \mathbb{S}$ such that f_α is a set and

$$f_\alpha(\beta) = H(f_\alpha \restriction \Downarrow\beta)\,, \qquad \text{for all } \beta < \alpha\,.$$

Let $I \subseteq \mathrm{On}$ be the class of all α such that f_{α^+} is defined. (Note that I is an initial segment since if $\alpha \in I$ and $\beta < \alpha$ then $f_{\beta^+} = f_{\alpha^+} \restriction \Downarrow\beta$.)

We say that H defines the function F by *transfinite recursion* if

$$\operatorname{dom} F = I \qquad \text{and} \qquad F(\alpha) = f_{\alpha^+}(\alpha)\,, \quad \text{for all } \alpha \in \operatorname{dom} F\,.$$

Theorem 3.4 (Principle of Transfinite Recursion). *Every partial function $H : \mathbb{S}^{<\infty} \to \mathbb{S}$ defines a unique function F by transfinite recursion. We have $F \notin \operatorname{dom} H$ and*

$$F(\alpha) = H(F \restriction \Downarrow\alpha)\,, \qquad \textit{for all } \alpha \in \operatorname{dom} F\,.$$

Proof. The existence of F follows immediately from the definition. Note that, by the remark after Lemma 3.2, we have $f_\beta(\alpha) = f_\gamma(\alpha)$, for all $\beta, \gamma > \alpha$. Consequently,

$$F(\alpha) = f_{\alpha^+}(\alpha) = f_\beta(\alpha)\,, \quad \text{for all } \beta > \alpha\,,$$

which implies that

$$F \restriction {\downarrow}\alpha = f_\beta \restriction {\downarrow}\alpha\,, \quad \text{for all } \beta \geq \alpha\,.$$

Therefore, it follows that

$$F(\alpha) = f_{\alpha^+}(\alpha) = H(f_{\alpha^+} \restriction {\downarrow}\alpha) = H(F \restriction {\downarrow}\alpha)\,.$$

In particular, if F is a set then $F = f_\alpha$, for some α. Hence, we have $\operatorname{dom} F = \operatorname{dom} f_\alpha = {\downarrow}\alpha$. Since $\alpha \notin \operatorname{dom} F$ it follows that f_{α^+} does not exists. Hence, $H(f_\alpha) = H(F)$ is undefined and $F \notin \operatorname{dom} H$. If F is a proper class then we trivially have $F \notin \operatorname{dom} H$. ☐

Remark. After we have introduced the Axiom of Replacement we can actually show that, if $H : \mathbb{S}^{<\infty} \to \mathbb{S}$ is a total function then $\operatorname{dom} F = \mathrm{On}$.

At the moment we can prove this statement only for the special case where $\operatorname{rng} H$ is a set.

Lemma 3.5. *Let A be a set. If $H : A^{<\infty} \to A$ is a total function that defines the function F by transfinite recursion then F is a proper class with* $\operatorname{dom} F = \mathrm{On}$.

Proof. Note that $\operatorname{rng} F \subseteq \operatorname{rng} H \subseteq A$ is a set. If $\operatorname{dom} F = {\downarrow}\alpha \subset \mathrm{On}$ then $F \in A^{\downarrow\alpha} \subset A^{<\infty} = \operatorname{dom} H$ in contradiction to Theorem 3.4. ☐

Usually definitions by transfinite recursion have the following simpler form. Given an element $a \in A$ and two functions $s : A \to A$ and $h : \wp(A) \to A$ one can construct a unique function $f : I \to A$ such that

- $f(0) = a$;
- $f(\beta^+) = s(f(\beta))$; and
- $f(\delta) = h(f[{\downarrow}\delta])$, for limit ordinals δ.

Example. We can define addition and multiplication of ordinals as follows. By transfinite recursions, we first define the function $\beta \mapsto \alpha + \beta$ by

$$\begin{aligned} \alpha + 0 &:= \alpha\,, \\ \alpha + \beta^+ &:= (\alpha + \beta)^+\,, \\ \alpha + \delta &:= \sup\{\, \alpha + \beta \mid \beta < \delta \,\}\,, \qquad \text{for limit ordinals } \delta\,, \end{aligned}$$

and then we define the function $\beta \mapsto \alpha \cdot \beta$ by

$$\begin{aligned} \alpha \cdot 0 &:= 0\,, \\ \alpha \cdot \beta^+ &:= \alpha \cdot \beta + \alpha,\,, \\ \alpha \cdot \delta &:= \sup\{\, \alpha \cdot \beta \mid \beta < \delta \,\}\,, \qquad \text{for limit ordinals } \delta\,. \end{aligned}$$

By the above theorem, we know that these operations are defined on some initial segment of On and that they are uniquely determined by these equations. Below we will give a different, more concrete definition of addition and multiplication.

Definitions by transfinite recursion are special cases of so-called *inductive fixed points*. Consider a partial order $\langle A, \leq \rangle$ and a function $f : A \to A$. If certain conditions on $\langle A, \leq \rangle$ and f are satisfied, one can compute a fixed point of f in the following way. Starting with some element $a \in A$ we construct the sequence $a,\ f(a),\ f(f(a)), \dots$. If it converges, its limit will be a fixed point of f. The next definition formalises this process.

Definition 3.6. Let $\langle A, \leq \rangle$ be a partial order. A function $f : A \to A$ is *inductive* over an element $a \in A$ if there exists an increasing function $F : I \to A$ where $I \subset \mathrm{On}$ is an initial segment of On such that F is a proper class and we have

$$\begin{aligned} F(0) &= a\,, \\ F(\beta^+) &= f(F(\beta))\,, \\ \text{and} \quad F(\delta) &= \sup F[{\downarrow}\delta]\,, \qquad \text{for limits } \delta\,. \end{aligned}$$

We call F the *fixed-point induction* of f over a. The element $F(\alpha)$ is also called the α-th *stage* of the induction.

Remark. (a) Note that, if A is a set then, by the Principle of Transfinite Recursion, there exists a unique function $F : \mathrm{On} \to A$ satisfying the above equations provided we can show that, for every limit δ, the supremum $\sup F[\downarrow\delta]$ exists. If, furthermore, we can prove that $F(\beta^+) \geq F(\beta)$, for all β, then it follows that f is inductive.

(b) Every fixed-point induction F is continuous, by Lemma 1.13.

Example. (a) The function $f : \mathrm{On} \to \mathrm{On} : \alpha \mapsto \alpha^+$ is inductive. Its fixed-point induction over 0 is the identity function $F : \mathrm{On} \to \mathrm{On} : \alpha \mapsto \alpha$.

(b) Let $f : \mathbb{S} \to \mathbb{S}$ be the function with $f(a) := \wp(a)$. The fixed-point induction of f over $\varnothing$ is the function $F : \mathrm{On}_0 \to \mathbb{S}$ with

$$F(\alpha) := S_\alpha .$$

(c) The graph of addition

$$A := \{ (x, y, z) \in \mathbb{N}^3 \mid x + y = z \}$$

is the least fixed point of the function $f : \wp(\mathbb{N}^3) \to \wp(\mathbb{N}^3)$ with

$$f(R) := \{ (x, 0, x) \mid x \in \mathbb{N} \} \cup \{ (x, y + 1, z + 1) \mid (x, y, z) \in R \} .$$

Its fixed-point induction over $\varnothing$ is the function

$$F(\alpha) := \begin{cases} \{ (x, y, z) \mid x + y = z , y < \alpha \} & \text{if } \alpha < \omega , \\ A & \text{if } \alpha \geq \omega . \end{cases}$$

(d) Let $\langle V, E \rangle$ be a graph. The function

$$f : \wp(V \times V) \to \wp(V \times V)$$

defined by $f(R) := E \cup E \circ R$ is increasing. Let F be the fixed-point induction of f over $\varnothing$. Then

$$F(0) = \varnothing ,$$
$$F(1) = E ,$$
$$F(2) = E \cup E \circ E ,$$
$$F(3) = E \cup E \circ E \cup E \circ E \circ E ,$$

and, generally, we have

$$F(n) = \textstyle\bigcup_{k<n} E^k , \qquad \text{for } n < \omega ,$$
and
$$F(\alpha) = \textstyle\bigcup_{k<\omega} E^k , \qquad \text{for } \alpha \geq \omega .$$

Hence, the inductive fixed point of f is the relation $\bigcup_{k<\omega} E^k = \mathrm{TC}(E)$.

(e) We consider the following simple game between two players. It is played on a graph $\langle V, E \rangle$ where the set of vertices $V = V_0 \cup V_1$ is partitioned into vertices V_0 that belong to player 0 and vertices V_1 belonging to player 1. At the start of the game a pebble is placed on the starting position $v_0 \in V$. In every round one of the players moves this pebble along an edge to a new vertex. If the pebble is on a vertex in V_0 then player 0 can choose where to move, if it is on a vertex in V_1 then player 1 may move. Hence, a play of the game determines a path $v_0, \dots, v_n$ through the graph. If at some point the pebble is on a vertex in V_i without outgoing edge then player i loses. If none of the players manage to manœuvre his opponent into such a situation then the game never stops and *both* players lose. The *winning region* W_i for player i is the set of all vertices w such that, if we start the game in w, then player i has a strategy to win the game. We can compute these winning regions by the fixed-point induction F_i of the function

$$f_i(X) := \{ x \in V_i \mid \text{there is some } y \in X \text{ with } \langle x, y \rangle \in E \} \cup \{ x \in V_{1-i} \mid \text{every } y \text{ with } \langle x, y \rangle \in E \text{ is element of } X \} .$$

Note that $F_i(1)$ is the set of all vertices $x \in V_{1-i}$ without outgoing edge. Generally, $F_i(n)$ contains all vertices such that player i has a strategy to win the game in at most n rounds.

Exercise 3.1. Let $\langle V, E \rangle$ be a graph. Prove that $\mathrm{TC}(E) = \bigcup_{n<\omega} E^n$.

If the fixed point induction of a function f converges, its limit is a fixed point of f.

Lemma 3.7. *Let F be the fixed-point induction of a function f. If $F(\alpha) = F(\alpha^+)$ then $F(\alpha) \in \operatorname{fix} f$.*

Proof. $F(\alpha)$ is a fixed point of f since $f(F(\alpha)) = F(\alpha^+) = F(\alpha)$. □

Thus, we can use the fixed point induction F of f to compute a fixed point provided F converges.

Lemma 3.8. *Let F be the fixed-point induction of a function f. If $F(\alpha) = F(\alpha^+)$ then $F(\alpha) = F(\beta)$, for all $\beta \geq \alpha$.*

Proof. We prove the claim by induction on β. If $\beta = \alpha$ then the claim is trivial. For the successor step, we have

$$F(\beta^+) = f(F(\beta)) = f(F(\alpha)) = F(\alpha^+) = F(\alpha)\,.$$

Finally, if $\delta > \alpha$ is a limit ordinal, then

$$\begin{aligned} F(\delta) = \sup\{\, F(\beta) \mid \beta < \delta \,\} &= \sup\{\, F(\beta) \mid \alpha \leq \beta < \delta \,\} \\ &= \sup\{F(\alpha)\} = F(\alpha)\,. \end{aligned}$$ □

If the universe A is a set, every fixed-point induction stabilises at some ordinal. Intuitively, the reason is that the size of the universe A is bounded. Therefore, if we repeat the application of f long enough, we will obtain some element $a \in A$ that already appeared in the sequence.

Theorem 3.9. *Let $\langle A, \leq \rangle$ be a partial order where A is a set. Let $f : A \to A$ be inductive over $a \in A$ and $F : \mathrm{On} \to A$ the corresponding fixed-point induction. There exists some ordinal α such that $F(\alpha) = F(\beta)$, for all $\beta \geq \alpha$.*

Proof. By Theorem 2.12, there exists an ordinal γ such that there is no injective function $\downarrow\gamma \to A$. We claim that there is some $\alpha < \gamma$ such that $F(\alpha) = F(\alpha^+)$. By Lemma 3.8, it then follows that $F(\beta) = F(\alpha)$, for all $\beta \geq \alpha$. Suppose that $F(\alpha) \neq F(\alpha^+)$, for all $\alpha < \gamma$. Since F is increasing it follows that $F \upharpoonright \downarrow\gamma : \downarrow\gamma \to A$ is injective. This contradicts our choice of γ. □

Remark. This proof actually shows that $\alpha < |A|^+$ where $|A|$ is the cardinality of A (see Section A4.2).

Definition 3.10. Let $f : A \to A$ be inductive and $F : \mathrm{On} \to A$ the corresponding fixed-point induction. The minimal ordinal α such that $F(\alpha) = F(\alpha^+)$ is called the *closure ordinal* of the induction and the element $F(\infty) := F(\alpha)$ is the *inductive fixed point* of f over a.

Remark. If A is a set, every inductive function $f : A \to A$ has an inductive fixed point.

Example. Let $\langle A, R \rangle$ be a graph. The *well-founded part* of R is the maximal subset $B \subseteq A$ such that $\langle B, R|_B \rangle$ is well-founded and, for all $\langle a, b \rangle \in R$ with $b \in B$, we also have $a \in B$. We can compute B as inductive fixed point over $\varnothing$ of the function

$$f(X) := \{\, x \in A \mid R^{-1}(x) \subseteq X \cup \{x\} \,\}\,.$$

If we want to apply the above machinery to compute fixed points, we need methods to show that a given function f is inductive. Basically, there are two conditions a function f has to satisfy. The sequence obtained by iterating f has to be linearly ordered and its supremum must exists.

Definition 3.11. Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order.

(a) $\mathfrak{A}$ is *inductively ordered* if every chain $C \subseteq A$ that is a set has a supremum.

(b) A function $f : A \to A$ is *inflationary* if $f(a) \geq a$, for all $a \in A$.

Remark. (a) Every inductively ordered set has a least element $\perp$ since the set $\varnothing$ is linearly ordered.

(b) Every complete partial order is inductively ordered.

(c) $\langle \mathrm{On}, \leq \rangle$ is inductively ordered.

(d) If $\langle A, \leq \rangle$ is a well-order then according to Lemma 1.7 all strictly continuous functions $f : A \to A$ are inflationary.

Example. (a) The partial order $\langle F, \subseteq \rangle$ where

$$F := \{\, X \subseteq \mathbb{N} \mid X \text{ is finite} \,\}$$

is not inductively ordered since the chain

$$[0] \subset [1] \subset [2] \subset \cdots \subset [n] \subset \cdots$$

has no upper bound.

(b) Let V be a vector space over the field K and set

$$I := \{ B \subseteq V \mid B \text{ is linearly independent} \}.$$

We claim that $\langle I, \subseteq \rangle$ is inductively ordered.

Let $C \subseteq I$ be a chain. We show that $\sup C = \bigcup C$. By Corollary A2.3.10, it is sufficient to prove that $\bigcup C \in I$.

Suppose otherwise. Then $\bigcup C$ is not linearly independent and there are elements $v_0, \dots, v_n \in \bigcup C$ and $\lambda_0, \dots, \lambda_n \in K$ such that $\lambda_i \neq 0$, for all i, and

$$\lambda_0 v_0 + \cdots + \lambda_n v_n = 0.$$

For each v_i, fix some $B_i \in C$ with $v_i \in B_i$. Since C is linearly ordered so is the set $\{B_0, \dots, B_n\}$. This set is finite and, therefore, it has a maximal element B_k, that is, $B_i \subseteq B_k$, for all i. It follows that $v_0, \dots, v_n \in B_k$, which implies that B_k is not linearly independent. Contradiction.

Lemma 3.12. *Let $\mathfrak{A} = \langle A, \leq \rangle$ be inductively ordered.*

(a) *If $f : A \to A$ is inflationary, f is inductive over every element $a \in A$.*

(b) *If $f : A \to A$ is increasing, f is inductive over every element a with $f(a) \geq a$.*

(c) *If $f : A \to A$ is continuous, f is inductive over every element a with $f(a) \geq a$. Furthermore, if the inductive fixed point of f over a exists, its closure ordinal is at most ω.*

Proof. (a) By transfinite recursion, we construct an increasing function $F : I \to A$ satisfying the equations in Definition 3.6. Let $F(0) := a$. For the inductive step, suppose that $F(\alpha)$ is already defined. We set $F(\alpha^+) := f(F(\alpha))$. Since f is inflationary, it follows that $F(\alpha^+) = f(F(\alpha)) \geq F(\alpha)$. Finally, suppose that δ is a limit ordinal. If $F \restriction {\downarrow}\delta$ is a proper class, we are done. Otherwise, $F[{\downarrow}\delta]$ is a set which, furthermore, is linearly ordered because $F \restriction {\downarrow}\delta$ is increasing. As $\langle A, \leq \rangle$ is inductively ordered it follows that $F[{\downarrow}\delta]$ has a supremum and we can set $F(\delta) := \sup F[{\downarrow}\delta]$.

(b) Again we define an increasing function $F : I \to A$ by transfinite recursion. Let $F(0) := a$. For the inductive step, suppose that $F(\alpha)$ is already defined. We set $F(\alpha^+) := f(F(\alpha))$. To prove that $F(\alpha^+) \geq F(\alpha)$ we consider three cases. For $\alpha = 0$ we have $F(1) = f(a) \geq a = F(0)$. If $\alpha = \beta^+$ is a successor, we know by inductive hypothesis that $F(\beta^+) \geq F(\beta)$. Since f is increasing it follows that

$$F(\alpha^+) = f(F(\beta^+)) \geq f(F(\beta)) = F(\beta^+) = F(\alpha).$$

If α is a limit then $F(\alpha) = \sup F[{\downarrow}\alpha]$ and

$$F(\alpha^+) = f(\sup F[{\downarrow}\alpha]) \geq f(F(\beta)) = F(\beta^+), \qquad \text{for all } \beta < \alpha.$$

This implies that

$$F(\alpha^+) \geq \sup F[{\downarrow}\alpha] = F(\alpha).$$

Finally, let δ be a limit ordinal. Again, if $F \restriction {\downarrow}\delta$ is a proper class, we are done. Otherwise, $F[{\downarrow}\delta]$ is a set and, as above, $F(\delta) := \sup F[{\downarrow}\delta]$ exists.

(c) Since continuous functions are increasing it follows from (b) that f is inductive over a. Let F be the corresponding fixed-point induction. It remains to show that, if $\omega \in \operatorname{dom} F$ then $F(\infty) = F(\omega)$. Since f is continuous we have

$$\begin{aligned} F(\omega^+) &= f(\sup F[{\downarrow}\omega]) \\ &= \sup \{ f(F(\alpha)) \mid \alpha < \omega \} \\ &= \sup \{ F(\alpha^+) \mid \alpha < \omega \} \\ &= \sup F[{\downarrow}\omega] = F(\omega), \end{aligned}$$

as desired. □

Lemma 3.13. *Let $f : \mathrm{On} \to \mathrm{On}$ be strictly continuous and let $\alpha \in \mathrm{On}$.*

(a) f *is inductive over* α.

(b) *If* F *is the fixed-point induction of* f *over* α *then* $F(\infty)$ *exists if, and only if, the set* $\{ f^n(\alpha) \mid n < \omega \}$ *is bounded. In this case we have* $F(\infty) = F(\omega)$.

Proof. (a) In Lemma 1.7 we have shown that every strictly continuous function on a well-order is inflationary. Therefore, Lemma 3.12 implies that f is inductive over α.

(b) We prove by induction on $n < \omega$ that $n \in \operatorname{dom} F$. By definition we have $0 \in \operatorname{dom} F$. If $n \in \operatorname{dom} F$ then $f(F(n)) \geq F(n)$ since f is inflationary. Hence, $F(n+1) = f(F(n))$ is defined. If

$$\{ f^n(\alpha) \mid n < \omega \} = F[\downarrow\omega]$$

is bounded, it follows that $F(\omega) = \sup F[\downarrow\omega]$ is defined. Consequently, Lemma 3.12 implies that $F(\infty) = F(\omega)$. □

Exercise 3.2. Let $f : \wp(A) \to \wp(A)$ be inflationary and increasing, and let $c : \wp(A) \to \wp(A)$ be the function that maps $X \subseteq A$ to the inductive fixed point of f over X. Prove that c is a closure operator.

We conclude this section with two theorems which can be used to prove the existence of fixed points. The first one is an immediate consequence of the above results.

Theorem 3.14 (Bourbaki). *Let* $\langle A, \leq \rangle$ *be an inductively ordered graph. If* A *is a set then every inflationary function* $f : A \to A$ *has an inductive fixed point.*

Proof. By Lemma 3.12, f is inductive over $\perp$. Consequently, f has an inductive fixed point, by Theorem 3.9. □

Example. The condition of A being a set is necessary. For instance, $\mathfrak{On}$ is inductively ordered since every set of ordinals has a supremum and the function $f : \mathrm{On} \to \mathrm{On} : \alpha \mapsto \alpha^+$ is inflationary. But f has no fixed point.

The second theorem is a version of the Theorem of Knaster and Tarski which shows that we can compute the least fixed point of a function f by a fixed-point induction.

