

Set Theory

2016/17

BY PETER KOEPKE

*Die Mengenlehre ist das Fundament
der gesamten Mathematik
(FELIX HAUSDORFF,
Grundzüge der Mengenlehre, 1914)*

1 Introduction

1.1 The origin of set theory

GEORG CANTOR characterized sets as follows:

Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten, wohlunterschiedenen Objekten m unsrer Anschauung oder unseres Denkens (welche die "Elemente" von M genannt werden) zu einem Ganzen.

FELIX HAUSDORFF in *Grundzüge* formulated shorter:

Eine Menge ist eine Zusammenfassung von Dingen zu einem Ganzen, d.h. zu einem neuen Ding.

Sets are ubiquitous in mathematics. According to HAUSDORFF

Differential- und Integralrechnung, Analysis und Geometrie arbeiten in Wirklichkeit, wenn auch vielleicht in verschleiender Ausdrucksweise, beständig mit unendlichen Mengen.

1.2 Set theoretic foundations of mathematics

In current mathematics, *many* notions are explicitly defined using sets. The following example indicates that notions which are not set-theoretical *prima facie* can be construed set-theoretically:

f is a real funktion $\equiv f$ is a **set** of ordered pairs $(x, f(x))$ of real numbers, such that ... ;

(x, y) is an ordered pair $\equiv (x, y)$ is a **set** $\dots\{x, y\}\dots$;

x is a real number $\equiv x$ is a left half of a DEDEKIND cut in $\mathbb{Q} \equiv x$ is a **subset** of \mathbb{Q} , such that ... ;

r is a rational number $\equiv r$ is an **ordered pair** of integers, such that ... ;

z is an integer $\equiv z$ is an **ordered pair** of natural numbers (= non-negative integers);

$\mathbb{N} = \{0, 1, 2, \dots\}$;

0 is the empty **set**;

1 is the **set** $\{0\}$;

2 is the **set** $\{0, 1\}$; etc. etc.

We shall see that *all* mathematical notions can be reduced to the notion of *set*.

Besides this foundational role, set theory is also the mathematical study of the *infinite*. There are infinite sets like \mathbb{N} , \mathbb{Q} , \mathbb{R} which can be subjected to the constructions and analyses of set theory; there are various degrees of infinity which lead to a rich theory of infinitary combinatorics.

In this course, we shall first apply set theory to obtain the standard foundation of mathematics and then turn towards “pure” set theory.

1.3 The language of set theory

If m is an *element* of M one writes $m \in M$. If all mathematical objects are reducible to sets, *both sides* of these relation have to be sets. This means that set theory studies the \in -relation $m \in M$ for arbitrary *sets* m and M . As it turns out, this is sufficient for the purposes of set theory and mathematics. In set theory variables range over the class of all sets, the \in -relation is the only undefined structural component, every other notion will be defined from the \in -relation. Basically, set theoretical statement will thus be of the form

$$\dots \forall x \dots \exists y \dots x \in y \dots u \equiv v \dots,$$

belonging to the first-order predicate language with the only given predicate \in .

To deal with the complexities of set theory and mathematics one develops a comprehensive and intuitive language of abbreviations and definitions which, eventually, allows to write familiar statements like

$$e^{i\pi} = -1$$

and to view them as statements within set theory.

The language of set theory may be seen as a low-level, internal language. The language of mathematics possesses high-level “macro” expressions which abbreviate low-level statements in an efficient and intuitive way.

1.4 Russell’s paradox

CANTOR’s naive description of the notion of set suggests that for any mathematical statement $\varphi(x)$ in one free variable x there is a *set* y such that

$$x \in y \leftrightarrow \varphi(x),$$

i.e., y is the collection of all sets x which satisfy φ .

This axiom is a basic principle in GOTTLOB FREGE’s *Grundgesetze der Arithmetik*, 1893, Grundgesetz V, Grundgesetz der Wertverläufe.

BERTRAND RUSSELL noted in 1902 that setting $\varphi(x)$ to be $x \notin x$ this becomes

$$x \in y \leftrightarrow x \notin x,$$

and in particular for $x = y$:

$$y \in y \leftrightarrow y \notin y.$$

Contradiction.

This contradiction is usually called RUSSELL’s paradox, antinomy, contradiction. It was also discovered slightly earlier by ERNST ZERMELO. The paradox shows that the formation of sets as collections of sets by *arbitrary* formulas is not consistent.

2 The ZERMELO-FRAENKEL Axioms

The difficulties around RUSSELL’s paradox and also around the axiom of choice lead ZERMELO to the formulation of axioms for set theory in the spirit of the axiomatics of DAVID HILBERT of whom ZERMELO was an assistant at the time.

ZERMELO’s main idea was to restrict FREGE’s Axiom V to formulas which correspond to mathematically important formations of collections, but to avoid arbitrary formulas which can lead to paradoxes like the one exhibited by RUSSELL.

The original axiom system of ZERMELO was extended and detailed by ABRAHAM FRAENKEL (1922), DMITRY MIRIMANOFF (1917/20), and THORALF SKOLEM.

We shall discuss the axioms one by one and simultaneously introduce the logical language and useful conventions.

2.1 Set Existence

The *set existence axiom*

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

like all axioms, is expressed in a language with quantifiers \exists (“there exists”) and \forall (“for all”), which is familiar from the ϵ - δ -statements in analysis. The *language of set theory* uses variables x, y, \dots which may satisfy the binary relations \in or $=$: $x \in y$ (“ x is an *element of* y ”) or $x = y$. These elementary *formulas* may be connected by the *propositional connectives* \wedge (“and”), \vee (“or”), \rightarrow (“implies”), \leftrightarrow (“is equivalent”), and \neg (“not”). The use of this language will be demonstrated by the subsequent axioms.

The axiom expresses the existence of a set which has no elements, i.e., the existence of the *empty set*.

2.2 Extensionality

The *axiom of extensionality*

$$\forall x \forall x' (\forall y (y \in x \leftrightarrow y \in x') \rightarrow x = x')$$

expresses that a set is exactly determined by the collection of its elements. This allows to prove that there is exactly one empty set.

Lemma 1. $\forall x \forall x' (\forall y \neg y \in x \wedge \forall y \neg y \in x' \rightarrow x = x')$.

Proof. Consider x, x' such that $\forall y \neg y \in x \wedge \forall y \neg y \in x'$. Consider y . Then $\neg y \in x$ and $\neg y \in x'$. This implies $\forall y (y \in x \leftrightarrow y \in x')$. The axiom of extensionality implies $x = x'$. \square

Note that this proof is a usual mathematical argument, and it is also a *formal proof* in the sense of mathematical logic. The sentences of the proof can be derived from earlier ones by purely formal deduction rules. The rules of natural deduction correspond to common sense figures of argumentation which treat hypothetical objects as if they would concretely exist.

2.3 Pairing

The *pairing axiom*

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

postulates that for all sets x, y there is set z which may be denoted as

$$z = \{x, y\}.$$

This formula, including the new notation, is equivalent to the formula

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

In the sequel we shall extend the small language of set theory by hundreds of symbols and conventions, in order to get to the ordinary language of mathematics with notations like

$$\mathbb{N}, \mathbb{R}, \sqrt{385}, \pi, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \int_a^b f'(x) dx = f(b) - f(a), \text{ etc.}$$

Such notations are chosen for intuitive, pragmatic, or historical reasons.

Using the notation for unordered pairs, the pairing axiom may be written as

$$\forall x \forall y \exists z z = \{x, y\}.$$

By the axiom of extensionality, the term-like notation has the expected behaviour. E.g.:

Lemma 2. $\forall x \forall y \forall z \forall z' (z = \{x, y\} \wedge z' = \{x, y\} \rightarrow z = z')$.

Proof. Exercise. □

Note that we implicitly use several notational conventions: variables have to be chosen in a reasonable way, for example the symbols z and z' in the lemma have to be taken different and different from x and y . We also assume some operator priorities to reduce the number of brackets: we let \wedge bind stronger than \vee , and \vee stronger than \rightarrow and \leftrightarrow .

We used the “term” $\{x, y\}$ to occur within set theoretical formulas. This abbreviation is than to be expanded in a natural way, so that officially all mathematical formulas are formulas in the “pure” \in -language. We want to see the notation $\{x, y\}$ as an example of a *class term*. We define uniform notations and convention for such abbreviation terms.

2.4 Class Terms

The extended language of set theory contains class terms and notations for them. There are axioms for class terms that fix how extended formulas can be reduced to formulas in the unextended \in -language of set theory.

Definition 3. A class term is of the form $\{x|\varphi\}$ where x is a variable and $\varphi \in L^\in$. The usage of these class terms is defined recursively by the following axioms: If $\{x|\varphi\}$ and $\{y|\psi\}$ are class terms then

- $u \in \{x|\varphi\} \leftrightarrow \varphi_x^u$, where φ_x^u is obtained from φ by (reasonably) substituting the variable x by the variable u ;
- $u = \{x|\varphi\} \leftrightarrow \forall v (v \in u \leftrightarrow \varphi_x^v)$;
- $\{x|\varphi\} = u \leftrightarrow \forall v (\varphi_x^v \leftrightarrow v \in u)$;
- $\{x|\varphi\} = \{y|\psi\} \leftrightarrow \forall v (\varphi_x^v \leftrightarrow \psi_y^v)$;
- $\{x|\varphi\} \in u \leftrightarrow \exists v (v \in u \wedge v = \{x|\varphi\})$;
- $\{x|\varphi\} \in \{y|\psi\} \leftrightarrow \exists v (\psi_y^v \wedge v = \{x|\varphi\})$.

A term is either a variable or a class term.

Definition 4.

- a) $\emptyset := \{x|x \neq x\}$ is the empty set;
- b) $V := \{x|x = x\}$ is the universe (of all sets);
- c) $\{x, y\} := \{u|u = x \vee u = y\}$ is the unordered pair of x and y .

Lemma 5.

- a) $\emptyset \in V$.
- b) $\forall x, y \{x, y\} \in V$.

Proof. a) By the axioms for the reduction of abstraction terms, $\emptyset \in V$ is equivalent to the following formulas

$$\begin{aligned} & \exists v (v = v \wedge v = \emptyset) \\ & \exists v v = \emptyset \\ & \exists v \forall w (w \in v \leftrightarrow w \neq w) \\ & \exists v \forall w w \notin v \end{aligned}$$

which is equivalent to the axiom of set existence. So $\emptyset \in V$ is another way to write the axiom of set existence.

b) $\forall x, y \{x, y\} \in V$ abbreviates the formula

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

This can be expanded equivalently to the pairing axiom

$$\forall x, y \exists z \forall u (u \in z \leftrightarrow u = x \vee u = y). \quad \square$$

So a) and b) are concise equivalent formulations of the axiom Ex and Pair.

We also introduce *bounded quantifiers* to simplify notation.

Definition 6. Let A be a term. Then $\forall x \in A \varphi \leftrightarrow \forall x (x \in A \rightarrow \varphi)$ and $\exists x \in A \varphi \leftrightarrow \exists x (x \in A \wedge \varphi)$.

Definition 7. Let x, y, z, \dots be variables and X, Y, Z, \dots be class terms. Define

- a) $X \subseteq Y \leftrightarrow \forall x \in X x \in Y$, X is a subclass of Y ;
- b) $X \cup Y := \{x \mid x \in X \vee x \in Y\}$ is the union of X and Y ;
- c) $X \cap Y := \{x \mid x \in X \wedge x \in Y\}$ is the intersection of X and Y ;
- d) $X \setminus Y := \{x \mid x \in X \wedge x \notin Y\}$ is the difference of X and Y ;
- e) $\bigcup X := \{x \mid \exists y \in X x \in y\}$ is the union of X ;
- f) $\bigcap X := \{x \mid \forall y \in X x \in y\}$ is the intersection of X ;
- g) $\mathcal{P}(X) := \{x \mid x \subseteq X\}$ is the power class of X ;
- h) $\{X\} := \{x \mid x = X\}$ is the singleton set of X ;
- i) $\{X, Y\} := \{x \mid x = X \vee x = Y\}$ is the (unordered) pair of X and Y ;
- j) $\{X_0, \dots, X_{n-1}\} := \{x \mid x = X_0 \vee \dots \vee x = X_{n-1}\}$.

One can prove the well-known boolean properties for these operations. We only give a few examples.

Proposition 8. $X \subseteq Y \wedge Y \subseteq X \rightarrow X = Y$.

Proposition 9. $\bigcup \{x, y\} = x \cup y$.

Proof. We show the equality by two inclusions:

(\subseteq). Let $u \in \bigcup \{x, y\}$. $\exists v (v \in \{x, y\} \wedge u \in v)$. Let $v \in \{x, y\} \wedge u \in v$. ($v = x \vee v = y$) $\wedge u \in v$.

Case 1. $v = x$. Then $u \in x$. $u \in x \vee u \in y$. Hence $u \in x \cup y$.

Case 2. $v = y$. Then $u \in y$. $u \in x \vee u \in y$. Hence $u \in x \cup y$.

Conversely let $u \in x \cup y$. $u \in x \vee u \in y$.

Case 1. $u \in x$. Then $x \in \{x, y\} \wedge u \in x$. $\exists v (v \in \{x, y\} \wedge u \in v)$ and $u \in \bigcup \{x, y\}$.

Case 2. $u \in y$. Then $y \in \{x, y\} \wedge u \in y$. $\exists v (v \in \{x, y\} \wedge u \in v)$ and $u \in \bigcup \{x, y\}$. □

Exercise 1. Show: a) $\bigcup V = V$. b) $\bigcap V = \emptyset$. c) $\bigcup \emptyset = \emptyset$. d) $\bigcap \emptyset = V$.