Theorem 3.15. *Let* $\langle A, \leq \rangle$ *be an inductively ordered graph where* A *is a set and let* $f : A \to A$ *be an increasing function. If the least fixed point of* f *exists then it coincides with its inductive fixed point over* $\perp$.

Proof. Let $F : \mathrm{On} \to A$ be the fixed-point induction of f over $\perp$. Suppose that $a := \operatorname{lfp} f$ exists. We prove by induction on α that $F(\alpha) \leq a$. Then it follows that $F(\infty) \leq a$ and the minimality of a implies that $F(\infty) = a$.

Clearly, $F(0) = \perp \leq a$. For the inductive step, suppose that $F(\alpha) \leq a$. Since f is increasing it follows that

$$F(\alpha^+) = f(F(\alpha)) \leq f(a) = a\,.$$

Finally, if δ is a limit ordinal, the inductive hypothesis implies that

$$F(\delta) = \sup \{ F(\alpha) \mid \alpha < \delta \} \leq a\,.$$ □

4. Ordinal arithmetic

Many properties of natural numbers can be generalised to ordinals. We have already seen that ordinals allow proofs by induction. In this section we will show how to define addition, multiplication, and exponentiation for such numbers.

We start by defining these operations for arbitrary linear orders. Intuitively, the sum of two linear orders $\mathfrak{A}$ and $\mathfrak{B}$ is the order consisting of a copy of $\mathfrak{A}$ followed by a copy of $\mathfrak{B}$. Similarly, their product is obtained from $\mathfrak{B}$ by replacing every element by a copy of $\mathfrak{A}$.

Definition 4.1. Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be linear orders.

(a) The *sum* $\mathfrak{A} + \mathfrak{B}$ is the graph $\langle C, \leq_C \rangle$ where

$$C := A \uplus B = (\{0\} \times A) \cup (\{1\} \times B)$$

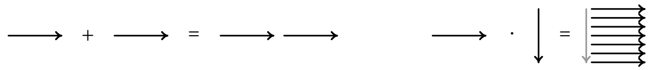


Figure 1.. Sum and product of linear orders

and the order is defined by

$$\begin{aligned}\langle i, a\rangle \leq_C \langle k, b\rangle \quad :\text{iff} \quad & i = k = 0 \text{ and } a \leq_A b \\ & \text{or } i = k = 1 \text{ and } a \leq_B b \\ & \text{or } i = 0 \text{ and } k = 1 .\end{aligned}$$

(b) The *product* $\mathfrak{A} \cdot \mathfrak{B}$ is the graph $\langle C, \leq_C\rangle$ where $C := A \times B$ and the order is defined by

$$\langle a, b\rangle \leq_C \langle a', b'\rangle \quad :\text{iff} \quad b <_B b' \text{ or } (b = b' \text{ and } a \leq_A a') .$$

(This is the reversed lexicographic ordering, see Definition B2.1.1.)

(c) If $\mathfrak{A}$ and $\mathfrak{B}$ are well-orders then we define $\mathfrak{A}^{(\mathfrak{B})} := \langle C, \leq_C\rangle$ where

$$C := \{\, f \in A^B \mid \text{there are only finitely many } b \in B \text{ with } f(b) \neq \bot \,\} ,$$

and the order is defined by

$$f <_C g \quad :\text{iff} \quad \text{the set } \{\, b \in B \mid f(b) \neq g(b) \,\} \text{ has a maximal element } b_0 \text{ and we have } f(b_0) <_A g(b_0) .$$

For natural numbers, these operations coincide with the usual ones.

Exercise 4.1. Let $\mathfrak{K} := \langle [k], \leq\rangle$ and $\mathfrak{M} := \langle [m], \leq\rangle$ where $k, m < \omega$. Prove that

(a) $\mathfrak{K} + \mathfrak{M} \cong \langle [k + m], \leq\rangle$,

(b) $\mathfrak{K} \cdot \mathfrak{M} \cong \langle [km], \leq\rangle$,

(c) $\mathfrak{K}^{(\mathfrak{M})} \cong \langle [k^m], \leq\rangle$.

Addition of linear orders is associative and the empty order is a neutral element. Below we will give an example showing that, in general, it is not commutative.

Lemma 4.2. *If $\mathfrak{A}$, $\mathfrak{B}$, and $\mathfrak{C}$ are linear orders then*

$$(\mathfrak{A} + \mathfrak{B}) + \mathfrak{C} \cong \mathfrak{A} + (\mathfrak{B} + \mathfrak{C}) .$$

Proof. Let $\mathfrak{A} = \langle A, \leq_A\rangle$, $\mathfrak{B} = \langle B, \leq_B\rangle$, and $\mathfrak{C} = \langle C, \leq_C\rangle$. We can define a bijection $f : (A \cup B) \cup C \to A \cup (B \cup C)$ by

$$\begin{aligned} f\langle 0, \langle 0, a\rangle\rangle &:= \langle 0, a\rangle && \text{for } a \in A , \\ f\langle 0, \langle 1, b\rangle\rangle &:= \langle 1, \langle 0, b\rangle\rangle && \text{for } b \in B , \\ f\langle 1, c\rangle &:= \langle 1, \langle 1, c\rangle\rangle && \text{for } c \in C . \end{aligned}$$

Since this bijection preserves the ordering it is the desired isomorphism. □

As we want to define arithmetic operations on ordinals we have to show that, if we apply the above operations to well-orders, we again obtain a well-order.

Lemma 4.3. *If $\mathfrak{A}$ and $\mathfrak{B}$ are well-orders then so are $\mathfrak{A} + \mathfrak{B}$, $\mathfrak{A} \cdot \mathfrak{B}$, and $\mathfrak{A}^{(\mathfrak{B})}$.*

Proof. Suppose that $\mathfrak{A} = \langle A, \leq_A\rangle$ and $\mathfrak{B} = \langle B, \leq_B\rangle$. We will prove the claim only for $\mathfrak{C} := \mathfrak{A}^{(\mathfrak{B})}$. The other operations are left as an exercise to the reader.

Let $\mathfrak{C} = \langle C, \leq_C\rangle$. The relation $<_C$ is irreflexive since, for each $f \in C$, the set $\{\, b \in B \mid f(b) \neq f(b) \,\}$ is empty and has no maximal element. Furthermore, $<_C$ is linear. For transitivity, let $f, g, h \in C$ be functions such that $f <_C g <_C h$. Let $b_0, b_1 \in B$ be the maximal elements such that, respectively, $f(b_0) \neq g(b_0)$ and $g(b_1) \neq h(b_1)$. By definition, we have $f(b_0) <_A g(b_0)$ and $g(b_1) <_A h(b_1)$. If $b_0 \leq_B b_1$ then

$$f(b_1) \leq g(b_1) <_A h(b_1)$$

and $\quad f(b) = g(b) = h(b)\,, \qquad$ for $b >_B b_1$,

implies that $f <_C h$. Similarly, if $b_1 <_B b_0$ then

$$f(b_0) <_A g(b_0) = h(b_0)$$

and $\quad f(b) = g(b) = h(b)\,, \qquad$ for $b >_B b_0$.

In both cases it follows that $f <_C h$. Consequently, $<_C$ is a strict linear order.

It remains to prove that every nonempty subset $X \subseteq C$ has a minimal element. We prove the claim by induction on $\beta := \operatorname{ord}(\mathfrak{B})$. If $\beta = 0$ then $C = A^{(\varnothing)} = \{\varnothing\}$ and we are done. Suppose that $\beta > 0$ and select an arbitrary element $f \in X$. If $f(b) = \perp$, for all $b \in B$, then f is the minimal element of X and we are done. Hence, we may assume that there is some $b \in B$ with $f(b) \neq \perp$. Since there are only finitely many such elements we may assume that b is the maximal one. Define

$$Y := \{\, g \in X \mid g(c) = \perp \text{ for all } c > b \,\}\,.$$

This set is nonempty since $f \in Y$. Set

$$a := \min\{\, g(b) \mid g \in Y \,\} \quad \text{and} \quad Z := \{\, g \in Y \mid g(b) = a \,\}\,.$$

By construction, we have $g <_C h$ whenever $g \in Z$ and $h \in X \setminus Z$. Consequently, if we can find a minimal element of Z, we also have the minimal element of X. Let

$$U := \{\, g \restriction {\downarrow} b \mid g \in Z \,\} \subseteq A^{({\downarrow} b)}.$$

Since $\operatorname{ord}({\downarrow} b) < \beta$ we can apply the inductive hypothesis and there exists a minimal element $h \in U$. Note that the restriction map

$$\rho : Z \to U : g \mapsto g \restriction {\downarrow} b$$

is a bijection since we have

$$g(c) = g'(c) \quad \text{for all } g, g' \in Z \text{ and every } c \geq b\,.$$

Furthermore, ρ preserves the ordering, that is, it is an isomorphism. It follows that $\rho^{-1}(h)$ is the minimal element of Z and of X. □

Exercise 4.2. Show that, if $\mathfrak{A}$ and $\mathfrak{B}$ are well-orders then so are $\mathfrak{A} + \mathfrak{B}$ and $\mathfrak{A} \cdot \mathfrak{B}$.

It is easy to see that $\mathfrak{A} \cong \mathfrak{A}'$ and $\mathfrak{B} \cong \mathfrak{B}'$ implies that the sums, products, and powers are also isomorphic. Therefore, we can define the corresponding operations on ordinals by taking representatives.

Definition 4.4. For $\alpha = \operatorname{ord}(\mathfrak{A})$ and $\beta = \operatorname{ord}(\mathfrak{B})$ we define

$$\begin{aligned} \alpha + \beta &:= \operatorname{ord}(\mathfrak{A} + \mathfrak{B})\,, \\ \alpha \cdot \beta &:= \operatorname{ord}(\mathfrak{A} \cdot \mathfrak{B})\,, \\ \alpha^{(\beta)} &:= \operatorname{ord}\big(\mathfrak{A}^{(\mathfrak{B})}\big)\,. \end{aligned}$$

Example. The following equations can be proved easily by the lemmas below. We encourage the reader to derive them directly from the definitions.

$$\begin{aligned} &1 + 1 = 2 && (3 + 6)\omega = 9\omega = \omega < \omega 2 = 3\omega + 6\omega \\ &\omega + \omega = \omega 2 && (\omega 6 + 17)\omega = \omega\omega = \omega^{(2)} \\ &1 + \omega = \omega < \omega + 1 && 2^{(\omega)} = \omega \\ &2\omega = \omega < \omega 2 && \end{aligned}$$

Exercise 4.3. Show that $\alpha + \beta$, $\alpha \cdot \beta$, and $\alpha^{(\beta)}$ are well-defined, for all $\alpha, \beta \in \mathrm{On}$.

Exercise 4.4. Show that $\alpha^+ = \alpha + 1$.

Ordinal addition

The properties of ordinal addition, multiplication, and exponentiation are similar to, but not quite the same as those for integers. The following sequence of lemmas summarises them. We start with addition.

Lemma 4.5. *Let $\alpha, \beta, \gamma \in \mathrm{On}$. If $\beta < \gamma$ then $\alpha + \beta < \alpha + \gamma$.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$, and $\gamma = \mathrm{ord}(\mathfrak{C})$. There exists an isomorphism $f : B \to I \subset C$ between B and some proper initial segment I of C. We define an isomorphism $g : A \uplus B \to A \uplus I$ by

$$g(\langle 0, a\rangle) := \langle 0, a\rangle\,, \qquad \text{for } a \in A\,,$$
$$\text{and} \quad g(\langle 1, b\rangle) := \langle 1, f(b)\rangle\,, \qquad \text{for } b \in B\,.$$

Hence, $\mathfrak{A} + \mathfrak{B} < \mathfrak{A} + \mathfrak{C}$. □

In the last section we gave an inductive definition of addition. The next lemma shows that it is equivalent to the official definition above.

Lemma 4.6. *Let $\alpha, \beta \in \mathrm{On}$.*

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

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

(c) $\alpha + \delta = \sup\{\, \alpha + \beta \mid \beta < \delta \,\}$, *for limit ordinals δ.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$ and $\beta = \mathrm{ord}(\mathfrak{B})$.

(a) follows immediately since $\mathfrak{A} + \langle \varnothing, \leq\rangle \cong \mathfrak{A}$.

(b) By Lemma 4.2, we have

$$(\mathfrak{A} + \mathfrak{B}) + \mathfrak{C} \cong \mathfrak{A} + (\mathfrak{B} + \mathfrak{C})\,, \quad \text{for all linear orders } \mathfrak{A}, \mathfrak{B}, \mathfrak{C}\,.$$

Since $\beta^+ = \mathrm{ord}(\mathfrak{B} + \langle [1], \leq\rangle)$ the result follows.

(c) Let $X := \{\, \alpha + \beta \mid \beta < \delta \,\}$ and set $\gamma := \sup X$. By Lemma 4.5, we have $\alpha + \beta < \alpha + \delta$, for all $\beta < \delta$, which implies that $\gamma \leq \alpha + \delta$.

For a contradiction suppose that $\gamma < \alpha + \delta$. Fix representatives $\gamma = \mathrm{ord}(\mathfrak{C})$ and $\delta = \mathrm{ord}(\mathfrak{D})$. Since $\alpha + 0 < \gamma < \alpha + \delta$ there exists an isomorphism $f : C \to A \uplus I$, for some proper initial segment $\varnothing \subset I \subset D$. Let $C_0 := f^{-1}[A]$ and $C_1 := f^{-1}[I]$. Since f is an isomorphism we have

$$\mathfrak{A} \cong \langle C_0, \leq\rangle \quad \text{and} \quad \mathfrak{C} \cong \langle C_0, \leq\rangle + \langle C_1, \leq\rangle\,.$$

Set $\beta := \mathrm{ord}(\langle C_1, \leq\rangle)$. It follows that $\gamma = \alpha + \beta$. Furthermore, because of the inclusion map $I \to D$ we have $\beta < \delta$. By (b) it follows that

$$\gamma < (\alpha + \beta)^+ = \alpha + \beta^+ \leq \sup X\,.$$

Contradiction. □

Corollary 4.7. *The function $f_\alpha : \mathrm{On} \to \mathrm{On}$ with $f_\alpha(\beta) := \alpha + \beta$ is strictly continuous, for every $\alpha \in \mathrm{On}$.*

Proof. The claim follows immediately from the preceding lemma and Lemma 1.13. □

Since ordinal addition is not commutative there are two possible ways to subtract ordinals. Given $\alpha \geq \beta$ we can ask for some ordinal γ such that $\alpha = \beta + \gamma$, or we can ask for some γ with $\alpha = \gamma + \beta$. The next lemma shows that the first operation is well-defined. The second one is not since, for example, $1 + \omega = \omega = 2 + \omega$.

Lemma 4.8. *For all ordinals $\beta \leq \alpha$, there exists a unique ordinal γ such that $\alpha = \beta + \gamma$.*

Proof. By Corollary 4.7 and Lemma 1.14, there exists a greatest ordinal γ such that $\beta + \gamma \leq \alpha$. If $\beta + \gamma < \alpha$ then we would have

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

in contradiction to the choice of γ. Hence, $\beta + \gamma = \alpha$. The uniqueness of γ follows from the fact that the function $\gamma \mapsto \beta + \gamma$ is injective. □

The next lemma summarises the laws of ordinal addition.

Lemma 4.9. *Let $\alpha, \beta, \gamma \in \mathrm{On}$.*

(a) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.

(b) $\alpha + \beta = \alpha + \gamma$ *implies* $\beta = \gamma$.

(c) $\alpha \leq \beta$ *implies* $\alpha + \gamma \leq \beta + \gamma$.

(d) *If* $X \subseteq \mathrm{On}$ *is nonempty and bounded then*

$$\alpha + \sup X = \sup \{ \alpha + \beta \mid \beta \in X \} .$$

(e) $\beta \leq \alpha$ *if, and only if,* $\alpha = \beta + \gamma$, *for some* $\gamma \in \mathrm{On}$.

(f) $\beta < \alpha$ *if, and only if,* $\alpha = \beta + \gamma$, *for some* $\gamma \in \mathrm{On} \setminus \{0\}$.

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$ and $\gamma = \mathrm{ord}(\mathfrak{C})$.

(a) follows from Lemma 4.2; (b) follows from Lemma 4.8; and (d) follows from Corollary 4.7.

(c) We prove the claim by induction on γ. For $\gamma = 0$, we have

$$\alpha + 0 = \alpha \leq \beta = \beta + 0 .$$

For the successor step, note that $\alpha \leq \beta$ implies $\alpha^+ \leq \beta^+$. Hence, it follows that

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

It remains to consider the limit step. For every $\eta < \gamma$, the inductive hypothesis yields

$$\alpha + \eta \leq \beta + \eta < \beta + \gamma .$$

Therefore, Lemma 4.6 (c) implies that

$$\alpha + \gamma = \sup \{ \alpha + \eta \mid \eta < \gamma \} \leq \beta + \gamma .$$

(e) If $\beta < \alpha$, we obtain by Lemma 4.8 some $\gamma \in \mathrm{On}$ with $\alpha = \beta + \gamma$. Conversely, if $\beta + \gamma = \alpha$ then there exists an isomorphism

$$f : B \cup C \to A .$$

We can define an isomorphism $g : B \to I \subseteq A$ by

$$g(b) := f(\langle 0, b \rangle) .$$

This implies that $\mathfrak{B} \leq \mathfrak{A}$.

(f) follows immediately from (e). □

Ordinal multiplication

After addition we turn to ordinal multiplication. The development is analogous to the one above. First, we show that the function $\beta \mapsto \alpha\beta$ is strictly increasing.

Lemma 4.10. *Let $\alpha, \beta, \gamma \in \mathrm{On}$. If $\alpha \neq 0$ and $\beta < \gamma$ then $\alpha\beta < \alpha\gamma$.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$, and $\gamma = \mathrm{ord}(\mathfrak{C})$. By assumption, there exists an isomorphism $f : B \to I \subset C$ between B and a proper initial segment of C. We can define an isomorphism $g : A \times B \to A \times I$ by

$$g(\langle a, b \rangle) := \langle a, f(b) \rangle .$$

Since $A \times I$ is a proper initial segment of $A \times C$ it follows that $\alpha\beta < \alpha\gamma$. □

Again the inductive definition coincides with the official one.

Lemma 4.11. *Let $\alpha, \beta \in \mathrm{On}$.*

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

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

(c) $\alpha\delta = \sup \{ \alpha\beta \mid \beta < \delta \}$, *for limit ordinals* δ.

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$ and $\beta = \mathrm{ord}(\mathfrak{B})$.

(a) follows immediately from the fact that $\mathfrak{A} \cdot \langle \varnothing, \varnothing \rangle = \langle \varnothing, \varnothing \rangle$.

(b) The canonical bijection

$$A \times (B \cup [1]) \to (A \times B) \cup A$$

given by

$$\begin{aligned} \langle a, \langle 0, b \rangle \rangle &\mapsto \langle 0, \langle a, b \rangle \rangle , \\ \langle a, \langle 1, 0 \rangle \rangle &\mapsto \langle 1, a \rangle , \end{aligned}$$

induces an isomorphism

$$\mathfrak{A} \cdot (\mathfrak{B} + \langle [1], \leq \rangle) \to \mathfrak{A} \cdot \mathfrak{B} + \mathfrak{A} .$$

(c) Let $X := \{ \alpha\beta \mid \beta < \delta \}$ and set $\gamma := \sup X$. By Lemma 4.10, we have $\alpha\beta < \alpha\delta$, for all $\beta < \delta$. Hence, $\gamma = \sup X \leq \alpha\delta$.

For a contradiction suppose that $\gamma < \alpha\delta$. Fix representatives $\gamma = \mathrm{ord}(\mathfrak{C})$ and $\delta = \mathrm{ord}(\mathfrak{D})$. Since $\gamma < \alpha\delta$ there exists an isomorphism $f : C \to I$, for some proper initial segment $\varnothing \subset I \subset A \times D$. Let $\langle a, d \rangle$ be the minimal element of $A \times D \setminus I$. Then $I = (A \times {\downarrow} d) \cup ({\downarrow} a \times \{d\})$, which implies that

$$\gamma = \alpha \cdot \mathrm{ord}({\downarrow} d) + \mathrm{ord}({\downarrow} a) .$$

Since $\mathrm{ord}({\downarrow} a) < \alpha$ and $\beta := \mathrm{ord}({\downarrow} d) < \delta$ it follows that

$$\gamma < \alpha\beta + \alpha = \alpha\beta^+ \leq \sup X .$$

Contradiction. □

Corollary 4.12. *The function* $f_\alpha : \mathrm{On} \to \mathrm{On}$ *with* $f_\alpha(\beta) := \alpha\beta$ *is strictly continuous, for every* $\alpha > 0$.

Proof. The claim follows immediately from the preceding lemma and Lemma 1.13. □

We can also show that ordinals allow a limited form of division.

Lemma 4.13. *For all ordinals* $\alpha, \beta \in \mathrm{On}$ *with* $\beta \neq 0$, *there exist unique ordinals* γ *and* $\rho < \beta$ *such that* $\alpha = \beta\gamma + \rho$.

Proof. By Corollary 4.12 and Lemma 1.14, there exists a greatest ordinal γ such that $\beta\gamma \leq \alpha$, and, by Lemma 4.8, there exists some ordinal ρ such that $\beta\gamma + \rho = \alpha$. By choice of γ, we have

$$\beta\gamma + \beta = \beta(\gamma + 1) > \alpha = \beta\gamma + \rho ,$$

which implies that $\rho < \beta$.

Suppose there exist ordinals $\delta \neq \gamma$ and $\sigma < \beta$ such that $\beta\delta + \sigma = \alpha$. Since $\beta\delta \leq \alpha$ we have $\delta < \gamma$, which implies that

$$\alpha = \beta\gamma + \rho \geq \beta\delta^+ = \beta\delta + \beta > \beta\delta + \sigma = \alpha .$$

A contradiction. It follows that γ is unique. Hence, the uniqueness of ρ follows from Lemma 4.8. □

Lemma 4.14. α *is a limit ordinal if, and only if,* $\alpha = \omega\beta$, *for some* $\beta > 0$.

Proof. ($\Rightarrow$) By Lemma 4.13, we have $\alpha = \omega\beta + n$ for some $\beta \in \mathrm{On}$ and $n < \omega$. Suppose that $n \neq 0$. Then $n = m + 1$, for some $m < \omega$, and

$$\alpha = \omega\beta + (m + 1) = (\omega\beta + m) + 1 .$$

Consequently, α is a successor ordinal. Contradiction.

($\Leftarrow$) Suppose that $\omega\beta$ is a successor ordinal. That is, $\omega\beta = \gamma + 1$, for some γ. By Lemma 4.13, we can write γ as $\gamma = \omega\eta + n$, for some $n < \omega$. Hence,

$$\omega\beta = \gamma + 1 = \omega\eta + (n + 1) .$$

By Lemma 4.13, it follows that $\beta = \eta$ and $0 = n + 1$. Contradiction. □

The laws of ordinal multiplication are summarised in the following lemma.

Lemma 4.15. *Let* $\alpha, \beta, \gamma \in \mathrm{On}$.

(a) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$.

(b) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

(c) *If* $\alpha \neq 0$ *and* $\alpha\beta = \alpha\gamma$ *then* $\beta = \gamma$.

(d) $\alpha \leq \beta$ *implies* $\alpha\gamma \leq \beta\gamma$.

(e) *If* $X \subseteq \mathrm{On}$ *is nonempty and bounded then*

$$\alpha \cdot \sup X = \sup \{ \alpha\beta \mid \beta \in X \}.$$

Proof. (b) We prove the claim by induction on γ. For $\gamma = 0$, we have

$$\alpha(\beta + 0) = \alpha\beta = \alpha\beta + 0 = \alpha\beta + \alpha 0.$$

For the successor step, we have

$$\begin{aligned} \alpha(\beta + \gamma^+) &= \alpha(\beta + \gamma)^+ \\ &= \alpha(\beta + \gamma) + \alpha \\ &= \alpha\beta + \alpha\gamma + \alpha \\ &= \alpha\beta + \alpha\gamma^+. \end{aligned}$$

Finally, if γ is a limit ordinal then

$$\begin{aligned} \alpha(\beta + \gamma) &= \alpha \cdot \sup \{ \beta + \rho \mid \rho < \gamma \} \\ &= \sup \{ \alpha(\beta + \rho) \mid \rho < \gamma \} \\ &= \sup \{ \alpha\beta + \alpha\rho \mid \rho < \gamma \} \\ &= \alpha\beta + \sup \{ \alpha\rho \mid \rho < \gamma \} \\ &= \alpha\beta + \alpha\gamma. \end{aligned}$$

(a) and (d) can also be proved by induction on γ. We leave the details as an exercise to the reader.

(c) and (e) follow immediately from Corollary 4.12. □

Ordinal exponentiation

Finally, we consider ordinal exponentiation. Again, the basic steps are the same as for addition and multiplication.

Lemma 4.16. *Let* $\alpha, \beta, \gamma \in \mathrm{On}$. *If* $\alpha > 1$ *and* $\beta < \gamma$ *then* $\alpha^{(\beta)} < \alpha^{(\gamma)}$.

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$, and $\gamma = \mathrm{ord}(\mathfrak{C})$. There exists an isomorphism $f : B \to I \subset C$ between B and a proper initial segment I of C. The desired isomorphism

$$A^{(B)} \to A^{(I)} \subset A^{(C)}$$

is given by the mapping $g \mapsto g \circ f^{-1}$. □

Ordinal exponentiation can also be defined inductively.

Lemma 4.17. *Let* $\alpha, \beta \in \mathrm{On}$.

(a) $\alpha^{(0)} = 1$.

(b) $\alpha^{(\beta^+)} = \alpha^{(\beta)}\alpha$.

(c) $\alpha^{(\delta)} = \sup \{ \alpha^{(\beta)} \mid \beta < \delta \}$, *for limit ordinals* δ.

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$ and $\beta = \mathrm{ord}(\mathfrak{B})$.

(a) Since $\varnothing$ is the only function with empty domain we have $A^{(\varnothing)} = A^{\varnothing} = \{\varnothing\}$.

(b) There is a canonical bijection $A^{(B \cup [1])} \to A^{(B)} \times A$ given by

$$f \mapsto \langle f', f(\langle 1, 0 \rangle) \rangle$$

where the function $f' : B \to A$ is defined by $f'(b) := f(\langle 0, b \rangle)$. This bijection induces the desired isomorphism

$$\mathfrak{A}^{(\mathfrak{B} + \langle [1], \leq \rangle)} \to \mathfrak{A}^{(\mathfrak{B})} \cdot \mathfrak{A}.$$

(c) If $\alpha < 2$, the claim is trivial. Hence, we may assume that $\alpha > 1$. Let $X := \{ \alpha^{(\beta)} \mid \beta < \delta \}$ and set $\gamma := \sup X$. By Lemma 4.16, we have $\alpha^{(\beta)} < \alpha^{(\delta)}$, for all $\beta < \delta$. Hence, $\gamma = \sup X \leq \alpha^{(\delta)}$.

For a contradiction suppose that $\gamma < \alpha^{(\delta)}$. Fix representatives $\gamma = \mathrm{ord}(\mathfrak{C})$ and $\delta = \mathrm{ord}(\mathfrak{D})$. Since $\gamma < \alpha^{(\delta)}$, there exists an isomorphism $f : C \to I$, for some proper initial segment $I \subset A^{(D)}$. Let g be the minimal element of $A^{(D)} \setminus I$ and let $d_0 < \dots < d_n$ be the enumeration of the set $\{ d \in D \mid g(d) \neq 0 \}$. We can decompose I as $I = I_n \cup \dots \cup I_0$ where, for each $i \leq n$,

$$I_i := \{ h \in A^D \mid h(d_i) < g(d_i) \text{ and } h(x) = g(x)\,, \text{ for } x > d_i \}\,.$$

Set $\beta_i := \mathrm{ord}(\downarrow d_i) < \delta$ and $\eta_i := \mathrm{ord}(\downarrow g(d_i))$. It follows that

$$\begin{aligned}\gamma &= \alpha^{(\beta_n)} \cdot \eta_n + \dots + \alpha^{(\beta_0)} \cdot \eta_0 \\ &< \alpha^{(\beta_n)}\alpha + \dots + \alpha^{(\beta_0)}\alpha \\ &\leq \alpha^{(\beta_n)}\alpha + \dots + \alpha^{(\beta_n)}\alpha \\ &= \alpha^{(\beta_n+1)}(n+1)\,.\end{aligned}$$

Since $\alpha > 1$ there is some finite ordinal m such that $\alpha^{(m)} \geq n+1$. Therefore, it follows by (b) that

$$\gamma < \alpha^{(\beta_n+1)}\alpha^{(m)} = \alpha^{(\beta_n+m+1)} \leq \sup X\,.$$

Contradiction. □

Corollary 4.18. *The function $f_\alpha : \mathrm{On} \to \mathrm{On}$ with $f_\alpha(\beta) := \alpha^{(\beta)}$ is strictly continuous, for every $\alpha > 1$.*

Proof. The claim follows immediately from the preceding lemma and Lemma 1.13. □

Besides subtraction and division we can also take a limited form of logarithms.

Lemma 4.19. *For all ordinals $\alpha, \beta \in \mathrm{On}$ with $\alpha > 0$ and $\beta > 1$, there exist unique ordinals γ, η, and ρ with $0 < \gamma < \beta$ and $\rho < \beta^{(\eta)}$ such that $\alpha = \beta^{(\eta)}\gamma + \rho$.*

Proof. By Corollary 4.18 and Lemma 1.14, there exists a greatest ordinal η such that $\beta^{(\eta)} \leq \alpha$, and, by Lemma 4.13, there exist ordinals γ and $\rho < \beta^{(\eta)}$ such that $\beta^{(\eta)}\gamma + \rho = \alpha$. If $\gamma = 0$, we would have $\rho = \alpha \geq \beta^{(\eta)} > \rho$. A contradiction. And, if $\gamma \geq \beta$, we would have

$$\alpha < \beta^{(\eta+1)} = \beta^{(\eta)}\beta \leq \beta^{(\eta)}\gamma \leq \beta^{(\eta)}\gamma + \rho = \alpha\,.$$

Again a contradiction. Therefore, $0 < \gamma < \beta$.

Suppose there exist ordinals $\mu \neq \eta$, δ, and σ such that $\beta^{(\mu)}\delta + \sigma = \alpha$. Since $\beta^{(\mu)} \leq \alpha$ we have $\mu < \eta$, which implies that

$$\begin{aligned}\alpha = \beta^{(\eta)}\gamma + \rho \geq \beta^{(\mu^+)} &= \beta^{(\mu)}\beta \geq \beta^{(\mu)}(\delta + 1) = \beta^{(\mu)}\delta + \beta^{(\mu)} \\ &> \beta^{(\mu)}\delta + \sigma = \alpha\,.\end{aligned}$$

A contradiction. It follows that η is unique. Hence, the uniqueness of γ and ρ follows from Lemma 4.8. □

Let us summarise the laws of ordinal exponentiation.

Lemma 4.20. *Let $\alpha, \beta, \gamma \in \mathrm{On}$.*

(a) $\alpha^{(\beta+\gamma)} = \alpha^{(\beta)}\alpha^{(\gamma)}$.
(b) $\alpha^{(\beta\gamma)} = (\alpha^{(\beta)})^{(\gamma)}$.
(c) $\alpha > 1$ *implies* $\beta \leq \alpha^{(\beta)}$.
(d) *If* $\alpha > 1$ *and* $\alpha^{(\beta)} = \alpha^{(\gamma)}$ *then* $\beta = \gamma$.
(e) $\alpha \leq \beta$ *implies* $\alpha^{(\gamma)} \leq \beta^{(\gamma)}$.
(f) *If* $\alpha > 1$ *then we have* $\beta < \gamma$ *if, and only if,* $\alpha^{(\beta)} < \alpha^{(\gamma)}$.
(g) *If* $X \subseteq \mathrm{On}$ *is nonempty and bounded then we have*

$$\alpha^{(\sup X)} = \sup\{ \alpha^{(\beta)} \mid \beta \in X \}\,.$$

Proof. (a), (b) and (e) can be proved by a simple induction on γ. (c) follows from Lemma 1.7, while (d), (f) and (g) are immediate consequences of Corollary 4.18. □

Cantor normal form

We can apply the logarithm to decompose every ordinal in a canonical way.

Theorem 4.21. *For all ordinals $\alpha, \beta \in \mathrm{On}$ with $\beta > 1$, there are unique finite sequences $(\gamma_i)_{i<n}$ and $(\eta_i)_{i<n}$ of ordinal numbers such that*

$$\alpha = \beta^{(\eta_0)}\gamma_0 + \dots + \beta^{(\eta_{n-1})}\gamma_{n-1}\,,$$
$$\eta_0 > \dots > \eta_{n-1}\,, \quad \text{and} \quad 0 < \gamma_i < \beta\,, \quad \text{for } i < n\,.$$

Proof. We decompose α successively with the help of Lemma 4.19. We start by writing $\alpha = \beta^{(\eta_0)}\gamma_0 + \rho_0$. Applying the lemma to ρ_0 we get $\rho_0 = \beta^{(\eta_1)}\gamma_1 + \rho_1$. By induction on i, we obtain $\rho_i = \beta^{(\eta_{i+1})}\gamma_{i+1} + \rho_{i+1}$. If this process did not terminate then we would get an infinite decreasing sequence $\alpha > \rho_0 > \rho_1 > \dots$ of ordinals which is impossible. Consequently, there is some number n such that $\rho_n = 0$ and we have

$$\alpha = \beta^{(\eta_0)}\gamma_0 + \dots + \beta^{(\eta_{n-1})}\gamma_{n-1}\,. \qquad \square$$

Definition 4.22. Let α be an ordinal. The unique decomposition

$$\alpha = \omega^{(\eta_0)}\gamma_0 + \dots + \omega^{(\eta_n)}\gamma_n\,,$$
$$\text{with} \quad \eta_0 > \dots > \eta_n \quad \text{and} \quad 0 < \gamma_i < \omega\,, \quad \text{for } i \leq n\,.$$

is called the *Cantor normal form* of α.

The Cantor normal form is very convenient for ordinal calculations. Let us see how this is done. We start with addition.

Lemma 4.23. *$\alpha < \beta$ implies $\omega^{(\alpha)} + \omega^{(\beta)} = \omega^{(\beta)}$.*

Proof. Suppose that $\beta = \alpha + \gamma$, for $\gamma > 0$. We have

$$\begin{aligned}
\omega^{(\alpha)} + \omega^{(\beta)} &= \omega^{(\alpha)} + \omega^{(\alpha+\gamma)} \\
&= \omega^{(\alpha)} + \omega^{(\alpha)}\omega^{(\gamma)} \\
&= \omega^{(\alpha)}(1 + \omega^{(\gamma)}) \\
&= \omega^{(\alpha)}\omega^{(\gamma)} \\
&= \omega^{(\alpha+\gamma)} = \omega^{(\beta)}.
\end{aligned} \qquad \square$$

Corollary 4.24. *Let $\alpha, \beta \in \mathrm{On}$ be ordinals with Cantor normal form*

$$\alpha = \omega^{(\eta_0)}k_0 + \dots + \omega^{(\eta_{m-1})}k_{m-1}\,,$$
$$\beta = \omega^{(\gamma_0)}l_0 + \dots + \omega^{(\gamma_{n-1})}l_{n-1}\,.$$

If i is the maximal index such that $\eta_i \geq \gamma_0$ then we have

$$\alpha + \beta = \omega^{(\eta_0)}k_0 + \dots + \omega^{(\eta_i)}k_i + \omega^{(\gamma_0)}l_0 + \dots + \omega^{(\gamma_{n-1})}l_{n-1}\,.$$

Lemma 4.25. *An ordinal $\alpha > 0$ is of the form $\alpha = \omega^{(\eta)}$, for some η, if, and only if, $\beta + \gamma < \alpha$, for all $\beta, \gamma < \alpha$.*

Proof. ($\Rightarrow$) Let

$$\beta = \omega^{(\rho_m)}k_m + \dots + \omega^{(\rho_0)}k_0 \quad \text{and} \quad \gamma = \omega^{(\sigma_n)}l_n + \dots + \omega^{(\sigma_0)}l_0$$

be the Cantor normal forms of β and γ. If $\beta, \gamma < \omega^{(\eta)}$ then $\rho_m, \sigma_n < \eta$. By symmetry, we may assume that $\gamma \leq \beta$. Thus,

$$\begin{aligned}
\beta + \gamma &\leq \beta + \beta \\
&= \omega^{(\rho_n)}(k_m + k_m) + \omega^{(\rho_{m-1})}k_{m-1} + \dots + \omega^{(\rho_0)}k_0 \\
&< \omega^{(\eta)}.
\end{aligned}$$

($\Leftarrow$) Suppose that $\alpha = \omega^{(\eta)}k + \rho$ where $k < \omega$ and $\rho < \omega^{(\eta)}$. We have to show that $k = 1$ and $\rho = 0$.

If $k > 1$, we set $\beta := \omega^{(\eta)}(k-1) + \rho < \alpha$. It follows that

$$\beta + \beta = \omega^{(\eta)}(k + (k-2)) + \rho \geq \omega^{(\eta)}k + \rho = \alpha\,.$$

Contradiction.

Suppose that $k = 1$ but $\rho > 0$. In this case we can set $\beta := \omega^{(\eta)}$ and we have

$$\beta + \beta = \omega^{(\eta)} + \omega^{(\eta)} > \omega^{(\eta)} + \rho = \alpha\,.$$

Again a contradiction. □

The next two lemmas provide the laws of multiplication and exponentiation of ordinals in Cantor normal form.

Lemma 4.26. *If $\gamma > 0$, $0 \leq \rho < \omega^{(\eta)}$, and $0 < k < \omega$ then*

$$(\omega^{(\eta)}k + \rho)\omega^{(\gamma)} = \omega^{(\eta+\gamma)}.$$

Proof. We have

$$\begin{aligned} \omega^{(\eta)}\omega^{(\gamma)} &\leq \big(\omega^{(\eta)}k + \rho\big)\omega^{(\gamma)} \\ &\leq \big(\omega^{(\eta)}(k+1)\big)\omega^{(\gamma)} \\ &= \omega^{(\eta)}\big((k+1)\omega^{(\gamma)}\big) = \omega^{(\eta)}\omega^{(\gamma)}. \end{aligned}$$ □

Lemma 4.27. *If $\gamma, \eta > 0$, $0 \leq \rho < \omega^{(\eta)}$, and $0 < k < \omega$ then*

$$(\omega^{(\eta)}k + \rho)^{(\omega^{(\gamma)})} = \omega^{(\eta\omega^{(\gamma)})}.$$

Proof. We have

$$\begin{aligned} \omega^{(\eta\omega^{(\gamma)})} &= (\omega^{(\eta)})^{(\omega^{(\gamma)})} \\ &\leq (\omega^{(\eta)}k + \rho)^{(\omega^{(\gamma)})} \\ &\leq (\omega^{(\eta+1)})^{(\omega^{(\gamma)})} \\ &= \omega^{((\eta+1)\omega^{(\gamma)})} = \omega^{(\eta\omega^{(\gamma)})}. \end{aligned}$$ □

Example. By the above lemmas we have

$$\begin{aligned} &(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)})^{(\omega^{(2)}2+\omega+1)} \\ &= \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big)^{(\omega^{(2)}2)} \cdot \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big)^{(\omega)} \cdot \\ &\quad \cdot \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big) \\ &= \big(\omega^{((\omega^{(5)}+\omega 4+2)\omega^{(2)})}\big)^{(2)} \cdot \omega^{((\omega^{(5)}+\omega 4+2)\omega)} \cdot \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big) \\ &= \big(\omega^{(\omega^{(7)})}\big)^{(2)} \cdot \omega^{(\omega^{(6)})} \cdot \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big) \\ &= \omega^{(\omega^{(7)}2)} \cdot \omega^{(\omega^{(6)})} \cdot \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big) \\ &= \omega^{(\omega^{(7)}2+\omega^{(6)})} \cdot \big(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\big) \\ &= \omega^{(\omega^{(7)}2+\omega^{(6)})} \cdot \omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(\omega^{(7)}2+\omega^{(6)})} \cdot \omega^{(5)} \\ &= \omega^{(\omega^{(7)}2+\omega^{(6)}+\omega^{(5)}+\omega 4+2)} + \omega^{(\omega^{(7)}2+\omega^{(6)}+5)}. \end{aligned}$$

Exercise 4.5. Compute the cantor normal form of

$$\big(\omega^{(\omega^{(2)}7+\omega 3+4)}3 + \omega^{(\omega 6+3)}4 + \omega^{(4)}3 + 1\big)^{(\omega^{(2)}5+\omega 7+2)}$$

Remark. We will prove in Lemma A4.5.6 that we can find, for every β, arbitrarily large ordinals $\alpha_0, \alpha_1, \alpha_2$ such that

$$\alpha_0 = \beta + \alpha_0\,, \quad \alpha_1 = \beta\alpha_1\,, \quad \text{and} \quad \alpha_2 = \beta^{(\alpha_2)}.$$

In particular, there are ordinals ε such that $\varepsilon = \omega^{(\varepsilon)}$. By ε_α we denote the α-th ordinal such that $\beta^{(\varepsilon_\alpha)} = \varepsilon_\alpha$, for all $\beta < \varepsilon_\alpha$. Note that the Cantor normal form of ε_α is $\varepsilon_\alpha = \omega^{(\varepsilon_\alpha)}$.