2.5 Ordered Pairs

Combining objects into ordered pairs (x, y) is taken as an undefined fundamental operation of mathematics. We cannot use the unordered pair $\{x, y\}$ for this purpose, since it does not respect the order of entries:

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

We have to introduce some asymmetry between x and y to make them distinguishable. Following KURATOWSKI and WIENER we define:

Definition 10. $(x, y) := \{\{x\}, \{x, y\}\}$ is the ordered pair of x and y .

The definition involves substituting class terms within class terms. We shall see in the following how these class terms are eliminated to yield pure \in -formulas.

Lemma 11. $\forall x \forall y \exists z z = (x, y)$.

Proof. Consider sets x and y . By the pairing axiom choose u and v such that $u = \{x\}$ and $v = \{x, y\}$. Again by pairing choose z such that $z = \{u, v\}$. We argue that $z = (x, y)$. Note that $(x, y) = \{\{x\}, \{x, y\}\} = \{w \mid w = \{x\} \vee w = \{x, y\}\}$.

Then $z = (x, y)$ is equivalent to

$$\forall w (w \in z \leftrightarrow w = \{x\} \vee w = \{x, y\}),$$

$$\forall w (w = u \vee w = v \leftrightarrow (w = \{x\} \vee w = \{x, y\})),$$

and this is true by the choice of u and v . □

The KURATOWSKI-pair satisfies the fundamental property of ordered pairs:

Lemma 12. $(x, y) = (x', y') \rightarrow x = x' \wedge y = y'$.

Proof. Assume $(x, y) = (x', y')$, i.e.,

$$(1) \{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}.$$

Case 1. $x = y$. Then

$$\{x\} = \{x, y\},$$

$$\{\{x\}, \{x, y\}\} = \{\{x\}, \{x\}\} = \{\{x\}\},$$

$$\{\{x\}\} = \{\{x'\}, \{x', y'\}\},$$

$$\{x\} = \{x'\} \text{ and } x = x',$$

$$\{x\} = \{x', y'\} \text{ and } y' = x.$$

Hence $x = x'$ and $y = x = y'$ as required.

Case 2. $x \neq y$. (1) implies

$$\{x'\} = \{x\} \text{ or } \{x'\} = \{x, y\}.$$

The right-hand side would imply $x = x' = y$, contradicting the case assumption. Hence

$$\{x'\} = \{x\} \text{ and } x' = x.$$

Then (1) implies

$$\{x, y\} = \{x', y'\} = \{x, y'\} \text{ and } y = y'. \quad \square$$

Exercise 2.

a) Show that $\langle x, y \rangle := \{\{x, \emptyset\}, \{y, \{\emptyset\}\}\}$ also satisfies the fundamental property of ordered pairs (F. HAUSDORFF).

b) Can $\{x, \{y, \emptyset\}\}$ be used as an ordered pair?

Exercise 3. Give a set-theoretical formalization of an ordered-triple operation.

2.6 Relations and Functions

Ordered pairs allow to introduce *relations* and *functions* in the usual way. One has to distinguish between *sets* which are relations and functions, and *class terms* which are relations and functions.

Definition 13. A term R is a relation if all elements of R are ordered pairs, i.e., $R \subseteq V \times V$. Also write Rxy or xRy instead of $(x, y) \in R$. If A is a term and $R \subseteq A \times A$ then R is a relation on A .

Note that this definition is really an *infinite schema* of definitions, with instances for all terms R and A . The subsequent extensions of our language are also infinite definition schemas. We extend the term language by parametrized collections of terms.

Definition 14. Let $t(\vec{x})$ be a term in the variables \vec{x} and let φ be an \in -formula. Then $\{t(\vec{x}) \mid \varphi\}$ stands for $\{z \mid \exists \vec{x} (\varphi \wedge z = t(\vec{x}))\}$.

Definition 15. Let R, S, A be terms.

- a) The domain of R is $\text{dom}(R) := \{x \mid \exists y x R y\}$.
- b) The range of R is $\text{ran}(R) := \{y \mid \exists x x R y\}$.
- c) The field of R is $\text{field}(R) := \text{dom}(R) \cup \text{ran}(R)$.
- d) The restriction of R to A is $R \upharpoonright A := \{(x, y) \mid x R y \wedge x \in A\}$.
- e) The image of A under R is $R[A] := R''A := \{y \mid \exists x \in A x R y\}$.
- f) The preimage of A under R is $R^{-1}[A] := \{x \mid \exists y \in A x R y\}$.
- g) The composition of S and R (" S after R ") is $S \circ R := \{(x, z) \mid \exists y (x R y \wedge y S z)\}$.
- h) The inverse of R is $R^{-1} := \{(y, x) \mid x R y\}$.

Relations can play different roles in mathematics.

Definition 16. Let R be a relation.

- a) R is reflexive iff $\forall x \in \text{field}(R) x R x$.
- b) R is irreflexive iff $\forall x \in \text{field}(R) \neg x R x$.
- c) R is symmetric iff $\forall x, y (x R y \rightarrow y R x)$.
- d) R is antisymmetric iff $\forall x, y (x R y \wedge y R x \rightarrow x = y)$.
- e) R is transitive iff $\forall x, y, z (x R y \wedge y R z \rightarrow x R z)$.
- f) R is connex iff $\forall x, y \in \text{field}(R) (x R y \vee y R x \vee x = y)$.
- g) R is an equivalence relation iff R is reflexive, symmetric and transitive.
- h) Let R be an equivalence relation. Then $[x]_R := \{y \mid y R x\}$ is the equivalence class of x modulo R .

It is possible that an equivalence class $[x]_R$ is not a set: $[x]_R \notin V$. Then the formation of the collection of all equivalence classes modulo R may lead to contradictions. Another important family of relations is given by *order relations*.

Definition 17. Let R be a relation.

- a) R is a partial order iff R is reflexive, transitive and antisymmetric.
- b) R is a linear order iff R is a connex partial order.
- c) Let A be a term. Then R is a partial order on A iff R is a partial order and $\text{field}(R) = A$.
- d) R is a strict partial order iff R is transitive and irreflexive.
- e) R is a strict linear order iff R is a connex strict partial order.

Partial orders are often denoted by symbols like \leq , and strict partial orders by $<$. A common notation in the context of (strict) partial orders R is to write

$$\exists p R q \varphi \text{ and } \forall p R q \varphi \text{ for } \exists p (p R q \wedge \varphi) \text{ and } \forall p (p R q \rightarrow \varphi) \text{ resp.}$$

One of the most important notions in mathematics is that of a *function*.

Definition 18. Let F be a term. Then F is a function if it is a relation which satisfies

$$\forall x, y, y' (x F y \wedge x F y' \rightarrow y = y').$$

If F is a function then

$$F(x) := \{u \mid \forall y (x F y \rightarrow u \in y)\}$$

is the value of F at x .

If F is a function and xFy then $y = F(x)$. If there is no y such that xFy then

$$F(x) = \bigcap_{xFy} y = \bigcap \emptyset = V.$$

The “value” V at x may be read as “undefined”. A function can also be considered as the (indexed) sequence of its values, and we also write

$$(F(x))_{x \in A} \text{ or } (F_x)_{x \in A} \text{ instead of } F: A \rightarrow V.$$

We define further notions associated with functions.

Definition 19. Let F, A, B be terms.

- a) F is a function from A to B , or $F: A \rightarrow B$, iff F is a function, $\text{dom}(F) = A$, and $\text{range}(F) \subseteq B$.
- b) F is a partial function from A to B , or $F: A \dashrightarrow B$, iff F is a function, $\text{dom}(F) \subseteq A$, and $\text{range}(F) \subseteq B$.
- c) F is a surjective function from A to B iff $F: A \rightarrow B$ and $\text{range}(F) = B$.
- d) F is an injective function from A to B iff $F: A \rightarrow B$ and

$$\forall x, x' \in A (x \neq x' \rightarrow F(x) \neq F(x'))$$

- e) F is a bijective function from A to B , or $F: A \leftrightarrow B$, iff $F: A \rightarrow B$ is surjective and injective.
- f) ${}^A B := \{f \mid f: A \rightarrow B\}$ is the class of all functions from A to B .

One can check that these functional notions are consistent and agree with common usage:

Exercise 4. Define a relation \sim on V by

$$x \sim y \leftrightarrow \exists f f: x \leftrightarrow y.$$

One says that x and y are *equinumerous* or *equipollent*. Show that \sim is an equivalence relation on V . What is the equivalence class of \emptyset ? What is the equivalence class of $\{\emptyset\}$?

Exercise 5. Consider functions $F: A \rightarrow B$ and $F': A \rightarrow B$. Show that

$$F = F' \text{ iff } \forall a \in A F(a) = F'(a).$$

2.7 Unions

The *union axiom* reads

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

Lemma 20. The union axiom is equivalent to $\forall x \bigcup x \in V$.

Proof. Observe the following equivalences:

$$\forall x \bigcup x \in V$$

$$\leftrightarrow \forall x \exists y (y = \bigcup x)$$

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

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

which is equivalent to the union axiom. □

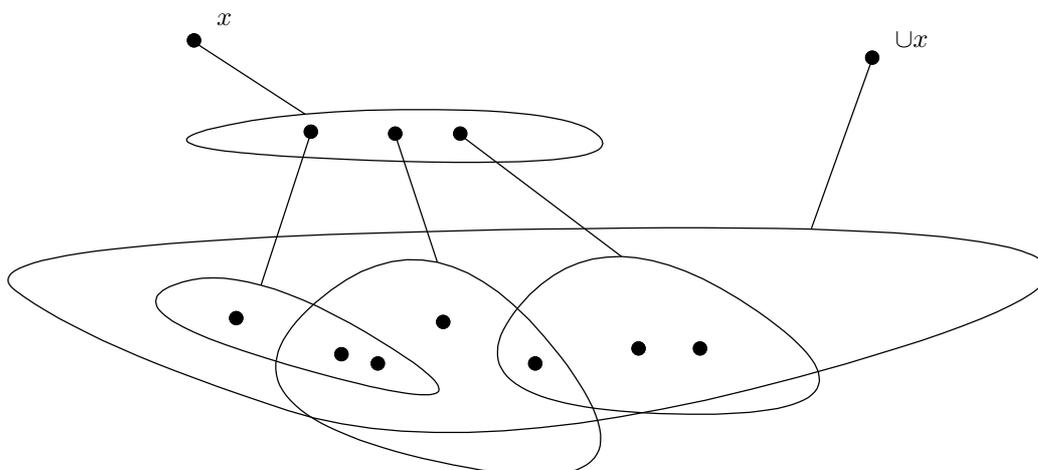
Note that the union of x is usually viewed as the union of all *elements* of x :

$$\bigcup x = \bigcup_{w \in x} w,$$

where we define

$$\bigcup_{a \in A} t(a) = \{z \mid \exists a \in A z \in t(a)\}.$$

Graphically $\bigcup x$ can be illustrated like this:



Combining the axioms of pairing and unions we obtain:

Lemma 21. $\forall x_0, \dots, x_{n-1} \{x_0, \dots, x_{n-1}\} \in V$.

Note that this is a *schema* of lemmas, one for each ordinary natural number n . We prove the schema by complete induction on n .

Proof. For $n = 0, 1, 2$ the lemma states that $\emptyset \in V$, $\forall x \{x\} \in V$, and $\forall x, y \{x, y\} \in V$ resp., and these are true by previous axioms and lemmas. For the induction step assume that the lemma holds for n , $n \geq 1$. Consider sets x_0, \dots, x_n . Then

$$\{x_0, \dots, x_n\} = \{x_0, \dots, x_{n-1}\} \cup \{x_n\}.$$

The right-hand side exists in V by the inductive hypothesis and the union axiom. \square

Remark 22. We are developing the axiom systems ZF and ZFC. These will be infinite schemas, lists, or *sets* of formulas. These schemas are formulated in the common mathematical language, which is able to speak about formulas, in particular \in -formulas, and is also able to speak about infinite collections of formulas. If we assume infinitely many axioms, it should also be possible to conclude infinitely many consequences, like the above Lemma: $\forall x_0, \dots, x_{n-1} \{x_0, \dots, x_{n-1}\} \in V$. We view the common mathematical language as a *meta language* which is able to speak about an *object language* like the language of set theory. The meta language has common mathematical tools available. For example induction and recursion on the common natural numbers, to perform the recursion in the *previous schema of lemmas*. We shall approach the problem of meta theory versus object theory in an informal naive way.

2.8 Separation

It is common to form a subset of a given set consisting of all elements which satisfy some condition. This is codified by the *separation schema*. For every \in -formula $\varphi(z, x_1, \dots, x_n)$ postulate:

$$\forall x_1 \dots \forall x_n \forall x \exists y \forall z (z \in y \leftrightarrow z \in x \wedge \varphi(z, x_1, \dots, x_n)).$$

Using class terms the schema can be reformulated as: for every term A postulate

$$\forall x A \cap x \in V.$$

The crucial point is the restriction to the given set x . The unrestricted, FREGEAN version $A \in V$ for every term A leads to the RUSSELL antinomy. We turn the antinomy into a consequence of the separation schema:

Theorem 23. $V \notin V$.

Proof. Assume that $V \in V$. Then $\exists x x = V$. Take x such that $x = V$. Let R be the RUSSELLIAN class:

$$R := \{x \mid x \notin x\}.$$

By separation, $y := R \cap x \in V$. Note that $R \cap x = R \cap V = R$. Then

$$y \in y \leftrightarrow y \in R \leftrightarrow y \notin y,$$

contradiction. □

This simple but crucial theorem leads to the distinction:

Definition 24. Let A be a term. Then A is a proper class iff $A \notin V$.