Let us summarise the picture of On that we have obtained. The first ordinals are

$$\begin{aligned}
&0, 1, 2, 3, \dots \\
&\dots, \omega, \omega + 1, \omega + 2, \dots \\
&\dots, \omega 2, \omega 2 + 1, \omega 2 + 2, \dots \\
&\dots, \omega 3, \dots, \omega 4, \dots, \omega^{(2)}, \dots, \omega^{(3)}, \dots \\
&\dots, \omega^{(\omega)}, \dots, \omega^{(\omega^{(\omega)})}, \dots \\
&\dots, \varepsilon_0, \dots, \varepsilon_0^{(\varepsilon_0)}, \dots, \varepsilon_1, \dots, \varepsilon_2, \dots, \varepsilon_\omega, \dots \\
&\dots, \omega_1, \dots, \omega_2, \dots, \omega_\omega, \dots
\end{aligned}$$

The ordinals ω_α will be defined in Section A4.2.

A4. *Zermelo-Fraenkel set theory*

1. *The Axiom of Choice*

We have seen that induction is a powerful technique to prove statements and to construct objects. But in order to use this tool we have to relate the sets we are interested in to ordinals. In basic set theory this is not always possible. Therefore, we will introduce a new axiom which states that, for every set A, there is a well-order over A. Before doing so, let us present several statements that are equivalent to this axiom. We need two new notions.

Definition 1.1. A set $F \subseteq \wp(A)$ has *finite character* if, for all sets $x \subseteq A$, we have

$$x \in F \quad \text{iff} \quad x_0 \in F\,, \text{ for every finite set } x_0 \subseteq x\,.$$

Lemma 1.2. *Suppose that $F \subseteq \wp(A)$ has finite character.*

(a) *F is an initial segment of $\wp(A)$.*

(b) *If $X \subseteq F$ is nonempty then $\bigcap X \in F$.*

(c) *If $C \subseteq F$ is a chain and $\bigcup C$ is a set then $\bigcup C \in F$.*

Proof. (a) follows immediately from the definition and (b) is a consequence of (a). For (c), let $C \subseteq F$ be a chain such that $X := \bigcup C$ is a set. If $X_0 \subseteq X$ is finite, there exists some element $Z \in C$ with $X_0 \subseteq Z \in F$. Hence, $X_0 \in F$, for all finite subsets $X_0 \subseteq X$. This implies that $X \in F$. □

Lemma 1.3. *If F has finite character then $\langle F, \subseteq \rangle$ is inductively ordered.*

Proof. Let $C \subseteq F$ be a linearly ordered subset of F. By Corollary A2.3.10 and Lemma 1.2 (c), it follows that $\sup C = \bigcup C \in F$. □

Example. Let V be a vector space over the field K. The set

$$F := \{ B \subseteq V \mid B \text{ is linearly independent} \}$$

has finite character.

The second notion we need is that of a choice function. Intuitively, a choice function is a function that, given some set A, selects an element of A.

Definition 1.4. A function f is a *choice function* if $f(a) \in a$, for all $a \in \operatorname{dom} f$.

Exercise 1.1. Let $\mathcal{I}$ be the set of all open intervals (a, b) of real numbers $a, b \in \mathbb{R}$ with $a < b$. Define a choice function $\mathcal{I} \to \mathbb{R}$.

Lemma 1.5. *Let A be a set and C the set of all choice functions f with* $\operatorname{dom} f \subseteq \wp(A)$.

(a) *C has finite character.*

(b) *If f is a $\subseteq$-maximal element of C then* $\operatorname{dom} f = \wp(A) \setminus \{\emptyset\}$.

Proof. (a) Suppose that f is a binary relation such that every finite $f_0 \subseteq f$ is a choice function. If $\langle a, b\rangle, \langle a, c\rangle \in f$ then $\{\langle a, b\rangle, \langle a, c\rangle\} \in C$ implies that $b = c$. Hence, f is a partial function. Furthermore, if $\langle a, b\rangle \in f$ then $\{\langle a, b\rangle\} \in C$ implies that $b \in a$. Consequently, f is a choice function.

(b) Let $f \in C$ be $\subseteq$-maximal. Since f is a choice function we have $\emptyset \notin \operatorname{dom} f$. Therefore, $\operatorname{dom} f \subseteq \wp(A) \setminus \{\emptyset\}$. Suppose that there is some element $B \in (\wp(A) \setminus \{\emptyset\}) \setminus \operatorname{dom} f$. Since $B \neq \emptyset$ we can choose some element $b \in B$. The relation $f \cup \{\langle B, b\rangle\} \supset f$ is again a choice function in contradiction to the maximality of f. □

Lemma 1.6. *Let A be a set. Given a choice function $f : \wp(A) \setminus \{\emptyset\} \to A$ we can define a well-order R on A.*

Proof. Let $f : \wp(A) \setminus \{\emptyset\} \to A$ be a choice function. We define a function $g : \wp(A) \to \wp(A)$ by

$$g(X) := \begin{cases} A & \text{if } X = A\,, \\ X \cup \{f(A \setminus X)\} & \text{if } X \neq A\,. \end{cases}$$

Since $g(X) \supseteq X$ this function is inflationary. Furthermore, the partial order $\langle \wp(A), \subseteq \rangle$ is complete. By Theorem A3.3.14, g has an inductive fixed point. Since $g(X) \neq X$, for $X \neq A$, it follows that this fixed point is A. Let $G : \mathrm{On} \to \wp(A)$ be the fixed-point induction of g over $\emptyset$ and let α be the closure ordinal. For every $\beta < \alpha$, there exists a unique element a_β such that $G(\beta + 1) \setminus G(\beta) = \{a_\beta\}$. We define a function, $h : {\downarrow}\alpha \to A$ by $h(\beta) := a_\beta$. Since $G(0) = \emptyset$ it follows that $\operatorname{rng} h = G(\infty) = A$. Hence, $h : {\downarrow}\alpha \to A$ is bijective and we can define the desired well-order R over A by

$$R := \{ \langle a, b\rangle \mid h^{-1}(a) \leq h^{-1}(b) \}\,.$$ □

Each of the following statements cannot be proved in basic set theory.

Theorem 1.7. *The following statements are equivalent:*

(1) *For every set A, there exists a well-order R over A.*

(2) *For every set A, there exists a choice function $f : \wp(A) \setminus \{\emptyset\} \to A$.*

(3) *If $(A_i)_{i \in I}$ is a sequence of nonempty sets then $\prod_{i \in I} A_i \neq \emptyset$.*

(4) *If $(A_i)_{i \in I}$ is a sequence of disjoint nonempty sets then $\prod_{i \in I} A_i \neq \emptyset$.*

(5) *Every inductively ordered partial order has a maximal element.*

(6) *If F is a set of finite character and $A \in F$, there exists a maximal element $B \in F$ with $A \subseteq B$.*

(7) *For all sets A and B, there exists an injective function $f : A \to B$ or an injective function $f : B \to A$.*

(8) *For every surjective function $f : A \to B$ where A is a set, there exists a function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$.*

Proof. (2) ⇒ (3) If $\prod_{i \in I} A_i$ is a proper class, it is nonempty and we are done. Hence, we may assume that it is a set. Then $A := \bigcup \{ A_i \mid i \in I \}$ is also a set. By (2) there exists a choice function $f : \wp(A) \setminus \{\varnothing\} \to A$. Let $g : I \to A$ be the function defined by $g(i) := f(A_i)$. Since $g(i) \in A_i$ it follows that $g \in \prod_{i \in I} A_i \neq \varnothing$.

(3) ⇒ (4) is trivial.

(4) ⇒ (2) Let $I := \wp(A) \setminus \{\varnothing\}$ and set $A_X := X \times \{X\}$, for $X \in I$. Since $\prod_{X \in I} A_X \neq \varnothing$ there exists some element $f \in \prod_{X \in I} A_X$. We can define the desired choice function $g : \wp(A) \setminus \{\varnothing\} \to A$ by

$$g(X) = a \quad :\text{iff} \quad f(X) = \langle a, X \rangle .$$

(2) ⇒ (1) was proved in Lemma 1.6.

(1) ⇒ (5) Suppose that $\langle A, \leq \rangle$ is inductively ordered, but A has no maximal element. For every $a \in A$, we can find some $b \in A$ with $b > a$. By assumption, there exists a well-order R over A. Let $f : A \to A$ be the function such that $f(a)$ is the R-minimal element $b \in A$ with $b > a$. By definition, we have $f(a) > a$, for all $a \in A$. Hence, f is inflationary and, by Theorem A3.3.14, f has a fixed point a. But $f(a) = a$ contradicts the definition of f.

(5) ⇒ (6) Let F be a set of finite character and $A \in F$. It is sufficient to prove that the subset $F_0 := \{ X \in F \mid A \subseteq X \}$ is inductively ordered by $\subseteq$. By Lemma 1.3, we know that $\langle F, \subseteq \rangle$ is inductively ordered. Let C be a chain in F_0. Then $C \subseteq F_0 \subseteq F$ and C is also a chain in F. Consequently, it has a least upper bound $B \in F$. Since $A \subseteq X$, for all $X \in C$, it follows that $A \subseteq B$, that is, $B \in F_0$ and B is also the least upper bound of C in F_0.

(6) ⇒ (2) Let A be a set. By Lemma 1.5 (a), the set C of choice functions f with $\operatorname{dom} f \subseteq \wp(A) \setminus \{\varnothing\}$ has finite character and, therefore, there is a maximal element $f \in C$. By Lemma 1.5 (b), it follows that f is the desired choice function.

(1) ⇒ (7) Fix well-orders R and S on, respectively, A and B. By Corollary A3.1.12, exactly one of the following conditions is satisfied:

$$\langle A, R \rangle < \langle B, S \rangle \quad \text{or} \quad \langle A, R \rangle \cong \langle B, S \rangle \quad \text{or} \quad \langle A, R \rangle > \langle B, S \rangle .$$

In the first two cases there exists an injection $A \to B$ and in the second and third case there exists an injection $B \to A$ in the other direction.

(7) ⇒ (1) Let A be a set. By Theorem A3.2.12, there exists an ordinal α such that there is no injective function $\downarrow\alpha \to A$. Consequently, there exists an injective function $f : A \to \downarrow\alpha$. We define a relation R on A by

$$R := \{ \langle a, b \rangle \mid f(a) < f(b) \} .$$

Since f is injective and $\operatorname{rng} f \subseteq \downarrow\alpha$ is well-ordered it follows that R is the desired well-order on A.

(2) ⇒ (8) Let $h : \wp(A) \setminus \{\varnothing\} \to A$ be a choice function. We can define $g : B \to A$ by

$$g(b) := h(f^{-1}(b)) .$$

(8) ⇒ (4) Let $(A_i)_{i \in I}$ be a family of disjoint nonempty sets. We define a function $f : \bigcup \{ A_i \mid i \in I \} \to I$ by

$$f(a) = i \quad :\text{iff} \quad a \in A_i .$$

Since the A_i are disjoint and nonempty it follows that f is well-defined and surjective. Hence, there exists a function $g : I \to \bigcup \{ A_i \mid i \in I \}$ such that $f(g(i)) = i$, for all $i \in I$. By definition of f, this implies that $g(i) \in A_i$. Hence, $g \in \prod_{i \in I} A_i \neq \varnothing$. □

Axiom of Choice. *For every set A there exists a well-order R over A.*

Lemma 1.8. *A left-narrow partial order $\langle A, \leq \rangle$ is well-founded if, and only if, there exists no infinite strictly decreasing sequence $a_0 > a_1 > \dots$.*

Proof. One direction was already proved in Lemma A3.1.3. For the other one, fix a choice function $f : \wp(A) \setminus \varnothing \to A$. Suppose that there exists a nonempty set $A_0 \subseteq A$ without minimal element. We can define a descending chain $a_0 > a_1 > \dots$ by induction. Let $a_0 := f(A_0)$ and, for $k > 0$, set

$$a_k := f(\{ b \in A_0 \mid b < a_{k-1} \}) .$$

Note that a_k is well-defined since a_{k-1} is not a minimal element of A_0. ☐

Exercise 1.2. We call a set a *countable* if there exists a bijection $\downarrow\omega \to a$. Prove that a left-narrow partial order $\langle A, \leq \rangle$ is well-founded if, and only if, every countable nonempty subset $X \subseteq A$ has a minimal element.

Exercise 1.3. Let $\langle A, R \rangle$ be a well-founded partial order that is a set. Prove that there exists a well-order $\leq$ on A with $R \subseteq \leq$.

The following variant of the Axiom of Choice (statement (5) in the above theorem) is known as 'Zorn's Lemma'.

Lemma 1.9 (Kuratowski, Zorn). *Every inductively ordered partial order has a maximal element.*

Example. We have seen that the system of all linearly independent subsets of a vector space V is inductively ordered. It follows that every vector space contains a maximal linearly independent subset, that is, a basis.

This example can be generalised to a certain kind of closure operators.

Definition 1.10. Let c be a closure operator on A.

(a) c has the *exchange property* if

$$b \in c(X \cup \{a\}) \setminus c(X) \quad \text{implies} \quad a \in c(X \cup \{b\}) .$$

(b) A set $I \subseteq A$ is *c-independent* if

$$a \notin c(I \setminus \{a\}) , \qquad \text{for all } a \in I .$$

We call $D \subseteq A$ *c-dependent* if it is not *c-independent.*

(c) Let $X \subseteq A$. A set $I \subseteq X$ is a *c-basis* of X if I is c-independent and $c(I) = c(X)$.

Lemma 1.11. *Let c be a closure operator on A and let $F \subseteq \wp(A)$ be the class of all c-independent sets. If c has finite character then F has finite character.*

Proof. Let $I \in F$ and $I_0 \subseteq I$. For every $a \in I_0$, we have

$$a \notin c(I \setminus \{a\}) \supseteq c(I_0 \setminus \{a\}) .$$

Hence, I_0 is c-independent. Conversely, suppose that $I \notin F$. Then there is some $a \in I$ with

$$a \in c(I \setminus \{a\}) .$$

Since c has finite character we can find a finite subset $I_0 \subseteq I \setminus \{a\}$ with $a \in c(I_0)$. Thus, $I_0 \cup \{a\}$ is a finite subset of I that is not c-independent. ☐

Before proving the converse let us show with the help of the Axiom of Choice that there is always a c-basis. We start with an alternative description of the exchange property.

Lemma 1.12. *Let c be a closure operator on A with the exchange property. If $D \subseteq A$ is a minimal c-dependent set then*

$$a \in c(D \setminus \{a\}) , \quad \text{for all } a \in D .$$

Proof. Let $a \in D$. Since D is c-dependent there exists some element $b \in D$ with $b \in c(D \setminus \{b\})$. If $b = a$ then we are done. Hence, suppose that $b \neq a$ and let $D_0 := D \setminus \{a, b\}$. By minimality of D we have $b \notin c(D_0)$. Hence, $b \in c(D_0 \cup \{a\}) \setminus c(D_0)$ and the exchange property implies that $a \in c(D_0 \cup \{b\})$. ☐

Proposition 1.13. *Let c be a closure operator on A that has finite character and the exchange property. Every set $X \subseteq A$ has a c-basis.*

Proof. The family F of all c-independent subsets of X has finite character. By the Axiom of Choice, there exists a maximal c-independent set $I \subseteq X$. We claim that $c(I) = c(X)$, that is, I is a c-basis of X.

Clearly, $c(I) \subseteq c(X)$. If $X \subseteq c(I)$, it follows that

$$c(X) \subseteq c(c(I)) = c(I)$$

and we are done. Hence, it remains to consider the case that there is some element $a \in X \setminus c(I)$. We derive a contradiction to the maximality of I by showing that $I \cup \{a\}$ is c-independent.

Suppose that $I \cup \{a\}$ is not c-independent. Since F has finite character there exists a finite c-dependent subset $D \subseteq I \cup \{a\}$ with $a \in D$. Suppose that D is chosen minimal. By Lemma 1.12, it follows that $a \in c(D \setminus \{a\}) \subseteq c(I)$. A contradiction. □

Proposition 1.14. *Let c be a closure operator on A with the exchange property and let $F \subseteq \wp(A)$ be the class of all c-independent sets. Then c has finite character if, and only if, F has finite character.*

Proof. $(\Rightarrow)$ has already been proved in Lemma 1.11.

$(\Leftarrow)$ For a contradiction, suppose that there is a set $X \subseteq A$ such that

$$Z := \bigcup \{ c(X_0) \mid X_0 \subseteq X \text{ is finite} \}$$

is a proper subset of $c(X)$. Fix some element $a \in c(X) \setminus Z$. By Proposition 1.13 there exists a c-basis I for X. It follows that $a \in c(X) = c(I)$. Since F has finite character we can find a finite subset $I_0 \subseteq I$ such that $I_0 \cup \{a\}$ is c-dependent. By Lemma 1.12, it follows that $a \in c(I_0) \subseteq Z$. A contradiction. □

A more extensive treatment of closure operators with the exchange property will be given in Section F1.1.

2. Cardinals

The notion of the cardinality of a set is a very natural one. It is based on the same idea which led to the definition of the order type of a well-order. But instead of well-orders we consider just sets without any relation. Although conceptually simpler than ordinals we introduce cardinals quite late in the development of our theory since most of their properties cannot be proved without resorting to ordinals and the Axiom of Choice.

Intuitively, the cardinality of a set A measures its size, that is, the number of its elements. So, how do we count the elements of a set? We can say that 'A has α elements' if there exists an enumeration of A of length α, that is, a bijection $\downarrow\alpha \to A$. For infinite sets, such an enumeration is not unique. We can find several sequences $\downarrow\alpha \to A$ with different values of α. To get a well-defined number we therefore pick the least one.

Definition 2.1. The *cardinality* $|A|$ of a class A is the least ordinal α such that there exists a bijection $\downarrow\alpha \to A$. If there exists no such ordinal then we write $|A| := \infty$. Let $\mathrm{Cn} := \mathrm{rng}\,|\cdot| \subseteq \mathrm{On}$ be the range of this mapping. (We do not consider ∞ to be an element of the range.) We set $\mathfrak{Cn} := \langle \mathrm{Cn}, \leq \rangle$. The elements of Cn are called *cardinals*.

Remark. Clearly, if $|A|, |B| < \infty$ then we have $|A| = |B|$ iff there exists a bijection $A \to B$.

Lemma 2.2. *Every set A has a cardinality and we have $|A| < \infty$.*

Proof. Let A be a set. By the Axiom of Choice, we can find a well-order R over A. Set $\alpha := \mathrm{ord}\langle A, R \rangle$. By definition of an ordinal, there exists a bijection $\downarrow\alpha \to A$. In particular, the class of all ordinals β such that there exists a bijection $\downarrow\beta \to A$ is nonempty and, therefore, there exists a least such ordinal. □

Lemma 2.3. *Let A and B be nonempty sets. The following statements are equivalent:*

(1) $|A| \leq |B|$

(2) *There exists an injective function $A \to B$.*

(3) *There exists an surjective function $B \to A$.*

Proof. Set $\kappa := |A|$ and $\lambda := |B|$ and let $g : \downarrow\kappa \to A$ and $h : \downarrow\lambda \to B$ be the corresponding bijections.

(1) ⇒ (2) Since $\kappa \leq \lambda$ there exists an isomorphism $f : \downarrow\kappa \to I$ between $\downarrow\kappa$ and an initial segment $I \subseteq \downarrow\lambda$. In particular, f is injective. The composition $h \circ f \circ g^{-1} : A \to B$ is the desired injective function.

(2) ⇒ (1) For a contradiction, suppose that there exists an injective function $A \to B$ but we have $|A| > |B|$. By (1) ⇒ (2), the latter implies that there is an injective function $B \to A$. Hence, applying Theorem A2.1.12 we find a bijection $A \to B$. It follows that $|A| = |B|$. Contradiction.

(2) ⇒ (3) Let $f : A \to B$ be injective. By Lemma A2.1.10 (b), there exists a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$. Furthermore, it follows by Lemma A2.1.10 (d) that g is surjective.

(3) ⇒ (2) As above, given a surjective function $f : B \to A$ we can apply Lemma A2.1.10 (and the Axiom of Choice) to obtain an injective function $g : A \to B$ with $f \circ g = \mathrm{id}_B$. □

For every cardinal, there is a canonical set with this cardinality.

Lemma 2.4. *For every cardinal $\kappa \in \mathrm{Cn}$, we have $\kappa = |{\downarrow}\kappa|$. It follows that* $\mathrm{Cn} = \{ \alpha \in \mathrm{On} \mid |{\downarrow}\alpha| = \alpha \}$.

Exercise 2.1. Let α and β be ordinals such that $|\alpha| \leq \beta \leq \alpha$. Show that $|\alpha| = |\beta|$.

Exercise 2.2. Prove that $\alpha \in \mathrm{Cn}$, for every ordinal $\alpha \leq \omega$. *Hint.* Show, by induction on α, that there is no surjective function ${\downarrow}\alpha \to {\downarrow}\beta$ with $\alpha < \beta \leq \omega$.