Set theory deals with sets and proper classes. Sets are the favoured objects of set theory, the axioms mainly state favorable properties of sets and set existence. Sometimes one says that a term A *exists* if $A \in V$. The intention of set theory is to construe important mathematical classes like the collection of natural and real numbers as sets so that they can be treated set-theoretically. ZERMELO observed that this is possible by requiring some set existences together with the *restricted* separation principle.

Exercise 6. Show that the class $\{\{x\} \mid x \in V\}$ of *singletons* is a proper class.

2.9 Power Sets

The *power set axiom* in class term notation is

$$\forall x \mathcal{P}(x) \in V.$$

The power set axiom yields the existence of function spaces.

Definition 25. Let A, B be terms. Then

$$A \times B := \{(a, b) \mid a \in A \wedge b \in B\}$$

is the cartesian product of A and B .

Exercise 7.

By the specific implementation of KURATOWSKI ordered pairs:

Lemma 26. $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B))$.

Proof. Let $(a, b) \in A \times B$. Then

$$\begin{aligned} a, b &\in A \cup B \\ \{a\}, \{a, b\} &\subseteq A \cup B \\ \{a\}, \{a, b\} &\in \mathcal{P}(A \cup B) \\ (a, b) = \{\{a\}, \{a, b\}\} &\subseteq \mathcal{P}(A \cup B) \\ (a, b) = \{\{a\}, \{a, b\}\} &\in \mathcal{P}(\mathcal{P}(A \cup B)) \end{aligned}$$

□

Theorem 27.

$$a) \quad \forall x, y \quad x \times y \in V.$$

$$b) \quad \forall x, y \quad {}^x y \in V.$$

Proof. Let x, y be sets. a) Using the axioms of pairing, union, and power sets, $\mathcal{P}(\mathcal{P}(x \cup y)) \in V$. By the previous lemma and the axiom schema of separation,

$$x \times y = (x \times y) \cap \mathcal{P}(\mathcal{P}(x \cup y)) \in V.$$

b) ${}^x y \subseteq \mathcal{P}(x \times y)$ since a function $f: x \rightarrow y$ is a subset of $x \times y$. By the separation schema,

$${}^x y = {}^x y \cap \mathcal{P}(x \times y) \in V. \quad \square$$

Note that to “find” the sets in this theorem one has to apply the power set operation repeatedly. We shall see that the universe of all sets can be obtained by iterating the power set operation. The power set axiom leads to higher *cardinalities*. The theory of cardinalities will be developed later, but we can already prove CANTOR’s theorem:

Theorem 28. *Let $x \in V$.*

- a) *There is an injective map $f: x \rightarrow \mathcal{P}(x)$.*
- b) *There does not exist an injective map $g: \mathcal{P}(x) \rightarrow x$.*

Proof. a) Define the map $f: x \rightarrow \mathcal{P}(x)$ by $u \mapsto \{u\}$. This is a set since

$$f = \{(u, \{u\}) \mid u \in x\} \subseteq x \times \mathcal{P}(x) \in V.$$

f is injective: let $u, u' \in x$, $u \neq u'$. By extensionality,

$$f(u) = \{u\} \neq \{u'\} = f(u').$$

b) Assume there were an injective map $g: \mathcal{P}(x) \rightarrow x$. Define the CANTOREAN set

$$c = \{u \mid u \in x \wedge u \notin g^{-1}(u)\} \in \mathcal{P}(x)$$

similar to the class R in RUSSELL’s paradox.

Let $u_0 = g(c)$. Then $g^{-1}(u_0) = c$ and

$$u_0 \in c \leftrightarrow u_0 \notin g^{-1}(u_0) = c.$$

Contradiction. □

2.10 Replacement

If every element of a set is definably *replaced* by another set, the result is a set again. The *schema of replacement* postulates for every term F :

$$F \text{ is a function } \rightarrow \forall x F[x] \in V.$$

Lemma 29. *The replacement schema implies the separation schema.*

Proof. Let A be a term and $x \in V$.

Case 1. $A \cap x = \emptyset$. Then $A \cap x \in V$ by the axiom of set existence.

Case 2. $A \cap x \neq \emptyset$. Take $u_0 \in A \cap x$. Define a map $F: x \rightarrow x$ by

$$F(u) = \begin{cases} u, & \text{if } u \in A \cap x \\ u_0, & \text{else} \end{cases}$$

Then by replacement

$$A \cap x = F[x] \in V$$

as required. □

2.11 Infinity

All the axioms so far can be realized in a domain of finite sets, see exercise 12. The true power of set theory is set free by postulating the existence of *one* infinite set and continuing to assume the axioms. The *axiom of infinity* expresses that the set of “natural numbers” exists. To this end, some “number-theoretic” notions are defined.

Definition 30.

- a) $0 := \emptyset$ is the number zero.
 b) For any term t , $t + 1 := t \cup \{t\}$ is the successor of t .

These notions are reasonable in the later formalization of the natural numbers. The axiom of infinity postulates the existence of a set which contains 0 and is closed under successors

$$\exists x (0 \in x \wedge \forall n \in x \ n + 1 \in x).$$

Intuitively this says that there is a set which contains all natural numbers. Let us define set-theoretic analogues of the standard natural numbers:

Definition 31. *Define*

- a) $1 := 0 + 1$;
 b) $2 := 1 + 1$;
 c) $3 := 2 + 1$; ...

From the context it will be clear, whether “3”, say, is meant to be the standard number “three” or the set theoretical object

$$\begin{aligned} 3 &= 2 \cup \{2\} \\ &= (1 + 1) \cup \{1 + 1\} \\ &= (\{\emptyset\} \cup \{\{\emptyset\}\}) \cup \{\{\emptyset\} \cup \{\{\emptyset\}\}\} \\ &= \{\emptyset, \{\emptyset\}, \{\emptyset\} \cup \{\{\emptyset\}\}\}. \end{aligned}$$

The set-theoretic axioms will ensure that this interpretation of “three” has the important number-theoretic properties of “three”.

2.12 Foundation

The *axiom schema of foundation* provides structural information about the set theoretic universe V . It can be reformulated by postulating, for any term A :

$$A \neq \emptyset \rightarrow \exists x \in A \ A \cap x = \emptyset.$$

Viewing \in as some kind of order relation this means that every non-empty class has an \in -minimal element $x \in A$ such that the \in -predecessors of x are not in A . Foundation excludes circles in the \in -relation:

Lemma 32. *Let n be a natural number ≥ 1 . Then there are no x_0, \dots, x_{n-1} such that*

$$x_0 \in x_1 \in \dots \in x_{n-1} \in x_0.$$

Proof. Assume not and let $x_0 \in x_1 \in \dots \in x_{n-1} \in x_0$. Let

$$A = \{x_0, \dots, x_{n-1}\}.$$

$A \neq \emptyset$ since $n \geq 1$. By foundation take $x \in A$ such that $A \cap x = \emptyset$.

Case 1. $x = x_0$. Then $x_{n-1} \in A \cap x = \emptyset$, contradiction.

Case 2. $x = x_i$, $i > 0$. Then $x_{i-1} \in A \cap x = \emptyset$, contradiction. □

Exercise 8. Show that $x \neq x + 1$.

Exercise 9. Show that the successor function $x \mapsto x + 1$ is injective.

Exercise 10. Show that the term $\{x, \{x, y\}\}$ may be taken as an ordered pair of x and y .

Theorem 33. *The foundation scheme is equivalent to the following, PEANO-type, induction scheme: for every term B postulate*

$$\forall x (x \subseteq B \rightarrow x \in B) \rightarrow B = V.$$

This says that if a “property” B is inherited by x if all elements of x have the property B , then every set has the property B .

Proof. (\rightarrow) Assume B were a term which did not satisfy the induction principle:

$$\forall x (x \subseteq B \rightarrow x \in B) \text{ and } B \neq V.$$

Set $A = V \setminus B \neq \emptyset$. By foundation take $x \in A$ such that $A \cap x = \emptyset$. Then

$$u \in x \rightarrow u \notin A \rightarrow u \in B,$$

i.e., $x \subseteq B$. By assumption, B is inherited by x : $x \in B$. But then $x \notin A$, contradiction.

(\leftarrow) Assume A were a term which did not satisfy the foundation scheme:

$$A \neq \emptyset \text{ and } \forall x \in A A \cap x \neq \emptyset.$$

Set $B = V \setminus A$. Consider $x \subseteq B$. Then $A \cap x = \emptyset$. By assumption, $x \notin A$ and $x \in B$. Thus $\forall x (x \subseteq B \rightarrow x \in B)$. The induction principle implies that $B = V$. Then $A = \emptyset$, contradiction. \square

This proof shows, that the induction principle is basically an equivalent formulation of the foundation principle. The \in -relation is taken as some binary relation without reference to specific properties of this relation. This leads to:

Exercise 11. A relation R on a domain D is called *wellfounded*, iff for all terms A

$$\emptyset \neq A \wedge A \subseteq D \rightarrow \exists x \in A A \cap \{y \mid yRx\} = \emptyset.$$

Formulate and prove a principle for R -induction on D which corresponds to the assumption that R is wellfounded on D .

2.13 Set Theoretic Axiom Schemas

Note that the axiom system introduced is an infinite informal *set* of axioms. It seems unavoidable that we have to go back to some previously given set notions to be able to define the collection of set theoretical axioms - another example of the frequent circularity in foundational theories.

Definition 34. The system ZF of the ZERMELO-FRAENKEL axioms of set theory consists of the following axioms:

a) The set existence axiom (Ex):

$$\exists x \forall y \neg y \in x$$

- there is a set without elements, the empty set.

b) The axiom of extensionality (Ext):

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

- a set is determined by its elements, sets having the same elements are identical.

c) The pairing axiom (Pair):

$$\forall x \forall y \exists z \forall w (w \in z \leftrightarrow w = x \vee w = y).$$

- z is the unordered pair of x and y .

d) The union axiom (Union):

$$\forall x \exists y \forall z (z \in y \leftrightarrow \exists w (w \in x \wedge z \in w))$$

- y is the union of all elements of x .

e) The separation schema (Sep) postulates for every \in -formula $\varphi(z, x_1, \dots, x_n)$:

$$\forall x_1 \dots \forall x_n \forall x \exists y \forall z (z \in y \leftrightarrow z \in x \wedge \varphi(z, x_1, \dots, x_n))$$

- this is an infinite scheme of axioms, the set z consists of all elements of x which satisfy φ .

f) The powerset axiom (Pow):

$$\forall x \exists y \forall z (z \in y \leftrightarrow \forall w (w \in z \rightarrow w \in x))$$

- y consists of all subsets of x .

g) The replacement schema (Rep) postulates for every \in -formula $\varphi(x, y, x_1, \dots, x_n)$:

$$\forall x_1 \dots \forall x_n (\forall x \forall y \forall y' ((\varphi(x, y, x_1, \dots, x_n) \wedge \varphi(x, y', x_1, \dots, x_n)) \rightarrow y = y') \rightarrow \forall u \exists v \forall y (y \in v \leftrightarrow \exists x (x \in u \wedge \varphi(x, y, x_1, \dots, x_n))))$$

- v is the image of u under the map defined by φ .

h) The axiom of infinity (Inf):

$$\exists x (\exists y (y \in x \wedge \forall z \neg z \in y) \wedge \forall y (y \in x \rightarrow \exists z (z \in x \wedge \forall w (w \in z \leftrightarrow w \in y \vee w = y))))$$

- by the closure properties of x , x has to be infinite.

i) The foundation schema (Found) postulates for every \in -formula $\varphi(x, x_1, \dots, x_n)$:

$$\forall x_1 \dots \forall x_n (\exists x \varphi(x, x_1, \dots, x_n) \rightarrow \exists x (\varphi(x, x_1, \dots, x_n) \wedge \forall x' (x' \in x \rightarrow \neg \varphi(x', x_1, \dots, x_n))))$$

- if φ is satisfiable then there are \in -minimal elements satisfying φ .

2.14 ZF in Class Notation

Using class terms, the ZF can be formulated concisely:

Theorem 35. *The ZF axioms are equivalent to the following system; we take all free variables of the axioms to be universally quantified:*

- a) *Ex:* $\emptyset \in V$.
- b) *Ext:* $x \subseteq y \wedge y \subseteq x \rightarrow x = y$.
- c) *Pair:* $\{x, y\} \in V$.
- d) *Union:* $\bigcup x \in V$.
- e) *Sep:* $A \cap x \in V$.
- f) *Pow:* $\mathcal{P}(x) \in V$.
- g) *Rep:* F is a function $\rightarrow F[x] \in V$.
- h) *Inf:* $\exists x (0 \in x \wedge \forall n \in x \ n + 1 \in x)$.
- i) *Found:* $A \neq \emptyset \rightarrow \exists x \in A \ A \cap x = \emptyset$.

This axiom system can be used as a foundation for all of mathematics. Axiomatic set theory considers various axiom systems of set theory.

Definition 36. *The axiom system ZF^- consists of the ZF-axioms except the power set axiom. The axiom System ST (“set theory”) consists of the ZF-axioms except the axiom of infinity. The system EML (“elementary set theory”) consists of the axioms Ex, Ext, Pair, and Union.*

Exercise 12. Consider the axiom system HF consisting of the axioms of EML together with the induction principle: for every term B postulate

$$\forall x, y (x \subseteq B \wedge y \in B \rightarrow x \cup \{y\} \in B) \rightarrow B = V.$$

Show that every axiom of ZF except Inf is provable in HF, and that HF proves the *negation* of Inf (HF axiomatizes the **hereditarily finite sets**, i.e., those sets such that the set itself and all its iterated elements are finite).

3 Ordinal Numbers

We had defined some “natural numbers” in set theory. Recall that

$$\begin{aligned} 0 &= \emptyset \\ 1 &= 0 + 1 = 0 \cup \{0\} = \{0\} \\ 2 &= 1 + 1 = 1 \cup \{1\} = \{0, 1\} \\ 3 &= 2 + 1 = 2 \cup \{2\} = \{0, 1, 2\} \\ &\vdots \end{aligned}$$

We would then like to have $\mathbb{N} = \{0, 1, 2, 3, \dots\}$. To obtain a set theoretic formalization of numbers we note some properties of the informal presentation:

1. “Numbers” are ordered by the \in -relation:

$$m < n \text{ iff } m \in n.$$

E.g., $1 \in 3$ but not $3 \in 1$.

2. On each “number”, the \in -relation is a *strict linear order*: $3 = \{0, 1, 2\}$ is strictly linearly ordered by \in .
3. “Numbers” are “complete” with respect to smaller “numbers”

$$i < j < m \rightarrow i \in m.$$

This can be written with the \in -relation as

$$i \in j \in m \rightarrow i \in m.$$

Definition 37.

- a) A is transitive, $\text{Trans}(A)$, iff $\forall y \in A \forall x \in y x \in A$.
- b) x is an ordinal (number), $\text{Ord}(x)$, if $\text{Trans}(x) \wedge \forall y \in x \text{Trans}(y)$.
- c) Let $\text{Ord} := \{x \mid \text{Ord}(x)\}$ be the class of all ordinal numbers.

We shall use small greek letter α, β, \dots as variables for ordinals. So $\exists \alpha \varphi$ stands for $\exists \alpha \in \text{Ord} \varphi$, and $\{\alpha \mid \varphi\}$ for $\{\alpha \mid \text{Ord}(\alpha) \wedge \varphi\}$.

Exercise 13. Show that arbitrary unions and intersections of transitive sets are again transitive.

We shall see that the ordinals extend the standard natural numbers. Ordinals are particularly adequate for enumerating infinite sets.

Theorem 38.

- a) $0 \in \text{Ord}$.
- b) $\forall \alpha \alpha + 1 \in \text{Ord}$.

Proof. a) $\text{Trans}(\emptyset)$ since formulas of the form $\forall y \in \emptyset \dots$ are tautologously true. Similarly $\forall y \in \emptyset \text{Trans}(y)$.

b) Assume $\alpha \in \text{Ord}$.

(1) $\text{Trans}(\alpha + 1)$.

Proof. Let $u \in v \in \alpha + 1 = \alpha \cup \{\alpha\}$.

Case 1. $v \in \alpha$. Then $u \in \alpha \subseteq \alpha + 1$, since α is transitive.

Case 2. $v = \alpha$. Then $u \in \alpha \subseteq \alpha + 1$. *qed*(1)

(2) $\forall y \in \alpha + 1 \text{Trans}(y)$.

Proof. Let $y \in \alpha + 1 = \alpha \cup \{\alpha\}$.

Case 1. $y \in \alpha$. Then $\text{Trans}(y)$ since α is an ordinal.

Case 2. $y = \alpha$. Then $\text{Trans}(y)$ since α is an ordinal. □

Exercise 14.

- a) Let $A \subseteq \text{Ord}$ be a term, $A \neq \emptyset$. Then $\bigcap A \in \text{Ord}$.

b) Let $x \subseteq \text{Ord}$ be a set. Then $\bigcup x \in \text{Ord}$.

Theorem 39. $\text{Trans}(\text{Ord})$.

Proof. This follows immediately from the transitivity definition of Ord . \square

Exercise 15. Show that Ord is a proper class. (Hint: if $\text{Ord} \in V$ then $\text{Ord} \in \text{Ord}$.)

Theorem 40. *The class Ord is strictly linearly ordered by \in , i.e.,*

- a) $\forall \alpha, \beta, \gamma (\alpha \in \beta \wedge \beta \in \gamma \rightarrow \alpha \in \gamma)$.
- b) $\forall \alpha \alpha \notin \alpha$.
- c) $\forall \alpha, \beta (\alpha \in \beta \vee \alpha = \beta \vee \beta \in \alpha)$.

Proof. a) Let $\alpha, \beta, \gamma \in \text{Ord}$ and $\alpha \in \beta \wedge \beta \in \gamma$. Then γ is transitive, and so $\alpha \in \gamma$.

b) follows immediately from the non-circularity of the \in -relation.

c) Assume that there are “incomparable” ordinals. By the foundation schema choose $\alpha_0 \in \text{Ord}$ \in -minimal such that $\exists \beta \neg(\alpha_0 \in \beta \vee \alpha_0 = \beta \vee \beta \in \alpha_0)$. Again, choose $\beta_0 \in \text{Ord}$ \in -minimal such that $\neg(\alpha_0 \in \beta_0 \vee \alpha_0 = \beta_0 \vee \beta_0 \in \alpha_0)$. We obtain a contradiction by showing that $\alpha_0 = \beta_0$:

Let $\alpha \in \alpha_0$. By the \in -minimality of α_0 , α is comparable with β_0 : $\alpha \in \beta_0 \vee \alpha = \beta_0 \vee \beta_0 \in \alpha$. If $\alpha = \beta_0$ then $\beta_0 \in \alpha_0$ and α_0, β_0 would be comparable, contradiction. If $\beta_0 \in \alpha$ then $\beta_0 \in \alpha_0$ by the transitivity of α_0 and again α_0, β_0 would be comparable, contradiction. Hence $\alpha \in \beta_0$.

For the converse let $\beta \in \beta_0$. By the \in -minimality of β_0 , β is comparable with α_0 : $\beta \in \alpha_0 \vee \beta = \alpha_0 \vee \alpha_0 \in \beta$. If $\beta = \alpha_0$ then $\alpha_0 \in \beta_0$ and α_0, β_0 would be comparable, contradiction. If $\alpha_0 \in \beta$ then $\alpha_0 \in \beta_0$ by the transitivity of β_0 and again α_0, β_0 would be comparable, contradiction. Hence $\beta \in \alpha_0$.

But then $\alpha_0 = \beta_0$ contrary to the choice of β_0 . \square

Definition 41. Let $< := \in \cap (\text{Ord} \times \text{Ord}) = \{(\alpha, \beta) \mid \alpha \in \beta\}$ be the natural strict linear ordering of Ord by the \in -relation.

Theorem 42. Let $\alpha \in \text{Ord}$. Then $\alpha + 1$ is the immediate successor of α in the \in -relation:

- a) $\alpha < \alpha + 1$;
- b) if $\beta < \alpha + 1$, then $\beta = \alpha$ or $\beta < \alpha$.

Definition 43. Let α be an ordinal. α is a successor ordinal, $\text{Succ}(\alpha)$, iff $\exists \beta \alpha = \beta + 1$. α is a limit ordinal, $\text{Lim}(\alpha)$, iff $\alpha \neq 0$ and α is not a successor ordinal. Also let

$$\text{Succ} := \{\alpha \mid \text{Succ}(\alpha)\} \text{ and } \text{Lim} := \{\alpha \mid \text{Lim}(\alpha)\}.$$

The existence of limit ordinals will be discussed together with the formalization of the natural numbers.

3.1 Ordinal induction

Ordinals satisfy an *induction theorem* which generalizes *complete induction* on the integers:

Theorem 44. Let $\varphi(x, v_0, \dots, v_{n-1})$ be an \in -formula and $x_0, \dots, x_{n-1} \in V$. Assume that the property $\varphi(x, x_0, \dots, x_{n-1})$ is inductive, i.e.,

$$\forall \alpha (\forall \beta \in \alpha \varphi(\beta, x_0, \dots, x_{n-1}) \rightarrow \varphi(\alpha, x_0, \dots, x_{n-1})).$$

Then φ holds for all ordinals:

$$\forall \alpha \varphi(\alpha, x_0, \dots, x_{n-1}).$$

Proof. It suffices to show that

$$B = \{x \mid x \in \text{Ord} \rightarrow \varphi(x, x_0, \dots, x_{n-1})\} = V.$$

Theorem 33 implies

$$\forall x (x \subseteq B \rightarrow x \in B) \rightarrow B = V$$

and it suffices to show

$$\forall x (x \subseteq B \rightarrow x \in B).$$

Consider $x \subseteq B$. If $x \notin \text{Ord}$ then $x \in B$. So assume $x \in \text{Ord}$. For $\beta \in x$ we have $\beta \in B$, $\beta \in \text{Ord}$, and so $\varphi(\beta, x_0, \dots, x_{n-1})$. By the inductivity of φ we get $\varphi(x, x_0, \dots, x_{n-1})$ and again $x \in B$. \square

Induction can be formulated in various forms:

Exercise 16. Prove the following transfinite induction principle: Let $\varphi(x) = \varphi(x, v_0, \dots, v_{n-1})$ be an \in -formula and $x_0, \dots, x_{n-1} \in V$. Assume

- a) $\varphi(0)$ (the initial case),
- b) $\forall \alpha (\varphi(\alpha) \rightarrow \varphi(\alpha + 1))$ (the successor step),
- c) $\forall \lambda \in \text{Lim} (\forall \alpha < \lambda \varphi(\alpha) \rightarrow \varphi(\lambda))$ (the limit step).

Then $\forall \alpha \varphi(\alpha)$.

3.2 Natural Numbers

We have $0, 1, \dots \in \text{Ord}$. We shall now define and study the set of *natural numbers/integers* within set theory. Recall the axiom of infinity:

$$\exists x (0 \in x \wedge \forall u \in x u + 1 \in x).$$

The set of natural numbers should be the \subseteq -smallest such x .

Definition 45. Let $\omega = \bigcap \{x \mid 0 \in x \wedge \forall u \in x u + 1 \in x\}$ be the set of natural numbers. Sometimes we write \mathbb{N} instead of ω .

Theorem 46.

- a) $\omega \in V$.
- b) $\omega \subseteq \text{Ord}$.
- c) $(\omega, 0, +1)$ satisfy the second order PEANO axiom, i.e.,

$$\forall x \subseteq \omega (0 \in x \wedge \forall n \in x n + 1 \in x \rightarrow x = \omega).$$

- d) $\omega \in \text{Ord}$.
- e) ω is a limit ordinal.

Proof. a) By the axiom of infinity take a set x_0 such that

$$0 \in x_0 \wedge \forall u \in x_0 u + 1 \in x_0.$$

Then

$$\omega = \bigcap \{x \mid 0 \in x \wedge \forall u \in x u + 1 \in x\} = x_0 \cap \bigcap \{x \mid 0 \in x \wedge \forall u \in x u + 1 \in x\} \in V$$

by the separation schema.

b) By a), $\omega \cap \text{Ord} \in V$. Obviously $0 \in \omega \cap \text{Ord} \wedge \forall u \in \omega \cap \text{Ord} u + 1 \in \omega \cap \text{Ord}$. So $\omega \cap \text{Ord}$ is one factor of the intersection in the definition of ω and so $\omega \subseteq \omega \cap \text{Ord}$. Hence $\omega \subseteq \text{Ord}$.

c) Let $x \subseteq \omega$ and $0 \in x \wedge \forall u \in x u + 1 \in x$. Then x is one factor of the intersection in the definition of ω and so $\omega \subseteq x$. This implies $x = \omega$.

d) By b), every element of ω is transitive and it suffices to show that ω is transitive. Let

$$x = \{n \mid n \in \omega \wedge \forall m \in n m \in \omega\} \subseteq \omega.$$

We show that the hypothesis of c) holds for x . $0 \in x$ is trivial. Let $u \in x$. Then $u + 1 \in \omega$. Let $m \in u + 1$. If $m \in u$ then $m \in \omega$ by the assumption that $u \in x$. If $m = u$ then $m \in x \subseteq \omega$. Hence $u + 1 \in x$ and $\forall u \in x u + 1 \in x$. By b), $x = \omega$. So $\forall n \in \omega n \in x$, i.e.,

$$\forall n \in \omega \forall m \in n m \in \omega.$$

e) Of course $\omega \neq 0$. Assume for a contradiction that ω is a successor ordinal, say $\omega = \alpha + 1$. Then $\alpha \in \omega$. Since ω is closed under the $+1$ -operation, $\omega = \alpha + 1 \in \omega$. Contradiction. \square

Thus the axiom of infinity implies the existence of the set of natural numbers, which is also the smallest limit ordinal. The axiom of infinity can now be reformulated equivalently as:

h) Inf: $\omega \in V$.

3.3 Ordinal recursion

Recursion, often called induction, over the natural numbers is a ubiquitous method for defining mathematical objects. We prove the following *recursion theorem* for ordinals.

Theorem 47. *Let $G: V \rightarrow V$. Then there is a canonical class term F , given by the subsequent proof, such that*

$$F: \text{Ord} \rightarrow V \text{ and } \forall \alpha F(\alpha) = G(F \upharpoonright \alpha).$$

We then say that F is defined recursively (over the ordinals) by the recursion rule G . F is unique in the sense that if another term F' satisfies

$$F': \text{Ord} \rightarrow V \text{ and } \forall \alpha F'(\alpha) = G(F' \upharpoonright \alpha)$$

then $F = F'$.

Proof. We say that $H: \text{dom}(H) \rightarrow V$ is G -recursive if

$$\text{dom}(H) \subseteq \text{Ord}, \text{dom}(H) \text{ is transitive, and } \forall \alpha \in \text{dom}(H) H(\alpha) = G(H \upharpoonright \alpha).$$

(1) Let H, H' be G -recursive. Then H, H' are compatible, i.e., $\forall \alpha \in \text{dom}(H) \cap \text{dom}(H') H(\alpha) = H'(\alpha)$.