Using the notion of cardinality we can restate Theorem A2.1.13 in the following way.

Theorem 2.5. *We have $|A| < |\wp(A)|$, for every set A.*

Proof. By Theorem A2.1.13, there exists an injective function $A \to \wp(A)$ but no surjective one. By Lemma 2.3, it follows that $|A| \leq |\wp(A)|$ and $|\wp(A)| \nleq |A|$. □

Cn is a proper class since it is an unbounded subclass of On.

Lemma 2.6. Cn *is a proper class.*

Proof. For a contradiction, suppose otherwise. By Lemma A3.2.8, it follows that there is some $\alpha \in \mathrm{On}$ such that $\kappa < \alpha$, for all cardinals κ. But, by Theorem A3.2.12, there exists some ordinal β such that $\lambda := |{\downarrow}\beta| > |{\downarrow}\alpha|$, which implies that $\lambda > \alpha$. A contradiction. □

Lemma 2.7. $\mathfrak{On}_0 \leq \mathfrak{Cn} \leq \mathfrak{On}$.

Proof. Since $\mathrm{Cn} \subseteq \mathrm{On}$ it follows that $\mathfrak{Cn}$ is a well-order. Therefore, there exists an isomorphism $h : \mathrm{Cn} \to I$, for some initial segment $I \subseteq \mathrm{On}$.

By Theorem 2.5 we know that the function $f : \mathrm{On}_0 \to \mathrm{Cn}$ with $f(\alpha) := |S_\alpha|$ is strictly increasing. Consequently, we have $\mathfrak{On}_0 \leq \mathfrak{Cn}$, by Lemma A3.2.11. □

Remark. With the Axiom of Replacement which we will introduce in Section 5 we can actually prove that $\langle \mathrm{On}_0, \in \rangle \cong \langle \mathrm{On}, < \rangle$. Therefore, all three orders are isomorphic.

Definition 2.8. (a) By the preceding lemma and Lemma A3.1.10, there exists a unique isomorphism $h : I \to \mathrm{Cn}$ where I is an initial segment of On. We define $\aleph_\alpha := h(\omega + \alpha)$ (*'aleph alpha'*), for all α such that $\omega + \alpha \in I$. Furthermore, we denote by ω_α the minimal ordinal such that $|\omega_\alpha| = \aleph_\alpha$.

(b) A set A is *finite* if $|A| < \aleph_0$. Otherwise, A is called *infinite*. Similarly, we say that A is *countable* if $|A| \leq \aleph_0$, and A is *uncountable*, if $|A| > \aleph_0$. A countable set that is not finite is called *countably infinite*.

(c) For cardinals κ, we will denote by κ^+ the minimal *infinite* cardinal greater than κ.

Note that, by our definition of a cardinal, we have $\omega_\alpha = \aleph_\alpha$ and $\aleph_0 = \omega_0 = \omega$. Furthermore, $\aleph_\alpha^+ = \aleph_{\alpha+1}$. Since we have defined the operation κ^+ differently for cardinals and ordinals we will use this notation only for cardinals in the remainder of this book. If we consider the successor of an ordinal α we will write $\alpha + 1$.

3. Cardinal arithmetic

Similarly to ordinals we can define arithmetic operations on cardinals. Note that, except for finite cardinals, these operations are different from the ordinal operations. Therefore, we have chosen different symbols to denote them.

Definition 3.1. Let $\kappa, \lambda \in \mathrm{Cn}$ be cardinals. We define

$$\kappa \oplus \lambda := |{\downarrow}\kappa \cup {\downarrow}\lambda|\,, \quad \kappa \otimes \lambda := |{\downarrow}\kappa \times {\downarrow}\lambda|\,, \quad \kappa^\lambda := |{\downarrow}\kappa^{\downarrow\lambda}|\,.$$

The following lemmas follows immediately from the definition if one recalls that, for $\kappa := |A|$ and $\lambda := |B|$, there exist bijections $A \to {\downarrow}\kappa$ and $B \to {\downarrow}\lambda$.

Lemma 3.2. *Let A and B be sets.*

$$|A \cup B| = |A| \oplus |B|\,, \quad |A \times B| = |A| \otimes |B|\,, \quad |A^B| = |A|^{|B|}.$$

Corollary 3.3. *For all* $\alpha, \beta \in \mathrm{On}$, *we have*

$$|{\downarrow}(\alpha + \beta)| = |{\downarrow}\alpha| \oplus |{\downarrow}\beta| \quad \text{and} \quad |{\downarrow}(\alpha\beta)| = |{\downarrow}\alpha| \otimes |{\downarrow}\beta|\,.$$

The corresponding equation for ordinal exponentiation will be delayed until Lemma 4.4.

Exercise 3.1. Prove that, if A is a set then $|\wp(A)| = 2^{|A|}$. *Hint.* Take the obvious bijection $\wp(A) \to 2^A$.

For finite cardinals these operations coincide with the usual ones.

Lemma 3.4. *For* $m, n < \omega$, *we have*

$$m \oplus n = m + n\,, \quad m \otimes n = mn\,, \quad m^n = m^n\,,$$

where the operations on the left are the ones defined above while those on the right are the usual arithmetic operations.

Let us summarise the basic properties of cardinal arithmetic. The proofs are similar to, but much simpler than, the corresponding ones for ordinal arithmetic.

Lemma 3.5. *Let* $\kappa, \lambda, \mu \in \mathrm{Cn}$.

(a) $(\kappa \oplus \lambda) \oplus \mu = \kappa \oplus (\lambda \oplus \mu)$

(b) $\kappa \oplus \lambda = \lambda \oplus \kappa$

(c) $\kappa \oplus 0 = \kappa$

(d) $\kappa \leq \lambda$ *if, and only if, there is some* μ *with* $\lambda = \kappa \oplus \mu$.

(e) $\lambda \leq \mu$ *implies* $\kappa \oplus \lambda \leq \kappa \oplus \mu$.

(f) $\kappa \geq \aleph_0$ *if, and only if,* $\kappa \oplus 1 = \kappa$

Proof. (a) There is a canonical bijection $(A \cup B) \cup C \to A \cup (B \cup C)$ with

$$\begin{aligned} \langle 0, \langle 0, a\rangle\rangle &\mapsto \langle 0, a\rangle\,, \\ \langle 0, \langle 1, b\rangle\rangle &\mapsto \langle 1, \langle 0, b\rangle\rangle\,, \\ \langle 1, c\rangle &\mapsto \langle 1, \langle 1, c\rangle\rangle\,. \end{aligned}$$

(b) There is a canonical bijection $A \cup B \to B \cup A$ with $\langle 0, a\rangle \mapsto \langle 1, a\rangle$ and $\langle 1, b\rangle \mapsto \langle 0, b\rangle$.

(c) $A \cup \varnothing = \{0\} \times A$. We can define a bijection $A \to \{0\} \times A$ by $a \mapsto \langle 0, a\rangle$.

(d) If $\kappa \leq \lambda$, there exists an injective function $f : {\downarrow}\kappa \to {\downarrow}\lambda$. Let $X := {\downarrow}\lambda \setminus \mathrm{rng}\, f$ and $\mu := |X|$. We can define a bijection ${\downarrow}\kappa \cup X \to {\downarrow}\lambda$ by

$$\langle 0, a\rangle \mapsto f(a) \quad \text{and} \quad \langle 1, a\rangle \mapsto a\,.$$

(e) If there is an injective function $f : B \to C$, we can define an injective function $A \cup B \to A \cup C$ by

$$\langle 0, a\rangle \mapsto \langle 0, a\rangle \quad \text{and} \quad \langle 1, b\rangle \mapsto \langle 1, f(b)\rangle\,.$$

(f) If $\kappa \geq \aleph_0 = \omega$ then $\kappa = \omega + \alpha$, for some $\alpha \in \mathrm{On}$. We can define a bijection ${\downarrow}\omega \to {\downarrow}(\omega+1)$ by $0 \mapsto \omega$ and $n \mapsto n-1$, for $n > 0$. This function can be extended to a bijection ${\downarrow}\omega \cup {\downarrow}\alpha \to {\downarrow}\omega \cup {\downarrow}\alpha \cup [1]$. Conversely, if $\kappa < \omega$ then $\kappa \oplus 1 = \kappa + 1 > \kappa$. □

Lemma 3.6. *Let* $\kappa, \lambda, \mu \in \mathrm{Cn}$.

(a) $(\kappa \otimes \lambda) \otimes \mu = \kappa \otimes (\lambda \otimes \mu)$

(b) $\kappa \otimes \lambda = \lambda \otimes \kappa$

(c) $\kappa \otimes 0 = 0$, $\kappa \otimes 1 = \kappa$, $\kappa \otimes 2 = \kappa \oplus \kappa$.

(d) $\kappa \otimes (\lambda \oplus \mu) = (\kappa \otimes \lambda) \oplus (\kappa \otimes \mu)$

(e) $\lambda \leq \mu$ *implies* $\kappa \otimes \lambda \leq \kappa \otimes \mu$.

Proof. (a) There is a canonical bijection $(A \times B) \times C \to A \times (B \times C)$ with $\langle\langle a, b\rangle, c\rangle \mapsto \langle a, \langle b, c\rangle\rangle$.

(b) There is a canonical bijection $A \times B \to B \times A$ with $\langle a, b\rangle \mapsto \langle b, a\rangle$.

(c) $A \times \varnothing = \varnothing$. There are canonical bijections

$$A \times \{0\} \to A \quad \text{and} \quad A \cup A = [2] \times A \to A \times [2] .$$

(d) There exists a bijection $A \times (B \cup C) \to (A \times B) \cup (A \times C)$ with

$$\langle a, \langle 0, b\rangle\rangle \mapsto \langle 0, \langle a, b\rangle\rangle \quad \text{and} \quad \langle a, \langle 1, c\rangle\rangle \mapsto \langle 1, \langle a, c\rangle\rangle .$$

(e) Given an injective function $f : B \to C$ we define an injective function $A \times B \to A \times C$ by $\langle a, b\rangle \mapsto \langle a, f(b)\rangle$. □

Lemma 3.7. *Let* $\kappa, \lambda, \mu, \nu \in \mathrm{Cn}$.

(a) $(\kappa^\lambda)^\mu = \kappa^{\lambda \otimes \mu}$

(b) $(\kappa \otimes \lambda)^\mu = \kappa^\mu \otimes \lambda^\mu$

(c) $\kappa^{\lambda \oplus \mu} = \kappa^\lambda \otimes \kappa^\mu$

(d) $\kappa^0 = 1$, $\kappa^1 = \kappa$, $\kappa^2 = \kappa \otimes \kappa$.

(e) *If* $\kappa \leq \lambda$ *and* $\mu \leq \nu$ *then* $\kappa^\mu \leq \lambda^\nu$.

(f) $\kappa < 2^\kappa$

Proof. (a) There is a canonical bijection $(A^B)^C \to A^{B \times C}$ given by $f \mapsto g$ where $g(b, c) := f(c)(b)$.

(b) We define a bijection $A^C \times B^C \to (A \times B)^C$ by

$$\langle g, h\rangle \mapsto f \quad \text{where} \quad f(c) := \langle g(c), h(c)\rangle .$$

(c) We define a bijection $A^{B \cup C} \to A^B \times A^C$ by $f \mapsto \langle g, h\rangle$ where

$$g(b) := f(\langle 0, b\rangle) \quad \text{and} \quad h(c) := f(\langle 1, c\rangle) .$$

(d) $A^\varnothing = \{\varnothing\}$. A bijection $A^{[1]} \to A$ is given by $f \mapsto f(0)$, and a bijection $A^{[2]} \to A \times A$ by $f \mapsto \langle f(0), f(1)\rangle$.

(e) Suppose that $f : A \to B$ and $g : C \to D$ are injective. According to Lemma A2.1.10 (b), there exists a surjective function $g' : D \to C$ such that $g' \circ g = \mathrm{id}_C$. We define an injection $A^C \to B^D$ by $h \mapsto f \circ h \circ g'$. To show that this mapping is injective consider functions $h, h' \in A^C$ with $h \neq h'$. Fix some some $c \in C$ with $h(c) \neq h'(c)$ and set $d := g(c)$. Then $g'(d) = g'(g(c)) = \mathrm{id}_C(c) = c$. Since f is injective it follows that

$$(f \circ h \circ g')(d) = f(h(c)) \neq f(h'(c)) = (f \circ h' \circ g')(d) .$$

Consequently, $f \circ h \circ g' \neq f \circ h' \circ g'$.

(f) follows immediately from Theorem 2.5. □

We will show that addition and multiplication of infinite cardinals is especially simple since they just consist of taking the maximum of the operands. In particular, we have $\kappa \oplus \lambda = \kappa \otimes \lambda$ if at least one operand is infinite.

Exercise 3.2. Prove that $\aleph_0 \otimes \aleph_0 = \aleph_0$ by showing that the function

$$\downarrow\omega \times \downarrow\omega \to \downarrow\omega : \langle i, k\rangle \mapsto \tfrac{1}{2}(i + k)(i + k + 1) + k$$

is bijective.

We start by computing $\kappa \otimes \kappa$ by induction on $\kappa \geq \aleph_0$.

Theorem 3.8. *If* $\kappa \geq \aleph_0$ *then* $\kappa \otimes \kappa = \kappa$.

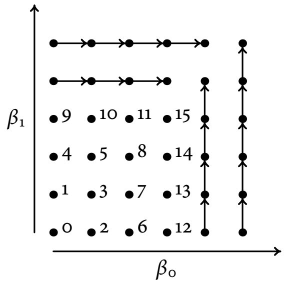


Figure 1.. Ordering on $\downarrow\kappa \times \downarrow\kappa$

Proof. We have $\kappa = \kappa \otimes 1 \le \kappa \otimes \kappa$. For the converse, we prove that $\kappa \otimes \kappa \le \kappa$ by induction on κ.

Note that, since κ is a cardinal we have $\alpha < \kappa$ if, and only if, $|\downarrow\alpha| < \kappa$, for all ordinals α. We define an order on $K := \downarrow\kappa \times \downarrow\kappa$ by

$$\begin{aligned} &\langle \beta_0, \beta_1 \rangle < \langle \gamma_0, \gamma_1 \rangle \\ :\text{iff} \quad &\max\{\beta_0, \beta_1\} < \max\{\gamma_0, \gamma_1\}\,, \text{ or} \\ &\max\{\beta_0, \beta_1\} = \max\{\gamma_0, \gamma_1\} \text{ and } \beta_0 < \gamma_0\,, \text{ or} \\ &\max\{\beta_0, \beta_1\} = \max\{\gamma_0, \gamma_1\} \text{ and } \beta_0 = \gamma_0 \text{ and } \beta_1 < \gamma_1\,. \end{aligned}$$

One can check easily that this order is a well-order. For every ordinal $\alpha \le \kappa$, the set

$$I(\alpha) := \downarrow\alpha \times \downarrow\alpha$$

is an initial subset of K. If $\omega \le \alpha < \kappa$, it follows by inductive hypothesis that

$$|I(\alpha)| = |\downarrow\alpha \times \downarrow\alpha| = |\downarrow\alpha| \otimes |\downarrow\alpha| = |\downarrow\alpha| < \kappa\,.$$

Similarly, if $\alpha < \omega$ then we have

$$|I(\alpha)| = |\downarrow\alpha| \otimes |\downarrow\alpha| = |\downarrow\alpha|^2 = \alpha^2 < \aleph_0 \le \kappa\,.$$

Hence, we have $\operatorname{ord} I(\alpha) < \kappa$, for all ordinals $\alpha < \kappa$.

We claim that $K = \bigcup\{\, I(\alpha) \mid \alpha < \kappa \,\}$. Let $\langle \alpha, \beta \rangle \in K$. Since $\alpha, \beta < \kappa$ and κ is a limit ordinal we have $\gamma := \max\{\alpha + 1, \beta + 1\} < \kappa$ and $\langle \alpha, \beta \rangle \in I(\gamma)$. It follows that

$$\operatorname{ord}\langle K, \le \rangle = \sup\{\, \operatorname{ord}\langle I(\alpha), \le \rangle \mid \alpha < \kappa \,\} \le \kappa\,.$$

In particular, there exists an isomorphism between K and some initial segment of κ. This implies that $\kappa \otimes \kappa = |K| \le \kappa$. ☐

The general case now follows easily.

Lemma 3.9. *If $\kappa > 0$ and $\lambda \ge \aleph_0$ then $\kappa \oplus \lambda = \kappa \otimes \lambda = \max\{\kappa, \lambda\}$.*

Proof. By symmetry, we may assume that $\kappa \le \lambda$. For $\kappa = 1$, the claim follows from Lemmas 3.5 and 3.6. Suppose that $\kappa > 1$. Then

$$\lambda \le \kappa \oplus \lambda \le \lambda \oplus \lambda = 2 \otimes \lambda \le \kappa \otimes \lambda \le \lambda \otimes \lambda = \lambda\,.$$ ☐

Corollary 3.10. *If $\kappa \ge \aleph_0$ then $\kappa^n = \kappa$, for all $n < \omega$.*

Example. We have

$$\begin{aligned} \aleph_4^{\aleph_3} \otimes (\aleph_5 \oplus \aleph_4^{\aleph_7})^{\aleph_2} &= \aleph_4^{\aleph_3} \otimes (\aleph_4^{\aleph_7})^{\aleph_2} = \aleph_4^{\aleph_3} \otimes \aleph_4^{\aleph_7 \otimes \aleph_2} \\ &= \aleph_4^{\aleph_3} \otimes \aleph_4^{\aleph_7} = \aleph_4^{\aleph_3 \oplus \aleph_7} = \aleph_4^{\aleph_7}\,. \end{aligned}$$

4. *Cofinality*

Frequently, we will construct objects as the union of an increasing sequence $A_0 \subseteq A_1 \subseteq \dots$ of sets. In this section we will study the cardinality of such unions.

Definition 4.1. For a sequence $(\kappa_i)_{i<\alpha}$ of cardinals, we define

$$\sum_{i<\alpha} \kappa_i := \Big|\dot{\bigcup_{i<\alpha}} \downarrow\kappa_i\Big| \quad \text{and} \quad \prod_{i<\alpha} \kappa_i := \Big|\prod_{i<\alpha} \downarrow\kappa_i\Big|\,.$$

Lemma 4.2. *If $\kappa \geq \aleph_0$ and $\lambda_i \geq 1$, for $i < \kappa$, then*

$$\sum_{i<\kappa} \lambda_i = \kappa \otimes \sup \{ \lambda_i \mid i < \kappa \} .$$

Proof. Let $\mu := \sup \{ \lambda_i \mid i < \kappa \}$. Note that

$$\kappa = \textstyle\sum_{i<\kappa} 1 \leq \sum_{i<\kappa} \lambda_i \quad \text{and} \quad \mu = \sup \{ \lambda_i \mid i < \kappa \} \leq \sum_{i<\kappa} \lambda_i$$

implies $\kappa \otimes \mu = \max \{\mu, \kappa\} \leq \sum_{i<\kappa} \lambda_i \leq \sum_{i<\kappa} \mu = \kappa \otimes \mu$. □

Corollary 4.3. *If $\kappa \geq \aleph_0$ and $\lambda_i \leq \kappa$, for $i < \kappa$, then $\sum_{i<\kappa} \lambda_i \leq \kappa$.*

We have seen in Lemma 3.7 (f) that $\kappa^\lambda > \kappa$, for infinite λ. Ordinal exponentiation, on the other hand, does not increase the cardinality.