Proof. We want to show that

$$\forall \alpha \in \text{Ord} (\alpha \in \text{dom}(H) \cap \text{dom}(H') \rightarrow H(\alpha) = H'(\alpha)).$$

By the induction theorem it suffices to show that $\alpha \in \text{dom}(H) \cap \text{dom}(H') \rightarrow H(\alpha) = H'(\alpha)$ is inductive, i.e.,

$$\forall \alpha \in \text{Ord} (\forall y \in \alpha (y \in \text{dom}(H) \cap \text{dom}(H') \rightarrow H(y) = H'(y)) \rightarrow (\alpha \in \text{dom}(H) \cap \text{dom}(H') \rightarrow H(\alpha) = H'(\alpha))).$$

So let $\alpha \in \text{Ord}$ and $\forall y \in \alpha (y \in \text{dom}(H) \cap \text{dom}(H') \rightarrow H(y) = H'(y))$. Let $\alpha \in \text{dom}(H) \cap \text{dom}(H')$. Since $\text{dom}(H)$ and $\text{dom}(H')$ are transitive, $\alpha \subseteq \text{dom}(H)$ and $\alpha \subseteq \text{dom}(H')$. By assumption

$$\forall y \in \alpha H(y) = H'(y).$$

Hence $H \upharpoonright \alpha = H' \upharpoonright \alpha$. Then

$$H(\alpha) = G(H \upharpoonright \alpha) = G(H' \upharpoonright \alpha) = H'(\alpha).$$

qed(1)

Let

$$F := \bigcup \{f \mid f \text{ is } G\text{-recursive}\}.$$

be the union of the class of all *approximations* to the desired function F .

(2) F is G -recursive.

Proof. By (1), F is a function. Its domain $\text{dom}(F)$ is the union of transitive classes of ordinals and hence $\text{dom}(F) \subseteq \text{Ord}$ is transitive.

Let $\alpha \in \text{dom}(F)$. Take some G -recursive function f such that $\alpha \in \text{dom}(f)$. Since $\text{dom}(f)$ is transitive, we have

$$\alpha \subseteq \text{dom}(f) \subseteq \text{dom}(F).$$

Moreover

$$F(\alpha) = f(\alpha) = G(f \upharpoonright \alpha) = G(F \upharpoonright \alpha).$$

qed(2)

(3) $\forall \alpha \alpha \in \text{dom}(F)$.

Proof. By induction on the ordinals. We have to show that $\alpha \in \text{dom}(F)$ is inductive in the variable α . So let $\alpha \in \text{Ord}$ and $\forall y \in \alpha y \in \text{dom}(F)$. Hence $\alpha \subseteq \text{dom}(F)$. Let

$$f = F \upharpoonright \alpha \cup \{(\alpha, G(F \upharpoonright \alpha))\}.$$

f is a function with $\text{dom}(f) = \alpha + 1 \in \text{Ord}$. Let $\alpha' < \alpha + 1$. If $\alpha' < \alpha$ then

$$f(\alpha') = F(\alpha') = G(F \upharpoonright \alpha') = G(f \upharpoonright \alpha').$$

if $\alpha' = \alpha$ then also

$$f(\alpha') = f(\alpha) = G(F \upharpoonright \alpha) = G(f \upharpoonright \alpha) = G(f \upharpoonright \alpha').$$

Hence f is G -recursive and $\alpha \in \text{dom}(f) \subseteq \text{dom}(F)$. qed(3)

The extensional uniqueness of F follows from (1) □

Theorem 48. Let $a_0 \in V$, $G_{\text{succ}}: \text{Ord} \times V \rightarrow V$, and $G_{\text{lim}}: \text{Ord} \times V \rightarrow V$. Then there is a canonically defined class term $F: \text{Ord} \rightarrow V$ such that

- a) $F(0) = a_0$;
- b) $\forall \alpha F(\alpha + 1) = G_{\text{succ}}(\alpha, F(\alpha))$;
- c) $\forall \lambda \in \text{Lim} F(\lambda) = G_{\text{lim}}(\lambda, F \upharpoonright \lambda)$.

Again F is unique in the sense that if some F' also satisfies a)-c) then $F = F'$.

We say that F is recursively defined by the properties a)-c).

Proof. We incorporate a_0 , G_{succ} , and G_{lim} into a single recursion rule $G: V \rightarrow V$,

$$G(f) = \begin{cases} a_0, & \text{if } f = \emptyset, \\ G_{\text{succ}}(\alpha, f(\alpha)), & \text{if } f: \alpha + 1 \rightarrow V, \\ G_{\text{lim}}(\lambda, f), & \text{if } f: \lambda \rightarrow V \text{ and } \text{Lim}(\lambda), \\ \emptyset, & \text{else.} \end{cases}$$

Then the term $F: \text{Ord} \rightarrow V$ defined recursively by the recursion rule G satisfies the theorem. □

In many cases, the *limit rule* will just require to form the union of the previous values so that

$$F(\lambda) = \bigcup_{\alpha < \lambda} F(\alpha).$$

Such recursions are called *continuous* (at limits).

3.4 Ordinal Arithmetic

We extend the recursion rules of standard integer arithmetic continuously to obtain transfinite version of the arithmetic operations. The initial operation of ordinal arithmetic is the $+1$ -operation defined before. Ordinal arithmetic satisfies some but not all laws of integer arithmetic.

Definition 49. Define ordinal addition $+: \text{Ord} \times \text{Ord} \rightarrow \text{Ord}$ recursively by

$$\begin{aligned} \delta + 0 &= \delta \\ \delta + (\alpha + 1) &= (\delta + \alpha) + 1 \\ \delta + \lambda &= \bigcup_{\alpha < \lambda} (\delta + \alpha), \text{ for limit ordinals } \lambda \end{aligned}$$

Definition 50. Define ordinal multiplication $\cdot: \text{Ord} \times \text{Ord} \rightarrow \text{Ord}$ recursively by

$$\begin{aligned} \delta \cdot 0 &= 0 \\ \delta \cdot (\alpha + 1) &= (\delta \cdot \alpha) + \delta \\ \delta \cdot \lambda &= \bigcup_{\alpha < \lambda} (\delta \cdot \alpha), \text{ for limit ordinals } \lambda \end{aligned}$$

Definition 51. Define ordinal exponentiation $_ - : \text{Ord} \times \text{Ord} \rightarrow \text{Ord}$ recursively by

$$\begin{aligned}\delta^0 &= 1 \\ \delta^{\alpha+1} &= \delta^\alpha \cdot \delta \\ \delta^\lambda &= \bigcup_{\alpha < \lambda} \delta^\alpha, \text{ for limit ordinals } \lambda\end{aligned}$$

Exercise 17. Explore which of the standard *ring axioms* hold for the ordinals with addition and multiplication. Give proofs and counterexamples.

Exercise 18. Show that for any ordinal α , $\alpha + \omega$ is a limit ordinal. Use this to show that the class Lim of all limit ordinals is a proper class.

4 Number Systems

We are now able to give set-theoretic formalizations of the standard number systems with their arithmetic operations.

4.1 Natural Numbers

Definition 52. The structure

$$\mathbb{N} := (\omega, + | (\omega \times \omega), \cdot | (\omega \times \omega), < | (\omega \times \omega), 0, 1)$$

is called the structure of natural numbers, or arithmetic. We sometimes denote this structure by

$$\mathbb{N} := (\omega, +, \cdot, <, 0, 1).$$

\mathbb{N} is an adequate formalization of arithmetic within set theory since \mathbb{N} satisfies all standard arithmetical axioms.

Exercise 19. Prove:

- $+ | [\omega \times \omega] := \{m + n \mid m \in \omega \wedge n \in \omega\} \subseteq \omega$.
- $\cdot | [\omega \times \omega] := \{m \cdot n \mid m \in \omega \wedge n \in \omega\} \subseteq \omega$.
- Addition and multiplication are commutative on ω .
- Addition and multiplication satisfy the usual monotonicity laws with respect to $<$.

Definition 53. We define the structure

$$\mathbb{Z} := (\mathbb{Z}, +^{\mathbb{Z}}, \cdot^{\mathbb{Z}}, <^{\mathbb{Z}}, 0^{\mathbb{Z}}, 1^{\mathbb{Z}})$$

of integers as follows:

- Define an equivalence relation \approx on $\mathbb{N} \times \mathbb{N}$ by

$$(a, b) \approx (a', b') \text{ iff } a + b' = a' + b.$$

- Let $a - b := [(a, b)]_{\approx}$ be the equivalence class of (a, b) in \approx . Note that every $a - b$ is a set.

- Let $\mathbb{Z} := \{a - b \mid a \in \mathbb{N} \wedge b \in \mathbb{N}\}$ be the set of integers.

- Define the integer addition $+^{\mathbb{Z}}: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ by

$$(a - b) +^{\mathbb{Z}} (a' - b') := (a + a') - (b + b').$$

- Define the integer multiplication $\cdot^{\mathbb{Z}}: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ by

$$(a - b) \cdot^{\mathbb{Z}} (a' - b') := (a \cdot a' + b \cdot b') - (a \cdot b' + a' \cdot b).$$

f) Define the strict linear order $<^{\mathbb{Z}}$ on \mathbb{Z} by

$$(a - b) <^{\mathbb{Z}} (a' - b') \text{ iff } a + b' < a' + b.$$

g) Let $0^{\mathbb{Z}} := 0 - 0$ and $1^{\mathbb{Z}} := 1 - 0$.

Exercise 20. Check that the above definitions are *sound*, i.e., that they do not depend on the choice of representatives of equivalence classes.

Exercise 21. Check that \mathbb{Z} satisfies (a sufficient number) of the standard axioms for rings.

The structure \mathbb{Z} extends the structure \mathbb{N} in a natural and familiar way: define an injective map $e: \mathbb{N} \rightarrow \mathbb{Z}$ by

$$n \mapsto n - 0.$$

The embedding e is a *homomorphism*:

- a) $e(0) = 0 - 0 = 0^{\mathbb{Z}}$ and $e(1) = 1 - 0 = 1^{\mathbb{Z}}$;
- b) $e(m + n) = (m + n) - 0 = (m + n) - (0 + 0) = (m - 0) +^{\mathbb{Z}} (n - 0) = e(m) +^{\mathbb{Z}} e(n)$;
- c) $e(m \cdot n) = (m \cdot n) - 0 = (m \cdot n + 0 \cdot 0) - (m \cdot 0 + n \cdot 0) = (m - 0) \cdot^{\mathbb{Z}} (n - 0) = e(m) \cdot^{\mathbb{Z}} e(n)$;
- d) $m < n \leftrightarrow m + 0 < n + 0 \leftrightarrow (m - 0) <^{\mathbb{Z}} (n - 0) \leftrightarrow e(m) <^{\mathbb{Z}} e(n)$.

By this injective homomorphism, one may consider \mathbb{N} as a *substructure* of \mathbb{Z} : $\mathbb{N} \subseteq \mathbb{Z}$.

4.2 Rational Numbers

Definition 54. We define the structure

$$\mathbb{Q}_0^+ := (\mathbb{Q}_0^+, +^{\mathbb{Q}}, \cdot^{\mathbb{Q}}, <^{\mathbb{Q}}, 0^{\mathbb{Q}}, 1^{\mathbb{Q}})$$

of non-negative rational numbers *as follows*:

a) Define an equivalence relation \simeq on $\mathbb{N} \times (\mathbb{N} \setminus \{0\})$ by

$$(a, b) \simeq (a', b') \text{ iff } a \cdot b' = a' \cdot b.$$

b) Let $\frac{a}{b} := [(a, b)]_{\simeq}$ be the equivalence class of (a, b) in \simeq . Note that $\frac{a}{b}$ is a set.

c) Let $\mathbb{Q}_0^+ := \{\frac{a}{b} \mid a \in \mathbb{N} \wedge b \in (\mathbb{N} \setminus \{0\})\}$ be the set of non-negative rationals.

d) Define the rational addition $+^{\mathbb{Q}}: \mathbb{Q}_0^+ \times \mathbb{Q}_0^+ \rightarrow \mathbb{Q}_0^+$ by

$$\frac{a}{b} +^{\mathbb{Q}} \frac{a'}{b'} := \frac{a \cdot b' + a' \cdot b}{b \cdot b'}.$$

e) Define the rational multiplication $\cdot^{\mathbb{Q}}: \mathbb{Q}_0^+ \times \mathbb{Q}_0^+ \rightarrow \mathbb{Q}_0^+$ by

$$\frac{a}{b} \cdot^{\mathbb{Q}} \frac{a'}{b'} := \frac{a \cdot a'}{b \cdot b'}.$$

f) Define the strict linear order $<^{\mathbb{Q}}$ on \mathbb{Q}_0^+ by

$$\frac{a}{b} <^{\mathbb{Q}} \frac{a'}{b'} \text{ iff } a \cdot b' < a' \cdot b.$$

g) Let $0^{\mathbb{Q}} := \frac{0}{1}$ and $1^{\mathbb{Q}} := \frac{1}{1}$.

Again one can check the soundness of the definitions and the well-known laws of standard non-negative rational numbers. Also one may assume \mathbb{N} to be embedded into \mathbb{Q}_0^+ as a substructure. The transfer from non-negative to *all* rationals, including negative rationals can be performed in analogy to the transfer from \mathbb{N} to \mathbb{Z} .

Definition 55. We define the structure

$$\mathbb{Q} := (\mathbb{Q}, +^{\mathbb{Q}}, \cdot^{\mathbb{Q}}, <^{\mathbb{Q}}, 0^{\mathbb{Q}}, 1^{\mathbb{Q}})$$

of rational numbers as follows:

a) Define an equivalence relation \approx on $\mathbb{Q}_0^+ \times \mathbb{Q}_0^+$ by

$$(p, q) \approx (p', q') \text{ iff } p + q' = p' + q.$$

b) Let $p - q := [(p, q)]_{\approx}$ be the equivalence class of (p, q) in \approx .

c) Let $\mathbb{Q} := \{p - q \mid p \in \mathbb{Q}_0^+ \wedge q \in \mathbb{Q}_0^+\}$ be the set of rationals.

Exercise 22. Continue the definition of the structure \mathbb{Q} and prove the relevant properties.

4.3 Real Numbers

Definition 56. $r \subseteq \mathbb{Q}_0^+$ is a positive real number if

a) $\forall p \in r \forall q \in \mathbb{Q}_0^+ (q <^{\mathbb{Q}} p \rightarrow q \in r)$, i.e., r is an initial segment of $(\mathbb{Q}_0^+, <^{\mathbb{Q}})$;

b) $\forall p \in r \exists q \in r p <^{\mathbb{Q}} q$, i.e., r is right-open in $(\mathbb{Q}_0^+, <^{\mathbb{Q}})$;

c) $0 \in r \neq \mathbb{Q}_0^+$, i.e., r is nonempty and bounded in $(\mathbb{Q}_0^+, <^{\mathbb{Q}})$.

Definition 57. We define the structure

$$\mathbb{R}^+ := (\mathbb{R}^+, +^{\mathbb{R}}, \cdot^{\mathbb{R}}, <^{\mathbb{R}}, 1^{\mathbb{R}})$$

of positive real numbers as follows:

a) Let \mathbb{R}^+ be the set of positive reals.

b) Define the real addition $+^{\mathbb{R}}: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ by

$$r +^{\mathbb{R}} r' = \{p +^{\mathbb{Q}} p' \mid p \in r \wedge p' \in r'\}.$$

c) Define the real multiplication $\cdot^{\mathbb{R}}: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ by

$$r \cdot^{\mathbb{R}} r' = \{p \cdot^{\mathbb{Q}} p' \mid p \in r \wedge p' \in r'\}.$$

d) Define the strict linear order $<^{\mathbb{R}}$ on \mathbb{R}^+ by

$$r <^{\mathbb{R}} r' \text{ iff } r \subseteq r' \wedge r \neq r'.$$

e) Let $1^{\mathbb{R}} := \{p \in \mathbb{Q}_0^+ \mid q <^{\mathbb{Q}} 1\}$.

We justify some details of the definition.

Lemma 58.

a) $\mathbb{R}^+ \in V$.

b) If $r, r' \in \mathbb{R}^+$ then $r +^{\mathbb{R}} r', r \cdot^{\mathbb{R}} r' \in \mathbb{R}^+$.

c) $<^{\mathbb{R}}$ is a strict linear order on \mathbb{R}^+ .

Proof. a) If $r \in \mathbb{R}^+$ then $r \subseteq \mathbb{Q}_0^+$ and $r \in \mathcal{P}(\mathbb{Q}_0^+)$. Thus $\mathbb{R}^+ \subseteq \mathcal{P}(\mathbb{Q}_0^+)$, and \mathbb{R}^+ is a set by the power set axiom and separation.

b) Let $r, r' \in \mathbb{R}^+$. We show that

$$r \cdot^{\mathbb{R}} r' = \{p \cdot^{\mathbb{Q}} p' \mid p \in r \wedge p' \in r'\} \in \mathbb{R}^+.$$

Obviously $r \cdot^{\mathbb{R}} r' \subseteq \mathbb{Q}_0^+$ is a non-empty bounded initial segment of $(\mathbb{Q}_0^+, <^{\mathbb{Q}})$.

Consider $p \in r \cdot^{\mathbb{R}} r', q \in \mathbb{Q}_0^+, q <^{\mathbb{Q}} p$. Let $p = \frac{a}{b} \cdot^{\mathbb{Q}} \frac{a'}{b'}$ where $\frac{a}{b} \in r$ and $\frac{a'}{b'} \in r'$. Let $q = \frac{c}{d}$. Then $\frac{c}{d} = \frac{c \cdot b'}{d \cdot a'} \cdot^{\mathbb{Q}} \frac{a'}{b'}$, where

$$\frac{c \cdot b'}{d \cdot a'} = q \cdot^{\mathbb{Q}} \frac{b'}{a'} <^{\mathbb{Q}} p \cdot^{\mathbb{Q}} \frac{b'}{a'} = \frac{a}{b} \cdot^{\mathbb{Q}} \frac{a'}{b'} \cdot^{\mathbb{Q}} \frac{b'}{a'} = \frac{a}{b} \in r.$$

Hence $\frac{c \cdot b'}{d \cdot a'} \in r$ and

$$\frac{c}{d} = \frac{c \cdot b'}{d \cdot a'} \cdot_{\mathbb{Q}} \frac{a'}{b'} \in r \cdot_{\mathbb{R}} r'.$$

Similarly one can show that $r \cdot_{\mathbb{R}} r'$ is open on the right-hand side.

c) The transitivity of $<^{\mathbb{R}}$ follows from the transitivity of the relation \subsetneq . To show that $<^{\mathbb{R}}$ is connex, consider $r, r' \in \mathbb{R}^+$, $r \neq r'$. Then r and r' are different subsets of \mathbb{Q}_0^+ . Without loss of generality we may assume that there is some $p \in r' \setminus r$. We show that then $r <^{\mathbb{R}} r'$, i.e., $r \subsetneq r'$. Consider $q \in r$. Since $p \notin r$ we have $p \not\leq^{\mathbb{Q}} q$ and $q \leq^{\mathbb{Q}} p$. Since r' is an initial segment of \mathbb{Q}_0^+ , $q \in r'$. \square

Exercise 23. Show that $(\mathbb{R}^+, \cdot, 1^{\mathbb{R}})$ is a multiplicative group.

We can now construct the complete real line \mathbb{R} from \mathbb{R}^+ just like we constructed \mathbb{Z} from \mathbb{N} . Details are left to the reader. We can also proceed to define the structure \mathbb{C} of complex numbers from \mathbb{R} .

Exercise 24. Formalize the structure \mathbb{C} of complex numbers such that $\mathbb{R} \subseteq \mathbb{C}$.

4.4 Discussion

The constructions carried out in the previous subsections contained many arbitrary choices. One could, e.g., define rational numbers as *reduced* fractions instead of equivalence classes of fractions, ensure that the canonical embeddings of number systems are inclusions, etc. If such choices have been made in reasonable ways we obtain the following theorem, which contains everything one wants to know about the number systems. So the statements of the following theorem can be seen as first- and second-order axioms for these systems.

Theorem 59. *There are structures $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and \mathbb{C} with the following properties:*

a) *the domains of these structures which are also denoted by $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and \mathbb{C} , resp., satisfy*

$$\omega = \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C};$$

b) *there are functions $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ and $\cdot: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ on \mathbb{C} which are usually written as binary infix operations;*

c) *$(\mathbb{C}, +, \cdot, 0, 1)$ is a field; for $a, b \in \mathbb{C}$ write $a - b$ for the unique element z such that $a = b + z$; for $a, b \in \mathbb{C}$ with $b \neq 0$ write $\frac{a}{b}$ for the unique element z such that $a = b \cdot z$;*

d) *there is a constant i , the imaginary unit, such that $i \cdot i + 1 = 0$ and*

$$\mathbb{C} = \{x + i \cdot y \mid x, y \in \mathbb{R}\};$$

e) *there is a strict linear order $<$ on \mathbb{R} such that $(\mathbb{R}, <, + \upharpoonright \mathbb{R}^2, \cdot \upharpoonright \mathbb{R}^2, 0, 1)$ is an ordered field.*

f) *$(\mathbb{R}, <)$ is complete, i.e., bounded subsets of \mathbb{R} possess suprema:*

$$\forall X \subseteq \mathbb{R} (X \neq \emptyset \wedge \exists b \in \mathbb{R} \forall x \in X x < b \longrightarrow \exists b \in \mathbb{R} (\forall x \in X x < b \wedge \neg \exists b' < b \forall x \in X x < b'))$$

g) *\mathbb{Q} is dense in $(\mathbb{R}, <)$:*

$$\forall r, s \in \mathbb{R} (r < s \longrightarrow \exists a, b, c \in \mathbb{Q} a < r < b < s < c);$$

h) *$(\mathbb{Q}, + \upharpoonright \mathbb{Q}^2, \cdot \upharpoonright \mathbb{Q}^2, 0, 1)$ is a field; moreover*

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\} \right\};$$

i) *$(\mathbb{Z}, + \upharpoonright \mathbb{Z}^2, \cdot \upharpoonright \mathbb{Z}^2, 0, 1)$ is a ring with a unit; moreover*

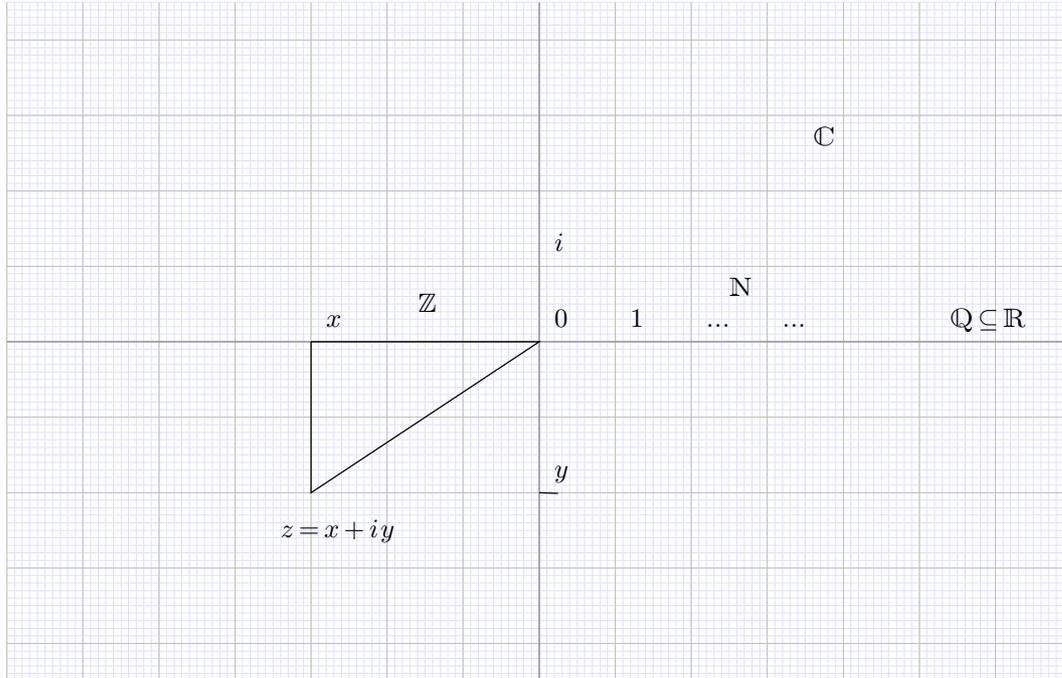
$$\mathbb{Z} = \{a - b \mid a, b \in \mathbb{N}\};$$

j) *$+ \upharpoonright \mathbb{N}^2$ agrees with ordinal addition on ω ; $\cdot \upharpoonright \mathbb{N}^2$ agrees with ordinal multiplication on ω ;*

k) $(\mathbb{N}, +1, 0)$ satisfies the second-order PEANO axioms, i.e., the successor function $n \mapsto n + 1$ is injective, 0 is not in the image of the successor function, and

$$\forall X \subseteq \mathbb{N} (0 \in X \wedge \forall n \in X n + 1 \in X \longrightarrow X = \mathbb{N}).$$

This theorem is all we require from the number systems. The details of the previous construction will not be used again. So we have the standard complex plane, possibly with the identification of \mathbb{N} and ω .



Remark 60. In set theory the set \mathbb{R} of reals is often identified with the sets ${}^\omega\omega$ or ω^2 , basically because all these sets have the same cardinality. We shall come back to this in the context of cardinality theory.

5 Sequences

The notion of a *sequence* is crucial in many contexts.

Definition 61.

- a) A set w is an α -sequence iff $w: \alpha \rightarrow V$; then α is called the length of the α -sequence w and is denoted by $|w|$. w is a sequence iff it is an α -sequence for some α . A sequence w is called finite iff $|w| < \omega$.
- b) A finite sequence $w: n \rightarrow V$ may be denoted by its enumeration w_0, \dots, w_{n-1} where we write w_i instead of $w(i)$. One also writes $w_0 \dots w_{n-1}$ instead of w_0, \dots, w_{n-1} , in particular if w is considered to be a word formed out of the symbols w_0, \dots, w_{n-1} .
- c) An ω -sequence $w: \omega \rightarrow V$ may be denoted by w_0, w_1, \dots where w_0, w_1, \dots suggests a definition of w .
- d) Let $w: \alpha \rightarrow V$ and $w': \alpha' \rightarrow V$ be sequences. Then the concatenation $w \hat{\ } w': \alpha + \alpha' \rightarrow V$ is defined by

$$(w \hat{\ } w') \upharpoonright \alpha = w \upharpoonright \alpha \text{ and } \forall i < \alpha' w \hat{\ } w'(\alpha + i) = w'(i).$$

- e) Let $w: \alpha \rightarrow V$ and $x \in V$. Then the adjunction wx of w by x is defined as

$$wx = w \hat{\ } \{(0, x)\}.$$

Sequences and the concatenation operation satisfy the algebraic laws of a *monoid* with cancellation rules.