Lemma 4.4. *If α and $\beta > 0$ are ordinals and at least one of them is infinite then*

$$\left|{\downarrow}(\alpha^{(\beta)})\right| = |{\downarrow}\alpha| \otimes |{\downarrow}\beta| .$$

Proof. If $\alpha = 0$ then $|{\downarrow}(\alpha^{(\beta)})| = 0 = |{\downarrow}\alpha| \otimes |{\downarrow}\beta|$. Otherwise, we obviously have $|{\downarrow}\alpha| \leq |{\downarrow}(\alpha^{(\beta)})|$ and $|{\downarrow}\beta| \leq |{\downarrow}(\alpha^{(\beta)})|$. Conversely,

$${\downarrow}(\alpha^{(\beta)}) = \bigcup_{n<\omega} \bigcup \{ ({\downarrow}\alpha)^X \mid X \subseteq {\downarrow}\beta , |X| = n \} .$$

Since $|({\downarrow}\alpha)^n| \leq |{\downarrow}\alpha| \oplus \aleph_0$, for $n < \omega$, it follows from Corollary 4.3 that

$$\begin{aligned}
\left|{\downarrow}(\alpha^{(\beta)})\right| &\leq \sum_{n<\omega} \sum_{i<|({\downarrow}\beta)^n|} |({\downarrow}\alpha)^n| \\
&\leq \sum_{n<\omega} |({\downarrow}\beta)^n| \otimes |{\downarrow}\alpha| \otimes \aleph_0 \\
&= \sum_{n<\omega} |{\downarrow}\alpha| \otimes |{\downarrow}\beta| \otimes \aleph_0 \\
&= \aleph_0 \otimes |{\downarrow}\alpha| \otimes |{\downarrow}\beta| \otimes \aleph_0 \\
&= |{\downarrow}\alpha| \otimes |{\downarrow}\beta| .
\end{aligned}$$

□

Corollary 4.5. *Let A and $B \neq \varnothing$ be sets, at least one of them infinite. There are exactly $|A| \oplus |B|$ functions $p : A_0 \to B$ with finite domain $A_0 \subseteq A$.*

Theorem 4.6 (Kőnig). *If $\kappa_i < \lambda_i$, for $i < \alpha$, then*

$$\sum_{i<\alpha} \kappa_i < \prod_{i<\alpha} \lambda_i .$$

Proof. We show that there is no surjective function

$$f : \textstyle\dot{\bigcup}_{i<\alpha} {\downarrow}\kappa_i \to \prod_{i<\alpha} {\downarrow}\lambda_i .$$

For a contradiction, suppose such a function exists and define

$$Z_k := \left\{ \beta_k < \lambda_k \;\middle|\; (\beta_i)_i = f\langle k, \gamma \rangle \text{ for some } \gamma < \kappa_k \right\} .$$

Then $|Z_k| \leq \kappa_k < \lambda_k$. Hence, ${\downarrow}\lambda_k \setminus Z_k \neq \varnothing$ and there is some sequence $(\beta_i)_i \in \prod_{i<\alpha}({\downarrow}\lambda_i \setminus Z_i)$. As f is surjective there must be some element $\langle k, \gamma \rangle$ with $f\langle k, \gamma \rangle = (\beta_i)_i$. But this implies that $\beta_k \in Z_k$. A contradiction. □

Consider some set A of cardinality $|A| = \kappa$. What is the shortest sequence of sets $(B_\alpha)_{\alpha<\lambda}$ of cardinality $|B_\alpha| < \kappa$ such that $A = \bigcup_{\alpha<\lambda} B_\alpha$? This question leads to the notion of cofinality.

Definition 4.7. (a) Let $\langle A, \leq \rangle$ be a linear order. A subset $X \subseteq A$ is *cofinal* in A if, for every $a \in A$, there is some element $x \in X$ with $a \leq x$.

We call a function $f : B \to A$ cofinal if $\operatorname{rng} f$ is cofinal in A.

(b) The *cofinality* $\operatorname{cf} \alpha$ of an ordinal α is the minimal ordinal β such that there exists a cofinal function $f : {\downarrow}\beta \to {\downarrow}\alpha$.

Exercise 4.1. Prove that every linear order $\langle A, \leq \rangle$ contains a cofinal subset $X \subseteq A$ such that $\langle X, \leq \rangle$ is well-ordered.

Lemma 4.8. *Let $\langle A, \leq \rangle$ be a linear order. If X is cofinal in A and Y is cofinal in X then Y is cofinal in A.*

We can restate the definition of the cofinality in a more useful form as follows.

Lemma 4.9. *If* $(\alpha_i)_{i<\lambda}$ *is a sequence of ordinals* $\alpha_i < \kappa$ *of length* $\lambda < \operatorname{cf}\kappa$ *then*

$$\sup\{\,\alpha_i \mid i < \lambda\,\} < \kappa\,.$$

Exercise 4.2. Prove that $\operatorname{cf}\omega = \omega$ and $\operatorname{cf}\aleph_\omega = \omega$.

The following lemmas provide tools to compute the cofinality of an ordinal.

Lemma 4.10. *For every ordinal* α*, we have*

$$\operatorname{cf}\alpha \leq \alpha \quad \text{and} \quad \operatorname{cf}(\alpha + 1) = 1\,.$$

Proof. For the first inequality, it is sufficient to note that the identity function $\mathrm{id}_{\downarrow\alpha} : {\downarrow\alpha} \to {\downarrow\alpha}$ is cofinal. The second claim follows from the fact that the function $f : [1] \to {\downarrow(\alpha + 1)}$ with $f(0) := \alpha$ is cofinal. □

Lemma 4.11. *If there exists a cofinal function* $f : {\downarrow\beta} \to {\downarrow\alpha}$*, we can construct such a function that is strictly increasing.*

Proof. The function $g : {\downarrow\beta} \to {\downarrow\alpha}$ with

$$g(\gamma) = \max\bigl\{f(\gamma),\ \sup\{\,g(\eta) + 1 \mid \eta < \gamma\,\}\bigr\}$$

is cofinal and increasing. □

Lemma 4.12. *If* $f : {\downarrow\alpha} \to {\downarrow\beta}$ *is strictly increasing and cofinal then* $\operatorname{cf}\alpha = \operatorname{cf}\beta$.

Proof. Let $g : {\downarrow\operatorname{cf}\alpha} \to {\downarrow\alpha}$ and $h : {\downarrow\operatorname{cf}\beta} \to {\downarrow\beta}$ be strictly increasing cofinal maps. Since the composition $f \circ g : {\downarrow\operatorname{cf}\alpha} \to {\downarrow\beta}$ is cofinal we have $\operatorname{cf}\alpha \leq \operatorname{cf}\beta$.

For the converse, we distinguish two cases. If $\alpha = \alpha_0 + 1$ is a successor, then $\operatorname{cf}\alpha = 1$ and $\{f(0)\}$ is cofinal in ${\downarrow\beta}$. Hence, $\beta = f(0) + 1$ is a successor and $\operatorname{cf}\beta = 1$. If α is a limit ordinal, we define a function $k : {\downarrow\operatorname{cf}\beta} \to {\downarrow\alpha}$ by

$$k(\gamma) := \min\{\,\eta \mid f(\eta) > h(\gamma)\,\}\,.$$

This function is cofinal since, given $\eta < \alpha$, there is some $\gamma < \operatorname{cf}\beta$ with $h(\gamma) \geq f(\eta)$. It follows that $k(\gamma) \geq \eta$ since $f(k(\gamma)) > h(\gamma) \geq f(\eta)$ and f is strictly increasing. □

Corollary 4.13. $\operatorname{cf}(\operatorname{cf}\alpha) = \operatorname{cf}\alpha$*, for every* $\alpha \in \mathrm{On}$.

We will see many examples showing that cardinals κ with $\operatorname{cf}\kappa = \kappa$ behave in a sane way while, for other cardinals, we might have to deal with pathological cases. Cardinals of the first kind are therefore called *regular,* the other ones are *singular.*

Definition 4.14. An ordinal α is called *regular* if α is a limit ordinal and $\operatorname{cf}\alpha = \alpha$. A cardinal which is not regular is called *singular.*

Remark. In Corollary 4.13 we have shown that every ordinal of the form $\operatorname{cf}\alpha$ is regular. It follows that the class of all regular ordinals is precisely the range $\operatorname{rng}(\operatorname{cf})$ of the function cf.

Example. ω and $\aleph_1$ are regular while $\aleph_\omega$ is singular.

The next lemma indicates that the notion of cofinality is mainly interesting for cardinals.

Lemma 4.15. *Every regular ordinal is a cardinal.*

Proof. Let $\alpha \in \mathrm{On} \setminus \mathrm{Cn}$ be an ordinal that is not a cardinal and set $\kappa := |\alpha| < \alpha$. By definition, there exists a bijection $f : {\downarrow\kappa} \to {\downarrow\alpha}$. This function is surjective and, hence, cofinal. Consequently, we have $\operatorname{cf}\alpha \leq \kappa < \alpha$. □

It turns out that all successor cardinals are regular while most limit cardinals are singular.

Lemma 4.16. *Every successor cardinal is regular.*

Proof. Suppose there exists a cardinal $\kappa \in \mathrm{Cn}$ such that $\alpha := \operatorname{cf}\kappa^+ < \kappa^+$. Let $f : {\downarrow\alpha} \to {\downarrow\kappa^+}$ be a cofinal map. Since κ^+ is a limit ordinal we have

$${\downarrow\kappa^+} = \bigcup\{\,{\downarrow f(\beta)} \mid \beta < \alpha\,\}\,.$$

By Corollary 4.3, it follows that

$$\kappa^+ = |{\downarrow}\kappa^+| = \Big|\bigcup\{\,{\downarrow}f(\beta) \mid \beta < \alpha\,\}\Big| \leq \sum_{\beta<\alpha} \kappa = \kappa\,.$$

A contradiction. □

Lemma 4.17. *If δ is a limit ordinal then* $\operatorname{cf} \aleph_\delta = \operatorname{cf} \delta$.

Proof. We can define a strictly increasing cofinal function $f : {\downarrow}\delta \to {\downarrow}\aleph_\delta$ by $f(\alpha) := \aleph_\alpha$. Hence, the claim follows from Lemma 4.12. □

It follows that regular limit cardinals are quite rare.

Corollary 4.18. *If δ is a limit ordinal such that $\aleph_\delta$ is regular then $\aleph_\delta = \delta$.*

Cardinal exponentiation is the least understood operation of those introduced so far. There are many open questions that the usual axioms of set theory are not strong enough to answer. For example, we do not know what the value of $2^{\aleph_0}$ is. Given an arbitrary model of set theory we can construct a new model where $2^{\aleph_0} = \aleph_1$, but we can also find models where $2^{\aleph_0}$ equals $\aleph_2$ or $\aleph_3$.

In the remainder of this section we present some elementary results that *can* be proved. The notion of cofinality appears at several places in these proofs. First, let us compute the cardinality of all stages S_α, by a simple induction.

Definition 4.19. We define the cardinal $\beth_\alpha(\kappa)$ (*'beth alpha'*), for $\alpha \in \mathrm{On}$ and $\kappa \in \mathrm{Cn}$, recursively by

$$\begin{aligned} \beth_0(\kappa) &:= \kappa\,, \\ \beth_{\alpha+1}(\kappa) &:= 2^{\beth_\alpha(\kappa)}\,, \\ \text{and} \quad \beth_\delta(\kappa) &:= \sup\{\,\beth_\alpha(\kappa) \mid \alpha < \delta\,\}\,, \quad \text{for limit ordinals } \delta\,. \end{aligned}$$

Further, let $\beth_\alpha := \beth_\alpha(\aleph_0)$.

Lemma 4.20. *For $\alpha \in \mathrm{On}_0$, we have*

$$|S_\alpha| = \beth_\alpha(0) \quad \textit{and} \quad |S_{\omega+\alpha}| = \beth_\alpha\,.$$

The next lemma shows that most questions about cardinal exponentiation can be reduced to the computation of the cardinality of power sets.

Lemma 4.21. *If $2 \leq \kappa \leq 2^\lambda$ and $\lambda \geq \aleph_0$ then $\kappa^\lambda = 2^\lambda$.*

Proof. $2^\lambda \leq \kappa^\lambda \leq (2^\lambda)^\lambda = 2^{\lambda\otimes\lambda} = 2^\lambda$. □

What is the value of κ^λ, for $\lambda < \kappa$? We can give only partial answers.

Lemma 4.22. *If $\kappa \geq \aleph_0$ and $\lambda \geq \operatorname{cf}\kappa$ then $\kappa^\lambda > \kappa$. In particular, $\kappa^{\operatorname{cf}\kappa} > \kappa$.*

Proof. Fix a cofinal function $f : {\downarrow}\lambda \to {\downarrow}\kappa$. By Theorem 4.6, we have

$$\kappa^\lambda = \big|({\downarrow}\kappa)^{{\downarrow}\lambda}\big| = \Big|\prod_{\alpha<\lambda} {\downarrow}\kappa\Big| > \Big|\dot{\bigcup_{\alpha<\lambda}} {\downarrow}f(\alpha)\Big| \geq |{\downarrow}\kappa| = \kappa\,.$$ □

Corollary 4.23. $\operatorname{cf} 2^\kappa > \kappa$.

Proof. $\operatorname{cf} 2^\kappa \leq \kappa$ would imply $(2^\kappa)^{\operatorname{cf} 2^\kappa} \leq (2^\kappa)^\kappa = 2^{\kappa\otimes\kappa} = 2^\kappa < (2^\kappa)^{\operatorname{cf} 2^\kappa}$. Contradiction. □

The next theorem summarises the extend of our knowledge about cardinal exponentiation. First, we introduce some abbreviations.

Definition 4.24. For cardinals κ and λ we set

$$({<}\kappa)^\lambda := \sup\{\,\mu^\lambda \mid \mu < \kappa\,\} \quad \text{and} \quad \kappa^{<\lambda} := \sup\{\,\kappa^\mu \mid \mu < \lambda\,\}\,.$$

Lemma 4.25. $\operatorname{cf}({<}\kappa)^\lambda \leq \operatorname{cf}\kappa$ *and* $\operatorname{cf}\kappa^{<\lambda} \leq \operatorname{cf}\lambda$.

Theorem 4.26. *Let $\kappa \geq 2$ and $\lambda \geq \aleph_0$.*

(a) *If $2 < \kappa \leq \lambda$ then $\kappa^\lambda = 2^\lambda = ({<}\kappa)^\lambda$.*

(b) *If $\operatorname{cf}\kappa \leq \lambda < \kappa$ then $\kappa < \kappa^\lambda = (({<}\kappa)^\lambda)^{\operatorname{cf}\kappa} \leq 2^\kappa$.*

(c) *If* $\lambda < \operatorname{cf}\kappa$ *then* $\kappa^\lambda = \kappa \oplus (<\kappa)^\lambda$.

Proof. (a) The first equality was proved in Lemma 4.21. For the second one, note that $\kappa > 2$ implies $2^\lambda \leq (<\kappa)^\lambda \leq \kappa^\lambda$.

(b) By (a) and Corollary 4.22, it follows that $\kappa < \kappa^\lambda \leq 2^\kappa$. Further, $(<\kappa)^\lambda \leq \kappa^\lambda$ implies that

$$((<\kappa)^\lambda)^{\operatorname{cf}\kappa} \leq (\kappa^\lambda)^{\operatorname{cf}\kappa} = \kappa^{\lambda \otimes \operatorname{cf}\kappa} = \kappa^\lambda .$$

For the converse, fix a cofinal function $f : \downarrow \operatorname{cf}\kappa \to \downarrow\kappa$. We have

$$\begin{aligned} \kappa^\lambda \leq \Big| \dot\bigcup_{\alpha < \operatorname{cf}\kappa} \downarrow f(\alpha) \Big|^\lambda &\leq \Big| \prod_{\alpha < \operatorname{cf}\kappa} \downarrow f(\alpha) \Big|^\lambda \\ &= \Big| \prod_{\alpha < \operatorname{cf}\kappa} \downarrow f(\alpha)^{\downarrow\lambda} \Big| \\ &\leq \Big| \prod_{\alpha < \operatorname{cf}\kappa} \downarrow (<\kappa)^\lambda \Big| \leq ((<\kappa)^\lambda)^{\operatorname{cf}\kappa} . \end{aligned}$$

(c) If $\lambda < \operatorname{cf}\kappa$ then

$$(\downarrow\kappa)^{\downarrow\lambda} = \bigcup \{ (\downarrow\mu)^{\downarrow\lambda} \mid \mu < \kappa \} ,$$

since the range of every function $\downarrow\lambda \to \downarrow\kappa$ is bounded by some $\mu < \kappa$. Hence,

$$\kappa^\lambda \leq \sum_{\mu < \kappa} \mu^\lambda \leq \sum_{\mu < \kappa} (<\kappa)^\lambda = \kappa \otimes (<\kappa)^\lambda .$$

If $\kappa = \mu^+$ then $(<\kappa)^\lambda = \mu^\lambda$ and

$$\kappa^\lambda \leq \kappa \otimes (<\kappa)^\lambda = \kappa \otimes \mu^\lambda \leq \kappa^\lambda .$$

Otherwise, κ is a limit and $(<\kappa)^\lambda \geq \sup \{ \mu \mid \mu < \kappa \} = \kappa$, which implies that

$$\kappa^\lambda \leq \kappa \otimes (<\kappa)^\lambda = (<\kappa)^\lambda \leq \kappa^\lambda . \qquad \square$$

Corollary 4.27. *If* κ *and* λ *are cardinals such that* $2^\mu = \mu^+$, *for all* $\mu \leq \kappa$, *then*

$$\kappa^\lambda = \begin{cases} 2^\lambda & \text{if } \kappa \leq \lambda , \\ \kappa^+ & \text{if } \operatorname{cf}\kappa \leq \lambda < \kappa , \\ \kappa & \text{if } \lambda < \operatorname{cf}\kappa . \end{cases}$$

Lemma 4.28. *Let* κ *be a cardinal. We have* $\kappa = \beth_\delta$, *for some limit ordinal* δ, *if and only if* $\kappa > \aleph_0$ *and* $2^\lambda < \kappa$, *for all* $\lambda < \kappa$.

Proof. ($\Rightarrow$) We have $\beth_\delta > \beth_0 = \aleph_0$. If $\lambda < \beth_\delta$ then $\lambda \leq \beth_\alpha$, for some $\alpha < \delta$. Hence, $2^\lambda \leq 2^{\beth_\alpha} = \beth_{\alpha+1} < \beth_\delta$.

($\Leftarrow$) Let $A := \{ \alpha + 1 \mid \beth_\alpha < \kappa \}$ and $\delta := \sup A$. By definition of A, it follows that $\beth_\delta \geq \kappa$. On the other hand,

$$\begin{aligned} \kappa &= \sup \{ 2^\lambda \mid \lambda < \kappa \} \\ &\geq \sup \{ 2^{\beth_\alpha} \mid \beth_\alpha < \kappa \} = \sup \{ \beth_\alpha \mid \alpha \in A \} = \beth_\delta . \end{aligned}$$

Hence, $\kappa = \beth_\delta$. Since $\beth_\delta = \kappa > \aleph_0$ we have $\delta > 0$. To show that δ is a limit suppose that $\delta = \alpha + 1$. Then $\beth_\alpha < \kappa$ implies $\beth_\delta = 2^{\beth_\alpha} < \kappa$. Contradiction. $\square$

We conclude this section with some results about sets of sequences indexed by ordinals. As we will see in Section B2.1, such a set forms the domain of a *tree*. Recall that a sequence indexed by an ordinal α is just a function $\downarrow\alpha \to A$.

Definition 4.29. If A is a set and $\alpha \in \mathrm{On}$, we define

$$A^\alpha := A^{\downarrow\alpha} \quad \text{and} \quad A^{<\alpha} := \bigcup_{\beta < \alpha} A^\beta .$$

Let us compute the cardinality of $A^{<\alpha}$. We are especially interested in the case where $\alpha = \omega$, i.e., in the set of all finite sequences.

Lemma 4.30. *If* $|A| > 1$ *then* $|A^{<\alpha}| = |A|^{<|\alpha|}$.

Lemma 4.31. *If $\kappa > 0$ then $\kappa^{<\aleph_0} = \kappa \oplus \aleph_0$.*

Proof. If $\kappa \geq \aleph_0$ then

$$\kappa^{<\aleph_0} = \sup\{\, \kappa^n \mid n < \aleph_0 \,\} = \sup\{\kappa\} = \kappa = \kappa \oplus \aleph_0 \,.$$

For $\kappa = 1$, we can define a bijection $[1]^{<\omega} \to {\downarrow}\omega$ by

$$\underbrace{\langle 0, \dots, 0 \rangle}_{n \text{ times}} \mapsto n \,.$$

Hence, $1^{<\aleph_0} = \aleph_0$. If $1 < \kappa < \aleph_0$, it follows that

$$\aleph_0 = 1^{<\aleph_0} \leq \kappa^{<\aleph_0} \leq \aleph_0^{<\aleph_0} = \aleph_0 \,. \qquad \square$$

Corollary 4.32. *$\kappa^{<\kappa} \geq \kappa$, for all $\kappa > 0$. If $\kappa \geq \aleph_0$ then $\kappa \leq 2^{<\kappa} \leq \kappa^{<\kappa}$.*

Proof. If $\kappa \geq \aleph_0$ then $2^{<\kappa} = \sup\{\, 2^\lambda \mid \lambda < \kappa \,\} \geq \sup\{\, \lambda^+ \mid \lambda < \kappa \,\} \geq \kappa$. □

Lemma 4.33. *If κ is an infinite regular cardinal then $\kappa^{<\kappa} = 2^{<\kappa}$.*

Proof. For $\aleph_0 \leq \lambda, \mu < \kappa$ we have

$$\lambda^\mu \leq (\lambda \oplus \mu)^{\lambda \oplus \mu} = 2^{\lambda \oplus \mu} \leq 2^{<\kappa}.$$

If $\operatorname{cf} \kappa = \kappa$, it follows by Theorem 4.26 and Corollary 4.32 that

$$\kappa^\mu = \kappa \oplus (<\kappa)^\mu = \kappa \oplus \sup\{\, \lambda^\mu \mid \lambda < \kappa \,\} \leq 2^{<\kappa}, \quad \text{for all } \mu < \kappa \,.$$

Consequently, $\kappa^{<\kappa} \leq 2^{<\kappa}$. □

Corollary 4.34. *Let κ be an infinite cardinal. We have $\kappa^{<\kappa} = \kappa$ if, and only if, κ is regular and $2^{<\kappa} = \kappa$.*

Proof. One direction follows from the preceding lemma. For the other one, note that $\operatorname{cf} \kappa < \kappa$ implies $\kappa^{<\kappa} \geq \kappa^{\operatorname{cf} \kappa} > \kappa$, and $2^{<\kappa} > \kappa$ implies $\kappa^{<\kappa} \geq 2^{<\kappa} > \kappa$. □

5. The Axiom of Replacement

At several times when mappings between classes were concerned we remarked that we need an additional axiom to prove the desired statement. This axiom is the generalisation of the following lemma to functions that are proper classes.