Proposition 62. *Let w, w', w'' be sequences. Then*

- a) $(w \hat{ } w') \hat{ } w'' = w \hat{ } (w' \hat{ } w'')$.
- b) $\emptyset \hat{ } w = w \hat{ } \emptyset = w$.
- c) $w \hat{ } w' = w \hat{ } w'' \rightarrow w' = w''$.

There are many other operations on sequences. One can *permute* sequences, substitute elements of a sequence, etc.

5.1 (ω -)Sequences of Reals

ω -sequences are particularly prominent in analysis. One may now define properties like

$$\lim_{i \rightarrow \infty} w_i = z \text{ iff } \forall \varepsilon \in \mathbb{R}^+ \exists m < \omega \forall i < \omega (i \geq m \rightarrow (z - \varepsilon < w_i \wedge w_i < z + \varepsilon))$$

or

$$\forall x: \omega \rightarrow \mathbb{R} (\lim_{i \rightarrow \infty} x_i = a \rightarrow \lim_{i \rightarrow \infty} f(x_i) = f(a)).$$

If x_0, x_1, \dots is given then the partial sums

$$\sum_{i=0}^n x_i$$

are defined recursively as

$$\sum_{i=0}^0 x_i = 0 \quad \text{and} \quad \sum_{i=0}^{n+1} x_i = \left(\sum_{i=0}^n x_i \right) + x_n.$$

The map $\varphi: {}^\omega 2 \rightarrow \mathbb{R}$ defined by

$$\varphi((x_i)_{i < \omega}) = \sum_{i=0}^{\infty} \frac{x_i}{2^{i+1}} = \lim_{n \rightarrow \infty} \sum_{i=0}^n \frac{x_i}{2^{i+1}}.$$

maps the function space ${}^\omega 2$ surjectively onto the real interval

$$[0, 1] = \{r \in \mathbb{R} \mid 0 \leq r \leq 1\}.$$

Such maps are the reason that one often identifies ${}^\omega 2$ with \mathbb{R} in set theory.

5.2 Symbols and Words

Languages are mathematical objects of growing importance. Mathematical logic takes terms and formulas as mathematical material. Terms and formulas are finite sequences of symbols from some alphabet. We represent the standard symbols $=, \in$, etc. by some set-theoretical terms $\doteq, \dot{\in}$, etc. Note that details of such a formalization are highly arbitrary. One really only has to *fix* certain sets to denote certain symbols.

Definition 63. *Formalize the basic set-theoretical symbols by*

- a) $\doteq = 0, \dot{\in} = 1, \dot{\wedge} = 2, \dot{\vee} = 3, \dot{\rightarrow} = 4, \dot{\leftrightarrow} = 5, \dot{\neg} = 6, \dot{=} = 7, \dot{)} = 8, \dot{\exists} = 9, \dot{\forall} = 10$.
- b) *Variables $\dot{v}_n = (1, n)$ for $n < \omega$.*
- c) *Let $L_\in = \{\doteq, \dot{\in}, \dot{\wedge}, \dot{\vee}, \dot{\rightarrow}, \dot{\leftrightarrow}, \dot{\neg}, \dot{=}, \dot{)}, \dot{\exists}, \dot{\forall}\} \cup \{(1, n) \mid n < \omega\}$ be the alphabet of set theory.*
- d) *A word over L_\in is a finite sequence with values in L_\in .*
- e) *Let $L_\in^* = \{w \mid \exists n < \omega w: n \rightarrow L_\in\}$ be the set of all words over L_\in .*
- f) *If φ is a standard set-theoretical formula, we let $\dot{\varphi} \in L_\in^*$ denote the formalization of φ . E.g., $\dot{\exists}x = \dot{\exists}\dot{v}_0 \dot{\forall}\dot{v}_1 \dot{\neg}\dot{v}_1 \dot{\in} \dot{v}_0$ is the formalization of the set existence axiom. If the intention is clear, one often omits the formalization dots and simply writes $\exists x = \exists v_0 \forall v_1 \neg v_1 \in v_0$.*

This formalization can be developed much further, so that the notions and theorems of first-order logic are available in the theory ZF. By carrying out the definition of the axiom system ZF *within* set theory, one obtains a term $\dot{Z}F$ which represents ZF within ZF. This (quasi) self-referentiality is the basis for limiting results like the GÖDEL incompleteness theorems.

6 The von Neumann Hierarchy

We use ordinal recursion to obtain more information on the universe of all sets.

Definition 64. Define the von Neumann Hierarchy $(V_\alpha)_{\alpha \in \text{Ord}}$ by recursion:

- a) $V_0 = \emptyset$;
- b) $V_{\alpha+1} = \mathcal{P}(V_\alpha)$;
- c) $V_\lambda = \bigcup_{\alpha < \lambda} V_\alpha$ for limit ordinals λ .

We show that the von Neumann hierarchy is indeed a (fast-growing) hierarchy

Lemma 65. Let $\beta < \alpha \in \text{Ord}$. Then

- a) $V_\beta \in V_\alpha$
- b) $V_\beta \subseteq V_\alpha$
- c) V_α is transitive

Proof. We conduct the proof by a simultaneous induction on α .

$\alpha = 0$: \emptyset is transitive, thus a)-c) hold at 0.

For the *successor case* assume that a)-c) hold at α . Let $\beta < \alpha + 1$. By the inductive assumption, $V_\beta \subseteq V_\alpha$ and $V_\beta \in \mathcal{P}(V_\alpha) = V_{\alpha+1}$. Thus a) holds at $\alpha + 1$. Consider $x \in V_\alpha$. By the inductive assumption, $x \subseteq V_\alpha$ and $x \in V_{\alpha+1}$. Thus $V_\alpha \subseteq V_{\alpha+1}$. Then b) at $\alpha + 1$ follows by the inductive assumption. Now consider $x \in V_{\alpha+1} = \mathcal{P}(V_\alpha)$. Then $x \subseteq V_\alpha \subseteq V_{\alpha+1}$ and $V_{\alpha+1}$ is transitive.

For the *limit case* assume that α is a limit ordinal and that a)-c) hold at all $\gamma < \alpha$. Let $\beta < \alpha$. Then $V_\beta \in V_{\beta+1} \subseteq \bigcup_{\gamma < \alpha} V_\gamma = V_\alpha$ hence a) holds at α . b) is trivial for limit α . V_α is transitive as a union of transitive sets. \square

The V_α are nicely related to the ordinal α .

Lemma 66. For every α , $V_\alpha \cap \text{Ord} = \alpha$.

Proof. Induction on α . $V_0 \cap \text{Ord} = \emptyset \cap \text{Ord} = \emptyset = 0$.

For the *successor case* assume that $V_\alpha \cap \text{Ord} = \alpha$. $V_{\alpha+1} \cap \text{Ord}$ is transitive, and every element of $V_{\alpha+1} \cap \text{Ord}$ is transitive. Hence $V_{\alpha+1} \cap \text{Ord}$ is an ordinal, say $\delta = V_{\alpha+1} \cap \text{Ord}$. $\alpha = V_\alpha \cap \text{Ord}$ implies that $\alpha \in V_{\alpha+1} \cap \text{Ord} = \delta$ and $\alpha + 1 \leq \delta$. Assume for a contradiction that $\alpha + 1 < \delta$. Then $\alpha + 1 \in V_{\alpha+1}$ and $\alpha + 1 \subseteq V_\alpha \cap \text{Ord} = \alpha$, contradiction. Thus $\alpha + 1 = \delta = V_{\alpha+1} \cap \text{Ord}$.

For the *limit case* assume that α is a limit ordinal and that $V_\beta \cap \text{Ord} = \beta$ holds for all $\beta < \alpha$. Then

$$V_\alpha \cap \text{Ord} = \left(\bigcup_{\beta < \alpha} V_\beta \right) \cap \text{Ord} = \bigcup_{\beta < \alpha} (V_\beta \cap \text{Ord}) = \bigcup_{\beta < \alpha} \beta = \alpha.$$

\square

The foundation schema implies that the V_α -hierarchy exhausts the universe V .

Theorem 67.

- a) $\forall x \subseteq \bigcup_{\alpha \in \text{Ord}} V_\alpha \exists \beta x \subseteq V_\beta$.
- b) $V = \bigcup_{\alpha \in \text{Ord}} V_\alpha$.

Proof. a) Let $x \subseteq \bigcup_{\alpha \in \text{Ord}} V_\alpha$. Define a function $f: x \rightarrow \text{Ord}$ by

$$f(u) = \min \{ \gamma \mid u \in V_\gamma \}.$$

By the axioms of replacement and union, $\beta = \bigcup \{ f(u) + 1 \mid u \in x \} \in V$ and $\beta \in \text{Ord}$. Let $u \in x$. Then $f(u) < f(u) + 1 \leq \beta$ and $u \in V_{f(u)} \subseteq V_\beta$. Thus $x \subseteq V_\beta$.

b) Let $B = \bigcup_{\alpha \in \text{Ord}} V_\alpha$. By the schema of \in -induction it suffices to show that

$$\forall x (x \subseteq B \rightarrow x \in B).$$

So let $x \subseteq B = \bigcup_{\alpha \in \text{Ord}} V_\alpha$. By a) take β such that $x \subseteq V_\beta$. Then $x \in V_{\beta+1} \subseteq \bigcup_{\alpha \in \text{Ord}} V_\alpha = B$. \square

The V_α -hierarchy ranks the elements of V into levels.

Definition 68. Define the rank (function) $\text{rk}: V \rightarrow \text{Ord}$ by

$$x \in V_{\text{rk}(x)+1} \setminus V_{\text{rk}(x)}.$$

The rank function satisfies a recursive law.

Lemma 69. $\forall x \text{rk}(x) = \bigcup_{y \in x} \text{rk}(y) + 1$.

Proof. Let us prove the statement

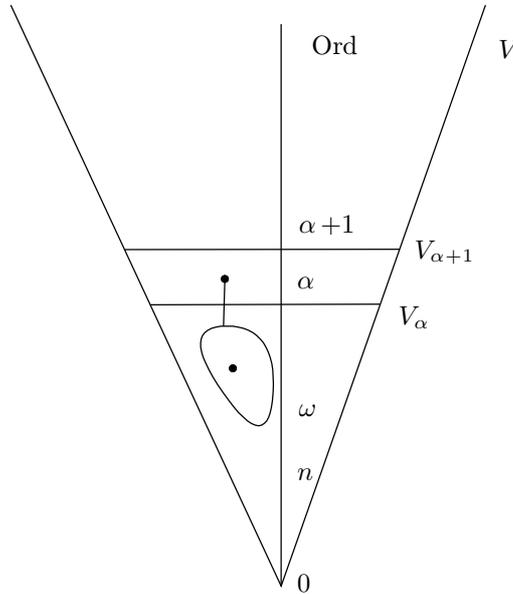
$$\forall x \in V_\alpha \text{rk}(x) = \bigcup_{y \in x} \text{rk}(y) + 1$$

by induction on α . The case $\alpha = 0$ is trivial. The limit case is obvious since $V_\lambda = \bigcup_{\alpha < \lambda} V_\alpha$ for limit λ .

For the successor case assume that the statement holds for α . Consider $x \in V_{\alpha+1}$. If $x \in V_\alpha$ the statement holds by the inductive assumption. So assume that $x \in V_{\alpha+1} \setminus V_\alpha$. Then $\text{rk}(x) = \alpha$. Let $y \in x \subseteq V_\alpha$. Then $y \in V_{\beta+1} \setminus V_\beta$ for some $\beta = \text{rk}(y) < \alpha$. $\text{rk}(y) + 1 \subseteq \alpha$. Thus $\bigcup_{y \in x} \text{rk}(y) + 1 \subseteq \alpha$. Assume that $\gamma = \bigcup_{y \in x} \text{rk}(y) + 1 < \alpha$. Let $y \in x$. Then $\text{rk}(y) + 1 \leq \gamma$ and $y \in V_{\text{rk}(y)+1} \subseteq V_\gamma$. Thus $x \subseteq V_\gamma$, $x \in V_{\gamma+1} \subseteq V_\alpha$, contradicting the assumption that $x \in V_{\alpha+1} \setminus V_\alpha$. \square

Lemma 70. Let A be a term. Then $A \in V$ iff $\exists \alpha A \subseteq V_\alpha$.

The previous analysis of the V_α -hierarchy suggest the following picture of the universe V .



7 The Axiom of Choice

Natural numbers $n \in \mathbb{N}$ are used to enumerate finite sets a as

$$a = \{a_0, a_1, \dots, a_{n-1}\}.$$

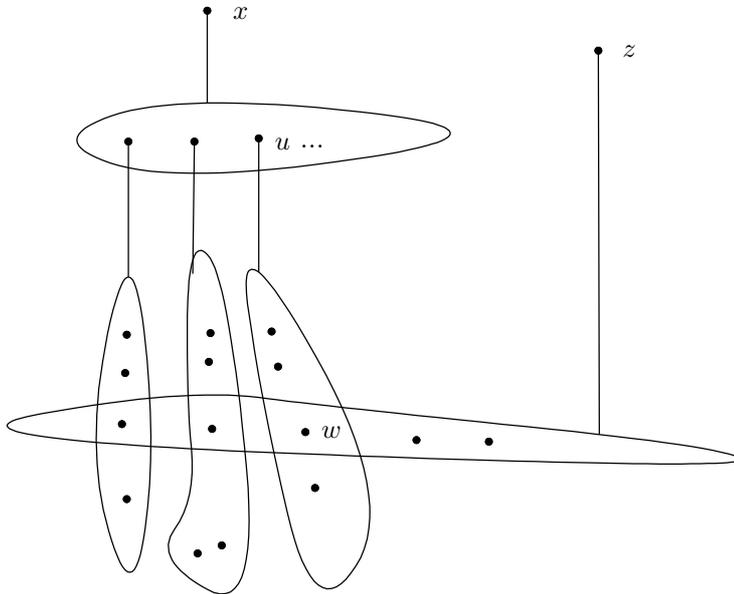
Assuming the *axiom of choice*, one can use ordinals to enumerate any set a as

$$a = \{a_i \mid i < \alpha\}.$$

Definition 71. *The Axiom of Choice, AC is the statement*

$$\forall x (\emptyset \notin x \wedge \forall u, v \in x (u \neq v \rightarrow u \cap v = \emptyset) \rightarrow \exists z \forall u \in x \exists w u \cap z = \{w\}).$$

The axiom expresses that for every set x consisting of nonempty pairwise disjoint elements there exists a choice set z , i.e., for every element $u \in x$ the intersection $u \cap z$ consists exactly of one element. Thus z “chooses” one element out of every element of x .