Lemma 5.1. *Let f be a function. If f is a set then so is $f[A]$, for all $A \subseteq \operatorname{dom} f$.*

Proof. Since f is a set so is $\operatorname{rng} f$. Therefore,

$$f[A] = \{\, y \in \operatorname{rng} f \mid y = f(x) \text{ for some } x \in A \,\}$$

is a set. □

Before stating the axiom let us collect several equivalent formulations of it.

Theorem 5.2. *The following statements are equivalent:*

(1) *If F is a function and $A \subseteq \operatorname{dom} F$ is a set then $F[A]$ is also a set.*

(2) *If F is a function and $\operatorname{dom} F$ is a set then so is $\operatorname{rng} F$.*

(3) *A function F is a set if, and only if, $\operatorname{dom} F$ is a set.*

(4) *There exists no bijection $F : a \to B$ between a set a and a proper class B.*

(5) *A class A is a set if, and only if, $|A| < \infty$.*

(6) *If $\alpha \in \mathrm{On}$ is an ordinal and $(A_i)_{i<\alpha}$ a sequence of sets then the class $\bigcup_{i<\alpha} A_i$ is also a set.*

Proof. (3) ⇒ (2) Let F be a function and suppose that $\operatorname{dom} F$ is a set. Then F is a set and so is $\operatorname{rng} F$.

(2) ⇒ (3) Clearly, if F is a set then so is $\operatorname{dom} F$. For the converse, let F be a function such that $\operatorname{dom} F$ is a set. By assumption, then $\operatorname{rng} F$ is also a set. Since $F \subseteq \operatorname{dom} F \times \operatorname{rng} F$ it follows that F is a set.

(2) ⇒ (1) Let F be a function and $A \subseteq \operatorname{dom} F$ a set. Let $G := F \upharpoonright A$ be the restriction of F to A. We apply the assumption to G. Since $\operatorname{dom} G = A$ is a set so is $\operatorname{rng} G = F[A]$.

(1) ⇒ (6) Let $F : \downarrow\alpha \to \mathbb{S}$ be the function with $F(i) = A_i$, for $i < \alpha$. By assumption, $B := F[\downarrow\alpha]$ is a set. Hence, so is

$$\bigcup B = \bigcup_{i<\alpha} A_i \,.$$

(6) ⇒ (2) Let $F : A \to B$ be a function and $A = \operatorname{dom} F$ a set. Let $\kappa := |A|$ and fix a bijection $g : \downarrow\kappa \to A$. We define a sequence $(B_i)_{i<\kappa}$ of sets by $B_i := S(F(g(i)))$. By assumption, $C := \bigcup_{i<\kappa} B_i$ is a set. For every $a \in A$, we have $S(F(a)) \subseteq C$ or, equivalently, $S(F(a)) \in \wp(C)$. It follows that $S(\operatorname{rng} F) = S(F[A]) \subseteq \wp(C)$. In particular, $\operatorname{rng} F$ is a set.

(2) ⇒ (5) If A is a set then $|A| < \infty$, by Lemma 2.2. For the converse, suppose that $\kappa := |A| < \infty$ and let $F : \downarrow\kappa \to A$ be a bijection. Since κ is a set it follows by assumption that $A = \operatorname{rng} F$ is also a set.

(5) ⇒ (4) Let $F : a \to B$ be a bijection where a is a set. Then $|B| = |a| < \infty$. Hence, B is also a set.

(4) ⇒ (2) Let $F : A \to B$ be a function where $A = \operatorname{dom} F$ is a set. Let $B_0 := \operatorname{rng} F$. Since the function $F : a \to B_0$ is surjective there exists a function $G : B_0 \to a$ such that $F \circ G = \operatorname{id}_{B_0}$. Let $A_0 := \operatorname{rng} G$. The restriction $F : A_0 \to B_0$ is a bijection. Since $A_0 \subseteq A$ is a set so is $B_0 = \operatorname{rng} F$. □

Axiom of Replacement. *If F is a function and* $\operatorname{dom} F$ *is a set then so is* $\operatorname{rng} F$.

Let us finally prove the results we promised in the preceding sections. First, up to isomorphism, $\mathfrak{On}$ is the only well-order that is a proper class.

Lemma 5.3. *Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be well-orders. If A and B are proper classes then $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Suppose that $\mathfrak{A} \ncong \mathfrak{B}$. By Theorem A3.1.11, there either exists an isomorphism $f : A \to \downarrow b$, for some $b \in B$, or some isomorphism $g : \downarrow a \to B$, for some $a \in A$. By symmetry, we may assume w.l.o.g. the latter. $\downarrow a$ is a set since $\leq_A$ is left-narrow. Hence, by the Axiom of Replacement, $B = g[\downarrow a]$ is also a set. Contradiction. □

It follows that it does not matter which of the two definitions of an ordinal we adopt.

Corollary 5.4. $\mathfrak{On}_0 \cong \mathfrak{Cn} \cong \mathfrak{On}$.

Finally, we state the general form of the Principle of Transfinite Recursion.

Theorem 5.5 (Principle of Transfinite Recursion)**.** *If $H : A^{<\infty} \to A$ is a total function that defines the function F by transfinite recursion then* $\operatorname{dom} F = \mathrm{On}$.

Proof. For a contradiction, suppose that $\operatorname{dom} F = \downarrow\alpha \subset \mathrm{On}$. In particular, $\operatorname{dom} F$ is a set. By the Axiom of Replacement, it follows that $\operatorname{rng} F$ is also a set. Since $\operatorname{rng} F \subseteq A$ we therefore have $F \in A^{<\infty} = \operatorname{dom} H$ in contradiction to Theorem A3.3.4. □

Lemma 5.6. *Every strictly continuous function $f : \mathrm{On} \to \mathrm{On}$ has arbitrarily large fixed points.*

Proof. For every $\alpha \in \mathrm{On}$ we have to find a fixed point $\gamma \geq \alpha$. If F is the fixed-point induction of f over α then $F[\downarrow\omega]$ exists. By Lemma A3.3.13 it follows that $\gamma := F(\infty) = F(\omega) \geq \alpha$ is a fixed point of f. □

Corollary 5.7. *There are arbitrarily large cardinals κ such that* $\operatorname{cf}\kappa = \aleph_0$ *and either $\aleph_\kappa = \kappa$ or $\beth_\kappa = \kappa$.*

Proof. The functions $f : \alpha \mapsto \aleph_\alpha$ and $g : \alpha \mapsto \beth_\alpha$ are strictly continuous. Furthermore, they are defined by transfinite recursion. Therefore, Theorem 5.5 implies that their domain is all of On. By Lemma A3.3.13 and Lemma 5.6, it follows that f and g have arbitrarily large inductive fixed points κ, and these fixed points are of the form

$$\kappa = \sup\{\, f^n(\alpha) \mid n < \omega \,\}, \qquad \text{for some } \alpha \,.$$

In particular, $\mathrm{cf}\,\kappa = \aleph_0$. □

Exercise 5.1. Prove that S_{ω_2} satisfies all axioms of set theory except for the Axiom of Replacement.

6. *Stationary sets*

There are many places in mathematics where one wants to argue that there are 'many' objects with a certain property. This has lead to several notions of 'large' and 'small' sets, for instance, being dense, being cofinite, having measure 1, or belonging to a given ultrafilter.

Example. Let κ be a regular cardinal and A a set of size $|A| = \kappa$. We call a subset $X \subseteq A$ *large* if it has size κ. A subset $X \subseteq A$ is *very large* if its complement $A \setminus X$ is not large. It is straightforward to check that the classes of large and very large sets have the following properties:

(a) Every very large set is large.

(b) A set X is large if, and only if, it has a non-empty intersection with every very large set.

(c) The intersection of less than κ very large sets is very large.

(d) The intersection of a very large set and a large one is large.

(e) Every large set can be partitioned into κ disjoint large subsets.

(f) If $f : X \to Y$ is a function from a large set X into a set Y that is not large, there is some element $y \in Y$ such that the fibre $f^{-1}(y)$ is large.

In this section we introduce two notions of 'largeness' for sets of ordinals which exhibit the same properties as the large and very large sets of the above example: *closed unbounded sets* correspond to the very large sets and *stationary sets* correspond to the large one. We will prove analogues to all of the above properties. We start with closed unbounded sets.

Definition 6.1. Let κ be a cardinal. A subset $C \subseteq \kappa$ is *closed unbounded* if it is cofinal in κ and, for every non-empty subset $X \subseteq C$ with $\sup X < \kappa$, we have $\sup X \in C$.

Example. For every ordinal $\alpha < \kappa$, the set $\Uparrow\alpha$ is obviously closed unbounded. Another example of a closed unbounded set is the set of all limit ordinals $\alpha < \kappa$.

Before verifying the above properties let us present two ways to construct closed unbounded subsets of a given closed unbounded set.

Lemma 6.2. *Let κ be an uncountable regular cardinal and $C \subseteq \kappa$ closed unbounded.*

(a) *The set $C' := \{ \alpha \in C \mid C \cap \alpha \text{ is cofinal in } \alpha \}$ is closed unbounded.*

(b) *For every cardinal λ such that $C \cap \lambda$ is cofinal in λ, the set $C \cap \lambda$ is closed unbounded in λ.*

Proof. (a) To show that C' is cofinal, let $\alpha < \kappa$. Since C is cofinal, we can construct an increasing sequence $\alpha < \beta_0 < \beta_1 < \dots$ of elements $\beta_n \in C$, for $n < \omega$. Since C is closed and κ is regular, it follows that $\delta := \sup_{n<\omega} \beta_n \in C$. Furthermore, the fact that all β_n belong to $C \cap \delta$ implies that $C \cap \delta$ is cofinal in δ. Hence, $\delta \in C'$.

It remains to show that C' is closed. Consider a set $X \subseteq C'$ such that $\delta := \sup X < \kappa$. If $\delta \in X \subseteq C'$, we are done. Hence, we may assume that $\delta \notin X$. Note that $X \subseteq C$ implies that $\delta \in C$. Furthermore, $X \subseteq C \cap \delta$ implies that $C \cap \delta$ is cofinal in δ. Consequently, $\delta \in C'$.

(b) By assumption, $C \cap \lambda$ is cofinal in λ. To show that it is also closed, let $X \subseteq C \cap \lambda$ be a set with $\sup X < \lambda$. Then $X \subseteq C$ implies that $\sup X \in C$. Hence, $\sup X \in C \cap \lambda$. □

The first property we check is that closed unbounded sets are closed under intersections. We consider two variants: ordinary intersections and so-called diagonal intersections.

Lemma 6.3. *Let κ be an uncountable regular cardinal. If $C, D \subseteq \kappa$ are closed unbounded then so is $C \cap D$.*

Proof. If $X \subseteq C \cap D$ and $\sup X < \kappa$ then $X \subseteq C$ implies $\sup X \in C$ and $X \subseteq D$ implies $\sup X \in D$. Consequently, we have $\sup X \in C \cap D$.

To show that $C \cap D$ is cofinal let $\alpha < \kappa$. Then there is some element $\beta_0 \in C$ with $\alpha \leq \beta_0$. Similarly, there is some element $\gamma_0 \in D$ with $\beta_0 \leq \gamma_0$. Continuing in this way we obtain an increasing sequence

$$\alpha \leq \beta_0 \leq \gamma_0 \leq \beta_1 \leq \gamma_1 \leq \dots$$

where $\beta_i \in C$ and $\gamma_i \in D$. Since $\operatorname{cf} \kappa > \omega$ it follows that

$$\delta := \sup_i \beta_i = \sup_i \gamma_i < \kappa \,.$$

As C and D are closed unbounded we have $\delta \in C$ and $\delta \in D$. Thus, we have found an element $\delta \in C \cap D$ with $\alpha \leq \delta$. □

Exercise 6.1. Show that this lemma fails for closed unbounded subsets of $\aleph_0$.

Proposition 6.4. *Let κ be an uncountable regular cardinal. If $\mathcal{C} \subseteq \wp(\kappa)$ is a family of closed unbounded sets with $|\mathcal{C}| < \kappa$ then $\bigcap \mathcal{C}$ is closed unbounded.*

Proof. Let $(C_i)_{i<\alpha}$ be a sequence of closed unbounded subsets of κ with $\alpha < \kappa$. By induction on α, we prove that $\bigcap_{i<\alpha} C_i$ is closed unbounded.

For $\alpha = 1$ there is nothing to do and the successor step follows immediately from the preceding lemma. Hence, we may assume that α is a limit ordinal. Furthermore, we know by inductive hypothesis that the sets $\bigcap_{i<\beta} C_i$, for $\beta < \alpha$ are closed unbounded. Therefore, replacing C_β by $\bigcap_{i\leq\beta} C_i$ we may assume that $C_0 \supseteq C_1 \supseteq \dots$.

Let $C := \bigcap_{i<\alpha} C_i$. If $X \subseteq C$ is a set with $\sup X < \kappa$, then $X \subseteq C_i$ implies that $\sup X \in C_i$, for all i. Consequently, we have $\sup X \in C$.

To show that C is cofinal let $\beta < \kappa$. We construct an increasing sequence $(\gamma_i)_{i<\alpha}$ as follows. Choose some $\gamma_0 \in C_0$ with $\beta \leq \gamma_0$. For $0 < i < \alpha$, let $\gamma_i \in C_i$ be some element with $\gamma_i \geq \sup \{ \gamma_k \mid k < i \}$. Since κ is regular it follows that $\delta := \sup_i \gamma_i < \kappa$. For $i < \alpha$, let

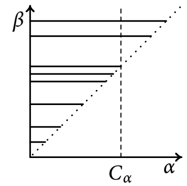
$$X_i := \{ \gamma_k \mid i \leq k < \alpha \} \,.$$

Then $X_i \subseteq C_i$. Since C_i is closed unbounded it follows that $\delta = \sup X_i \in C_i$. Consequently, we have found an element $\delta \in C$ with $\beta \leq \delta$. □

The second variant of intersection we consider has no correspondence in the above example since it relies on the presence of a linear order.

Definition 6.5. The *diagonal intersection* of a sequence $(C_\alpha)_{\alpha<\kappa}$ of subsets $C_\alpha \subseteq \kappa$ is the set

$$D := \{ \beta < \kappa \mid \beta \in C_\alpha \text{ for all } \alpha < \beta \} \,.$$



Remark. Note that, if D is the diagonal intersection of $(C_\alpha)_{\alpha<\kappa}$, then $D \setminus (\alpha + 1) \subseteq C_\alpha$, for all α.

Proposition 6.6. *Let κ be an uncountable regular cardinal. The diagonal intersection of a sequence $(C_\alpha)_{\alpha<\kappa}$ of closed unbounded sets is closed unbounded.*

Proof. Let $(C_\alpha)_{\alpha<\kappa}$ be a sequence of closed unbounded sets and let D be their diagonal intersection. By Proposition 6.4, the intersections $C'_\alpha := \bigcap_{\beta<\alpha} C_\beta$ are closed unbounded. Furthermore, the diagonal intersection of $(C'_\alpha)_{\alpha<\kappa}$ is also equal to D. Replacing C_α by C'_α, we may therefore assume that the sequence $(C_\alpha)_{\alpha<\kappa}$ is decreasing.

To show that D is closed, let $X \subseteq D$ be a set with $\delta := \sup X < \kappa$. For $\alpha < \delta$, consider the set $Y_\alpha := \{ \beta \in D \mid \alpha < \beta < \delta \}$. By the definition of the diagonal intersection, we have $Y_\alpha \subseteq D \setminus (\alpha + 1) \subseteq C_\alpha$. As C_α is closed, it follows that $\delta = \sup Y_\alpha \in C_\alpha$, for all $\alpha < \delta$. Consequently, $\delta \in D$.

To show that D is unbounded, let $\alpha < \kappa$. To find a bound $\delta \in D$ with $\alpha < \delta$, we construct an increasing sequence $(\beta_n)_{n<\omega}$ of ordinals as follows. Choose some element $\beta_0 \in C_0$ with $\beta_0 > \alpha$. If β_n is already defined, we choose an element $\beta_{n+1} \in C_{\beta_n}$ with $\beta_{n+1} > \beta_n$. We claim that

$\delta := \sup_{n<\omega} \beta_n \in D$. Hence, let $\gamma < \delta$. Then there is some $n < \omega$ with $\gamma < \beta_n$. Since $\beta_k \in C_{\beta_{k-1}} \subseteq C_{\beta_n}$, for $k > n$, it follows that $\delta = \sup_{k>n} \beta_k \in C_{\beta_n} \subseteq C_\gamma$. Hence, $\delta \in C_\gamma$, for all $\gamma < \delta$. This implies that $\delta \in D$. □

Our second notion of a large set is that of a stationary one. As definition we use the analogue of Property (b) from the above example.

Definition 6.7. Let κ be a cardinal. A set $S \subseteq \kappa$ is *stationary* if $S \cap C \neq \varnothing$, for every closed unbounded set $C \subseteq \kappa$.

We start by constructing several kinds of stationary sets.

Lemma 6.8. *Let κ be an uncountable regular cardinal.*

(a) *The set $\{ \alpha < \kappa \mid \operatorname{cf} \alpha = \lambda \}$ is stationary, for every regular $\lambda < \kappa$.*

(b) *Every closed unbounded set is stationary.*

(c) *If S is stationary and C closed unbounded, then $S \cap C$ is stationary.*

Proof. (a) Let $C \subseteq \kappa$ be closed unbounded. We have to find some element $\gamma \in C$ with cofinality λ. Let $f : \langle \kappa, \leq \rangle \to \langle C, \leq \rangle$ be an order isomorphism and set $\gamma := \sup f[\lambda]$. Since C is closed unbounded, we have $\gamma \in C$. As the function $f \restriction \lambda : \lambda \to \gamma$ is a strictly increasing and cofinal, it follows by Lemma 4.12 that $\operatorname{cf} \gamma = \operatorname{cf} \lambda = \lambda$.

(b) Let C be closed unbounded. For every closed unbounded set D, it follows by Lemma 6.3 that the intersection $C \cap D$ is also closed unbounded. In particular, $C \cap D \neq \varnothing$.

(c) If there were a closed unbounded set D with $(S \cap C) \cap D = \varnothing$, then S would not be stationary since $C \cap D$ is closed unbounded, by Lemma 6.3. □

Note that it follows from Lemma 6.8 (a) that there are disjoint stationary sets. Hence, the intersection of two stationary sets is not necessarily stationary.

The next theorem is a very strong version of Property (f) from the example.

Theorem 6.9 (Fodor). *Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and $f : S \to \kappa$ a function with $f(\alpha) < \alpha$, for all $\alpha \in S$. Then there exists an ordinal $\gamma < \kappa$ such that $f^{-1}(\gamma)$ is stationary.*

Proof. For a contradiction, suppose that $f^{-1}(\gamma)$ is non-stationary, for every $\gamma < \kappa$. For each $\gamma < \kappa$, choose a closed unbounded set $C_\gamma \subseteq \kappa$ such that $C_\gamma \cap f^{-1}(\gamma) = \varnothing$. By Proposition 6.6, the diagonal intersection D of $(C_\gamma)_{\gamma<\kappa}$ is closed unbounded. Consequently, Lemma 6.8 (c) implies that $S \cap D$ is stationary. Fix an element $\alpha \in S \cap D$. Then $\alpha \in C_\gamma$, for all $\gamma < \alpha$. Since $C_\gamma \cap f^{-1}(\gamma) = \varnothing$, it follows that $\alpha \notin f^{-1}(\gamma)$. Thus, $f(\alpha) \neq \gamma$, for all $\gamma < \alpha$, which implies that $f(\alpha) \geq \alpha$. A contradiction. □

Corollary 6.10. *Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and $f : S \to \lambda$ a function with $\lambda < \kappa$. Then there exists an ordinal $\gamma < \lambda$ such that $f^{-1}(\gamma)$ is stationary.*

Proof. By Lemma 6.8 (c), the set $S' := S \setminus \lambda$ is stationary. Since $f(\alpha) < \alpha$, for $\alpha \in S'$, we can apply the Theorem of Fodor to $f \restriction S'$ to find the desired ordinal γ. □

As an application, we prove the existence of so-called *sunflowers.*

Lemma 6.11 (Sunflower lemma). *Let κ be a regular cardinal and λ a cardinal such that $\mu^{<\lambda} < \kappa$, for all $\mu < \kappa$.*

For every family $(S_\alpha)_{\alpha<\kappa}$ of sets of size $|S_\alpha| < \lambda$, there exists a set U and a subset $I \subseteq \kappa$ of size $|I| = \kappa$ such that

$$S_\alpha \cap S_\beta = U\,, \quad \textit{for all distinct } \alpha, \beta \in I\,.$$

Proof. First, we consider the case where $\kappa = \aleph_0$. Then λ is finite and we can prove the claim by induction on λ. We distinguish two cases. If there is no element a that belongs to infinitely many sets S_α, we can choose a set $I \subseteq \kappa$ such that

$$S_\alpha \cap S_\beta = \varnothing\,, \quad \text{for all distinct } \alpha, \beta \in I\,.$$

Otherwise, choose such an element a and set $K := \{ \alpha < \kappa \mid a \in S_\alpha \}$. Applying the inductive hypothesis to the family $(S_\alpha \setminus \{a\})_{\alpha \in K}$, we obtain an infinite set $I \subseteq K$ and some set U' such that

$$(S_\alpha \setminus \{a\}) \cap (S_\beta \setminus \{a\}) = U', \quad \text{for all distinct } \alpha, \beta \in I.$$

Consequently, the sets I and $U := U' \cup \{a\}$ have the desired properties.

It remains to consider the case where κ is uncountable. Note that $\lambda \leq \kappa$. Hence, by choosing some injective function $\bigcup_{\alpha<\kappa} S_\alpha \to \kappa$ we may assume that $S_\alpha \subseteq \kappa$, for every α. According to Lemma 6.8 (a), the set

$$E := \{ \alpha < \kappa \mid \operatorname{cf} \alpha \geq \lambda \}$$

is stationary. We define a function $f : E \to \kappa$ by

$$f(\alpha) = \sup (S_\alpha \cap \alpha).$$

Note that $\operatorname{cf} \alpha \geq \lambda \geq |S_\alpha|$ implies that

$$f(\alpha) = \sup (S_\alpha \cap \alpha) < \alpha, \quad \text{for all } \alpha \in E.$$

Consequently, we can use the Theorem of Fodor to find a stationary subset $W \subseteq E$ and an ordinal γ such that

$$f(\alpha) = \gamma, \quad \text{for all } \alpha \in W.$$

Since there are at most $|\gamma|^{<\lambda} < \kappa$ sets of the form $S_\alpha \cap \gamma$, we can use Corollary 6.10 to find a stationary subset $W' \subseteq W$ and some set $U \subseteq \gamma$ such that

$$S_\alpha \cap \gamma = U, \quad \text{for all } \alpha \in W'.$$

We construct a strictly increasing sequence $(\xi_\alpha)_{\alpha<\kappa}$ of ordinals $\xi_\alpha \in W'$ as follows. Let ξ_0 be the minimal element of W'. For the inductive step, suppose that we have already defined ξ_α for all $\alpha < \beta$. Then we chose some element $\xi_\beta \in W'$ such that

$$\xi_\beta > \xi_\alpha \quad \text{and} \quad \xi_\beta > \sup S_{\xi_\alpha}, \quad \text{for all } \alpha < \beta.$$

Note that such an element exists since κ is regular.

Having constructed $(\xi_\alpha)_{\alpha<\kappa}$, it follows for $\alpha < \beta < \kappa$ that

$$S_{\xi_\alpha} \cap S_{\xi_\beta} = (S_{\xi_\alpha} \cap \xi_\beta) \cap S_{\xi_\beta} = S_{\xi_\alpha} \cap (S_{\xi_\beta} \cap \gamma) = U.$$

Consequently, the set $I := \{ \xi_\alpha \mid \alpha < \kappa \}$ has the desired properties. □

Exercise 6.2. Let $k, m, n < \omega$ be finite numbers with $n > k!(m-1)^{k+1}$. Prove that, for every family $(S_i)_{i<n}$ of sets of size $|S_i| = k$, there exists a subset $I \subseteq [n]$ of size $|I| = m$ and some set U such that

$$S_i \cap S_j = U, \quad \text{for all distinct } i, j \in I.$$

We conclude this section by proving that every stationary set can be partitioned into κ disjoint stationary subsets. We start with two technical lemmas.

Lemma 6.12. *Let κ be an uncountable regular cardinal and $S \subseteq \kappa$ a stationary set every element of which is an uncountable regular cardinal. Then the set*

$$W := \{ \lambda \in S \mid S \cap \lambda \text{ is not stationary in } \lambda \}$$

is stationary.

Proof. To show that W is stationary, let $C \subseteq \kappa$ be closed unbounded. By Lemma 6.2 (a), the set

$$C' := \{ \alpha \in C \mid C \cap \alpha \text{ is cofinal in } \alpha \}.$$

is closed unbounded. Hence, $S \cap C' \neq \varnothing$. Let λ be the minimal element of $S \cap C'$. Then λ is a regular cardinal and $C \cap \lambda$ is cofinal in λ. Consequently, it follows by Lemma 6.2 (b) that $C \cap \lambda$ is a closed unbounded subset of λ. Hence, Lemma 6.2 (a) implies that $C' \cap \lambda$ is also closed unbounded. Since, by choice of λ, the sets $C' \cap \lambda$ and $S \cap \lambda$ are disjoint, it follows that $S \cap \lambda$ is not stationary. Consequently, $\lambda \in W \cap C$, as desired. □

Lemma 6.13. *Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and, for every $\alpha \in S$, let $\gamma_\alpha : \mathrm{cf}\,\alpha \to \alpha$ be a cofinal and strictly increasing function. If either*

(i) *there is an infinite cardinal λ such that $\mathrm{cf}\,\alpha = \lambda$, for all $\alpha \in S$, or*

(ii) *every $\alpha \in S$ is a regular cardinal, the functions γ_α are continuous, and $S \cap \mathrm{rng}\,\gamma_\alpha = \varnothing$,*

then there exists an ordinal $\beta < \kappa$ such that, for every $\xi < \kappa$, the set

$$U_\xi := \{ \alpha \in S \mid \mathrm{cf}\,\alpha > \beta \text{ and } \gamma_\alpha(\beta) \geq \xi \}$$

is stationary.

Proof. For a contradiction, suppose otherwise. Then we can find, for every $\beta < \kappa$, an ordinal ξ_β and a closed unbounded set C_β such that $U_{\xi_\beta} \cap C_\beta = \varnothing$, that is,

$$\gamma_\alpha(\beta) < \xi_\beta\,, \quad \text{for all } \alpha \in S \cap C_\beta \text{ such that } \mathrm{cf}\,\alpha > \beta\,.$$

In Case (i) we set $\zeta := \sup_{\beta<\lambda} \xi_\beta$ and $D := \bigcap_{\beta<\lambda} C_\beta$. Then $\gamma_\alpha(\beta) < \zeta$, for all $\beta < \lambda$ and $\alpha \in S \cap D$. Choosing $\alpha \in S \cap D$ with $\alpha > \zeta$ it follows that $\sup_{\beta<\lambda} \gamma_\alpha(\beta) \leq \zeta < \alpha$. A contradiction to the cofinality of γ_α.

It remains to consider Case (ii). Let D be the diagonal intersection of $(C_\beta)_{\beta<\kappa}$. Then $\alpha \in S \cap D$ implies that $\alpha \in S \cap C_\beta$, for all $\beta < \alpha$. Hence,

$$\gamma_\alpha(\beta) < \xi_\beta\,, \quad \text{for } \beta < \alpha\,.$$

The set

$$E := \{ \alpha \in D \mid \xi_\beta < \alpha \text{ for all } \beta < \alpha \}$$

is closed unbounded since it can be written as the intersection of D and the diagonal intersection of the sets $\uparrow\xi_\beta$, $\beta < \kappa$, which are clearly closed unbounded. Hence, it follows by Lemma 6.8 (c) that $S \cap E$ is stationary. Let $\delta < \varepsilon$ be two elements of $S \cap E$. Then

$$\beta < \delta \quad \text{implies} \quad \gamma_\varepsilon(\beta) < \xi_\beta < \delta\,,$$

where the first inequality follows since $\varepsilon \in S \cap D$ and the second one follows since $\delta \in E$. By continuity of γ_ε,

$$\gamma_\varepsilon(\delta) = \sup_{\beta<\delta} \gamma_\varepsilon(\beta) \leq \delta\,.$$

Since γ_ε is strictly increasing, it therefore follows by Lemma A3.1.7 that $\gamma_\varepsilon(\delta) = \delta$. But $\delta \in S$ and $\gamma_\varepsilon(\delta) \in \mathrm{rng}\,\gamma_\varepsilon \subseteq \kappa \setminus S$. A contradiction. □

The first step in partitioning a stationary set into κ many stationary subsets consists in finding a decreasing chain of stationary subsets.

Lemma 6.14. *Let κ be an uncountable regular cardinal. For every stationary set $S \subseteq \kappa$, there exists a stationary subset $U \subseteq S$ and a function $f : U \to \kappa$ such that $f(\alpha) < \alpha$, for all $\alpha \in U$, and*

$$f^{-1}[\Uparrow\xi] \text{ is stationary,} \quad \text{for all } \xi < \kappa\,.$$

Proof. Consider the function

$$g : S \setminus \{0\} \to \kappa : \alpha \mapsto \begin{cases} \mathrm{cf}\,\alpha & \text{if } \mathrm{cf}\,\alpha < \alpha\,, \\ 0 & \text{if } \mathrm{cf}\,\alpha = \alpha\,. \end{cases}$$

Then $g(\alpha) < \alpha$, for all $\alpha \in S \setminus \{0\}$, and we can use the Theorem of Fodor to obtain a cardinal $\lambda < \kappa$ such that $T := g^{-1}(\lambda)$ is stationary. We distinguish two cases.

First, suppose that $\lambda > 0$. Note that the set T contains a limit ordinal, as the set of all limit ordinals is closed unbounded. This implies that λ is infinite. Therefore, for every $\alpha \in T$, we can choose by Lemma 4.11, a cofinal, strictly increasing function $\gamma_\alpha : \lambda \to \alpha$. By Lemma 6.13 , there exists an ordinal $\beta < \lambda$ such that, for every $\xi < \kappa$, the set

$$U_\xi := \{ \alpha \in T \mid \gamma_\alpha(\beta) \geq \xi \}$$

is stationary. Hence, we can set $U := T$ and define $f : T \to \kappa$ by

$$f(\alpha) := \gamma_\alpha(\beta)\,.$$

If $\lambda = 0$, the set T consists of regular cardinals and Lemma 6.12 implies that the set

$$W := \{ \alpha \in T \mid T \cap \alpha \text{ is not stationary in } \alpha \}$$

is stationary. For every $\alpha \in W$, we fix a closed unbounded set $C_\alpha \subseteq \alpha$ with $(T \cap \alpha) \cap C_\alpha = \emptyset$. Since C_α is well-ordered, there exists an order-isomorphism $\gamma_\alpha : \beta \to C_\alpha$, for some ordinal β. Note that β cannot be smaller than α, because γ_α is cofinal in α and α is regular. Therefore, $\gamma_\alpha : \alpha \to C_\alpha$. Furthermore, $\sup_{\beta<\delta} \gamma_\alpha(\beta) \in C_\alpha$, for each limit ordinal $\delta < \alpha$, since C_α is closed unbounded. Consequently, $\sup_{\beta<\delta} \gamma_\alpha(\beta)$ is the least element of C_α that is larger than every $\gamma_\alpha(\beta)$ with $\beta < \delta$. As this element is $\gamma_\alpha(\delta)$, we obtain

$$\sup_{\beta<\delta} \gamma_\alpha(\beta) = \gamma_\alpha(\delta) .$$

Hence, each γ_α is a strictly continuous function with $W \cap \operatorname{rng} \gamma_\alpha = \emptyset$. We can therefore use Lemma 6.13 to find an ordinal $\beta < \kappa$ such that, for every $\xi < \kappa$, the set

$$U_\xi := \{ \alpha \in W \mid \alpha > \beta \text{ and } \gamma_\alpha(\beta) \geq \xi \}$$

is stationary. Thus, we can set $U := W \cap {\Uparrow}\beta$ and define $f : U \to \kappa$ by $f(\alpha) := \gamma_\alpha(\beta)$. □

Theorem 6.15 (Solovay). *Let κ be an uncountable regular cardinal. Every stationary set $S \subseteq \kappa$ can be written as a disjoint union of κ stationary subsets of κ.*

Proof. By Lemma 6.14, there exists a stationary subset $U \subseteq S$ and a function $f : U \to \kappa$ such that $f(\alpha) < \alpha$ and the sets $U_\xi := f^{-1}[{\Uparrow}\xi]$ are stationary, for all $\xi < \kappa$. Applying the Theorem of Fodor to each restriction $f \upharpoonright U_\xi$, we obtain ordinals $\alpha_\xi < \kappa$ such that the sets $W_\xi := (f \upharpoonright U_\xi)^{-1}(\alpha_\xi)$ are stationary, for all $\xi < \kappa$. Note that $W_\xi \cap W_\zeta = \emptyset$, if $\alpha_\xi \neq \alpha_\zeta$. Furthermore, $W_\xi \neq \emptyset$ implies that $\alpha_\xi \geq \xi$. Hence, $\sup_{\xi<\kappa} \alpha_\xi = \kappa$ and it follows by regularity of κ that

$$\big|\{ W_\xi \mid \xi < \kappa \}\big| = \big|\{ \alpha_\xi \mid \xi < \kappa \}\big| = \kappa .$$

Thus, we have found a family of κ disjoint stationary subsets of S. Since every superset of a stationary set is also stationary, we can enlarge these subsets to obtain the desired partition of S. □

7. *Conclusion*

With the Axiom of Replacement we have introduced our final axiom. The theory consisting of the six axioms

- Extensionality
- Separation
- Infinity
- Creation
- Choice
- Replacement

is called *Zermelo-Fraenkel set theory*, ZFC for short.

We can classify these axioms into three parts. The Axioms of Extensionality and Creation specify what we mean by a set. They postulate that every set is uniquely determined by its elements and that the membership relation is well-founded. The remaining axioms speak about the existence of certain sets. Infinity and Replacement ensure that the cumulative hierarchy is long enough. There are as many stages as there are ordinals. The Axioms of Separation and Choice on the other hand make the hierarchy wide by ensuring that the power-set operation yields enough subsets. In particular, every definable subset exists and on every set there exists a well-ordering.

Finally, let us note that the usual definition of ZFC is based on a different axiomatisation where the Axiom of Creation is replaced by four other axioms and the Axiom of Infinity is stated in a slightly different way. Nevertheless, we are justified in calling the above theory ZFC since the two variants are equivalent: every model satisfying one of the axiom systems also satisfies the other one, and vice versa.

The following two theorems summarise the results of this section.

Theorem 6.12 (Cohen, Shelah). *Let T be a complete first-order theory. The following conditions are equivalent:*

(1) *T is stable.*

(2) *T has $\mathrm{Un}(\kappa, \lambda)$-representations, for some cardinals κ and λ.*

(3) *T has $\mathrm{Wf}(0, |T|)$-representations.*

(4) *T has $\mathrm{Wf}(|T|, |T|)$-representations.*

Proof. (2) $\Rightarrow$ (1) has been shown in Proposition 6.8 (a), the implications (4) $\Rightarrow$ (3) $\Rightarrow$ (2) follow from Lemmas 6.5 and 6.2, and (1) $\Rightarrow$ (4) follows by Proposition 6.11. □

Theorem 6.13 (Cohen, Shelah). *Let T be a complete first-order theory. The following conditions are equivalent:*

(1) *T is $\aleph_0$-stable.*

(2) *T has $\mathrm{Lf}(\aleph_0, \aleph_0)$-representations.*

Proof. (2) $\Rightarrow$ (1) follows by Proposition 6.8 (b) and (1) $\Rightarrow$ (2) follows by Proposition 6.11. □

Recommended Literature

Set theory

M. D. Potter, *Sets. An Introduction,* Oxford University Press 1990.

A. Lévy, *Basic Set Theory,* Springer 1979, Dover 2002.

K. Kunen, *Set Theory. An Introduction to Independence Proofs,* North-Holland 1983.

T. J. Jech, *Set Theory,* 3rd ed., Springer 2003.

Algebra and Category Theory

G. M. Bergman, *An Invitation to General Algebra and Universal Constructions,* 2nd ed., Springer 2015.

P. M. Cohn, *Universal Algebra,* 2nd ed., Springer 1981.

P. M. Cohn, *Basic Algebra,* Springer 2003.

S. Lang, *Algebra,* 3rd ed., Springer 2002.

F. Borceux, *Handbook of Categorical Algebra,* Cambridge University Press 1994.

S. MacLane, *Categories for the Working Mathematician,* 2nd ed., Springer 1998.

J. Adámek, J. Rosický, and M. Vitale, *Algebraic Theories,* Cambridge University Press 2011.

J. Adámek and J. Rosický, *Locally Presentable and Accessible Categories,* Cambridge University Press 1994.

Topology and lattice theory

R. Engelking, *General Topology,* 2nd ed., Heldermann 1989.

C.-A. Faure, A. Frölicher, *Modern Projective Geometry,* Kluwer 2000.

P. T. Johnstone, *Stone Spaces,* Cambridge University Press 1982.

G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, and D. S. Scott, *Continuous Lattices and Domains,* Cambridge University Press 2003.

Model theory

K. Tent and M. Ziegler, *A Course in Model Theory,* Cambridge University Press 2012.

W. Hodges, *Model Theory,* Cambridge University Press 1993.

B. Poizat, *A Course in Model Theory,* Springer 2000.

C. C. Chang and H. J. Keisler, *Model Theory,* 3rd ed., North-Holland 1990.

General model theory

J. Barwise and S. Feferman, eds., *Model-Theoretic Logics,* Springer 1985.

J. T. Baldwin, *Categoricity,* AMS 2010.

R. Diaconescu, *Institution-Independent Model Theory,* Birkhäuser 2008.

H.-D. Ebbinghaus and J. Flum, *Finite Model Theory,* Springer 1995.

Stability theory

S. Buechler, *Essential Stability Theory,* Springer 1996.

E. Casanovas, *Simple Theories and Hyperimaginaries,* Cambridge University Press 2011.

A. Pillay, *Geometric Stability Theory,* Oxford Science Publications 1996.

F. O. Wagner, *Simple Theories,* Kluwer Academic Publishers 2000.

S. Shelah, *Classification Theory,* 2nd ed., North-Holland 1990.

Symbol Index

Chapter A1

Chapter A2

Symbol Index

Chapter A3

Chapter A4

Chapter B1

Chapter B2

Chapter B3

Chapter B4

Chapter B5

Chapter B6

Chapter C1

Chapter C2

Chapter C3

Chapter C4

Chapter C5

Chapter D1

Chapter D2

Chapter D3

Chapter E2

Chapter E3

Chapter E4

Chapter E5

Chapter E6

Chapter E7

Chapter F1

Chapter F2

Chapter F3

Chapter F4

Chapter F5

Chapter G1

Chapter G2

Index

The Roman and Fraktur alphabets							
A	a	𝔄	𝔞	N	n	𝔑	𝔫
B	b	𝔅	𝔟	O	o	𝔒	𝔬
C	c	ℭ	𝔠	P	p	𝔓	𝔭
D	d	𝔇	𝔡	Q	q	𝔔	𝔮
E	e	𝔈	𝔢	R	r	ℜ	𝔯
F	f	𝔉	𝔣	S	s	𝔖	ſ 𝔰
G	g	𝔊	𝔤	T	t	𝔗	𝔱
H	h	ℌ	𝔥	U	u	𝔘	𝔲
I	i	ℑ	𝔦	V	v	𝔙	𝔳
J	j	𝔍	𝔧	W	w	𝔚	𝔴
K	k	𝔎	𝔨	X	x	𝔛	𝔵
L	l	𝔏	𝔩	Y	y	𝔜	𝔶
M	m	𝔐	𝔪	Z	z	ℨ	𝔷

The Greek alphabet					
Α	α	alpha	Ν	ν	nu
Β	β	beta	Ξ	ξ	xi
Γ	γ	gamma	Ο	ο	omicron
Δ	δ	delta	Π	π	pi
Ε	ε	epsilon	Ρ	ρ	rho
Ζ	ζ	zeta	Σ	σ	sigma
Η	η	eta	Τ	τ	tau
Θ	ϑ	theta	Υ	υ	upsilon
Ι	ι	iota	Φ	ϕ	phi
Κ	κ	kappa	Χ	χ	chi
Λ	λ	lambda	Ψ	ψ	psi
Μ	μ	mu	Ω	ω	omega