It seems intuitively clear that such choices are possible. On the other hand we shall see that the axiom of choice has unintuitive, paradoxical consequences.

Theorem 72. *The following statements are equivalent:*

- AC;
- $\forall x \exists g$ (g is a function with domain $x \wedge \forall u \in x (u \neq \emptyset \rightarrow g(u) \in u)$); such a function g is called a choice function for x ;
- $\forall x \exists \alpha \exists f f: \alpha \leftrightarrow x$.

Proof. $a) \rightarrow b)$ Assume AC. Let x be a set. We may assume that every element of x is nonempty. The class

$$x' = \{\{u\} \times u \mid u \in x\}$$

is the image of x under the set valued map $u \mapsto \{u\} \times u$, and thus a set by replacement. The elements $\{u\} \times u$ of x' are nonempty and pairwise disjoint. By AC, take a choice set z for x' . Define a choice function $g: x \rightarrow V$ by letting $g(u)$ be the unique element of u such that

$$(\{u\} \times u) \cap z = \{(u, g(u))\}.$$

b) \rightarrow c) Assume b). Let x be a set and let $g: \mathcal{P}(x) \setminus \{\emptyset\} \rightarrow V$ be a choice function for $\mathcal{P}(x) \setminus \{\emptyset\}$. Define a function $F: \text{Ord} \rightarrow x \cup \{x\}$ by ordinal recursion such that

$$F(\alpha) = \begin{cases} g(x \setminus F[\alpha]), & \text{if } x \setminus F[\alpha] \neq \emptyset; \\ x, & \text{if } x \setminus F[\alpha] = \emptyset. \end{cases}$$

At "time" α , the function F chooses an element $F(\alpha) \in x$ which has not been chosen before. If all elements of x have been chosen, this is signaled by F by the value x which is not an element of x .

(1) Let $\alpha < \beta$ and $F(\beta) \neq x$. Then $F(\alpha), F(\beta) \in x$ and $F(\alpha) \neq F(\beta)$.

Proof. $F(\beta) \neq x$ implies that $x \setminus F[\beta] \neq \emptyset$ and hence $F(\beta) = g(x \setminus F[\beta]) \in x \setminus F[\beta]$. Since $\alpha \in \beta$, $x \setminus F[\alpha] \neq \emptyset$ and $F(\alpha) = g(x \setminus F[\alpha]) \in x \setminus F[\alpha]$. $F(\alpha) \neq F(\beta)$ follows from $F(\beta) \in x \setminus F[\beta]$. *qed(1)*

(2) There is $\alpha \in \text{Ord}$ such that $F(\alpha) = x$.

Proof. Assume not. Then by (1), $F: \text{Ord} \rightarrow x$ is injective. Hence F^{-1} is a function and $\text{Ord} = F^{-1}[x]$. By replacement, Ord is a set, but this is a contradiction. *qed(2)*

By (2) let α be minimal such that $F(\alpha) = x$. Let $f = F \upharpoonright \alpha: \alpha \rightarrow x$. By the definition of F , $x \setminus F[\alpha] = \emptyset$, i.e., $F[\alpha] = x$ and f is surjective. By (1), f is also injective, i.e., $f: \alpha \leftrightarrow x$.

c) \rightarrow a) Assume c). Let the set x consist of nonempty pairwise disjoint elements. Apply c) to $\bigcup x$. Take an ordinal α and a function $f: \alpha \rightarrow \bigcup x$. Define a choice set z for x by setting

$$z = \{f(\xi) \mid \exists u \in x (f(\xi) \in u \wedge \forall \zeta < \xi f(\zeta) \notin u)\}.$$

So z chooses for every $u \in x$ that $f(\xi) \in u$ with ξ minimal. □

We shall later use the enumeration property c) to define the cardinality of a set. ZORN's Lemma is an important existence principle which is also equivalent to AC.

Definition 73. Let (P, \leq) be a partial order.

- a) $X \subseteq P$ is a chain in (P, \leq) if (X, \leq) is a linear order where (X, \leq) is a short notation for the structure $(X, \leq \cap X^2)$.
- b) An element $p \in P$ is an upper bound for $X \subseteq P$ iff $\forall x \in X x \leq p$.
- c) (P, \leq) is inductive iff every chain in (P, \leq) possesses an upper bound.
- d) An element $p \in P$ is a maximal element of (P, \leq) iff $\forall q \in P (q \geq p \rightarrow q = p)$.

Theorem 74. The axiom of choice is equivalent to the following principle, called Zorn's Lemma: every inductive partial order $(P, \leq) \in V$ possesses a maximal element.

Proof. Assume AC and let $(P, \leq) \in V$ be an inductive partial order. Let $g: \mathcal{P}(P) \setminus \{\emptyset\} \rightarrow V$ be a choice function for $\mathcal{P}(P) \setminus \{\emptyset\}$. Define a function $F: \text{Ord} \rightarrow P \cup \{P\}$ by ordinal recursion; if there is an upper bound for $F[\alpha]$ which is not an element of $F[\alpha]$ let

$$F(\alpha) = g(\{p \in P \setminus F[\alpha] \mid p \text{ is an upper bound for } F[\alpha]\});$$

otherwise set

$$F(\alpha) = P.$$

At "time" α , the function F chooses a strict upper bound of $F[\alpha]$ if possible. If this is not possible, this is signaled by F by the value P .

The definition of F implies immediately:

- (1) Let $\alpha < \beta$ and $F(\beta) \neq P$. Then $F(\alpha) < F(\beta)$.
- (2) There is $\alpha \in \text{Ord}$ such that $F(\alpha) = P$.

Proof. Assume not. Then by (1), $F: \text{Ord} \rightarrow P \in V$ is injective, and we get the same contradiction as in the proof of Theorem 72. *qed(2)*

By (2) let α be minimal such that $F(\alpha) = P$. By (1), $F[\alpha]$ is a chain in (P, \leq) . Since the partial order is inductive, take an upper bound p of $F[\alpha]$. We claim that p is a maximal element of (P, \leq) . Assume not and let $q \in P$, $q > p$. Then q is a strict upper bound of $F[\alpha]$ and $q \notin F[\alpha]$. But then the definition of F yields $F(\alpha) \neq P$, contradiction.

For the converse assume Zorn's Lemma and consider a set x consisting of nonempty pairwise disjoint elements. Define the set of "partial choice sets" which have empty or singleton intersection with every element of x :

$$P = \{z \subseteq \bigcup x \mid \forall u \in x (u \cap z = \emptyset \vee \exists w u \cap z = \{w\})\}.$$

P is partially ordered by \subseteq . If X is a chain in (X, \subseteq) then $\bigcup X$ is an upper bound for X . Hence (X, \subseteq) is inductive.

By Zorn's Lemma let z be a maximal element of (X, \subseteq) . We claim that z is a "total" choice set for x :

(3) $\forall u \in x \exists w u \cap z = \{w\}$.

Proof. If not, take $u \in x$ such that $u \cap z = \emptyset$. Take $w \in u$ and let $z' = z \cup \{w\}$. Then $z' \in P$, contrary to the \subseteq -maximality of z . \square

Theorem 75. *Every vector space $U \in V$ has a basis B , which is linearly independent and spans U .*

Proof. Let U be a vector space with scalar field K . Let

$$P = \{b \subseteq U \mid b \text{ is linearly independent in } U\}.$$

We shall apply Zorn's lemma to the partial order (P, \subseteq) .

(1) (P, \subseteq) is inductive.

Proof. Let $X \subseteq P$ be a chain. Let $c = \bigcup X \subseteq U$. We show that c is linearly independent. Consider a linear combination

$$k_0 \cdot v_0 + \dots + k_{n-1} \cdot v_{n-1} = 0,$$

where $v_0, \dots, v_{n-1} \in c$ and $k_0, \dots, k_{n-1} \in K$. Take $b_0, \dots, b_{n-1} \in X$ such that $v_0 \in b_0, \dots, v_{n-1} \in b_{n-1}$. Since X is a chain there is some $b_i, i < n$ such that $b_0, \dots, b_{n-1} \subseteq b_i$. Then $v_0, \dots, v_{n-1} \in b_i$. Since $b_i \in P$ is linearly independent, $k_0 = \dots = k_{n-1} = 0$. *qed*(1)

By Zorn's lemma, (P, \subseteq) has a maximal element, say B . B is linearly independent since $B \in P$.

(2) B spans U .

Proof. Let $v \in U$. If $v \in B$ then v is in the span of B . So consider the case that $v \notin B$. Then $B \cup \{v\}$ is a proper superset of B . By the \subseteq -maximality of B , $B \cup \{v\}$ is linearly dependent. So there is a non-trivial linear combination

$$k_0 \cdot v_0 + \dots + k_{n-1} \cdot v_{n-1} + k \cdot v = 0,$$

where $v_0, \dots, v_{n-1} \in B$ and at least one of the coefficients $k_0, \dots, k_{n-1}, k \in K$ is non-zero. If $k = 0$,

$$k_0 \cdot v_0 + \dots + k_{n-1} \cdot v_{n-1} = 0$$

would be a non-trivial representation of 0, contradicting that B is linearly independent. Hence $k \neq 0$ and

$$v = -\frac{k_0}{k} \cdot v_0 - \dots - \frac{k_{n-1}}{k} \cdot v_{n-1}.$$

So v is in the span of B . \square

Actually one can show the converse of this Theorem: if every vector space has a basis, then AC holds.

As another application of Zorn's lemma we consider *filters* which are collections of "large" subsets of some domain.

Definition 76. *Let Z be a set. We say that F is a filter on Z if*

- a) $F \subseteq \mathcal{P}(Z)$;
- b) $\emptyset \notin F$ and $Z \in F$;
- c) $X \in F$ and $X \subseteq Y \subseteq Z$ implies that $Y \in F$;
- d) $X, Y \in F$ implies that $X \cap Y \in F$.

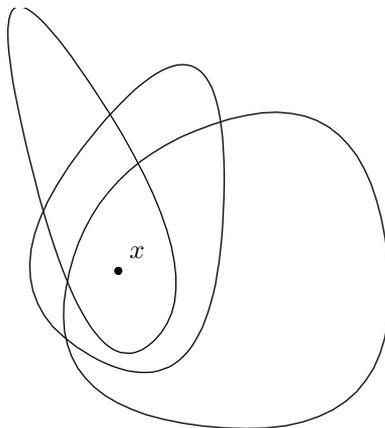
If moreover

$$X \subseteq Z \rightarrow X \in F \vee (Z \setminus X) \in F$$

we call F an ultrafilter on Z .

Important examples of filters are *neighbourhood filters* N_x of points x in some topological space (Z, \mathcal{T}) :

$$N_x = \{U \subseteq Z \mid U \text{ is a neighbourhood of } x\}.$$



A combinatorial example is the *Frechet filter* on ω :

$$F = \{X \subseteq \omega \mid \exists n \in \omega \forall m \in \omega (m > n \rightarrow m \in X)\}.$$

The expression “ $A(n)$ holds for almost all $n \in \omega$ ” is equivalent to

$$\{n \in \omega \mid A(n)\} \in F.$$

Theorem 77. *Let F be a filter on the set Z . Then there is an extension $G \supseteq F$ such that G is an ultrafilter on Z .*

Proof. Let

$$P = \{H \subseteq \mathcal{P}(Z) \mid H \text{ is a filter on } Z \text{ and } H \supseteq F\}.$$

We shall apply Zorn’s lemma to the partial order (P, \subseteq) .

(1) (P, \subseteq) is inductive.

Proof. Let $C \subseteq P$ be a chain. Let $H' = \bigcup X \subseteq \mathcal{P}(Z)$. We show that H' is a filter on Z . Trivially $\emptyset \notin H'$. Consider $X \in H'$ and $X \subseteq Y \subseteq Z$. Then $X \in H$ for some $H \in C$. Since H is a filter, $X \in H$ and so $Y \in H \subseteq H'$.

For the closure under intersections consider $X, Y \in H'$. Then $X \in H_0$ for some $H_0 \in C$, and $Y \in H_1$ for some $H_1 \in C$. Since C is a chain, we have, wlog, that $H_0 \subseteq H_1$. Then $X, Y \in H_1$, and $X \cap Y \in H_1 \subseteq H'$. *qed*(1)

By Zorn’s lemma, let $G \in P$ be a maximal element. Then G is a filter which extends F .

(2) G is an ultrafilter on Z .

Proof. Consider $X_0 \subseteq Z$. Assume that $X_0 \notin G$.

Claim. There exists $X_1 \in G$ such that $X_1 \cap X_0 = \emptyset$.

Proof. Assume instead that $X \cap X_0 \neq \emptyset$ for all $X \in G$.

Define

$$G' = \{Y \subseteq Z \mid \exists X \in G Y \supseteq X \cap X_0\}.$$

G' is a filter on Z ; we only check Definition 76, d): let $Y_1, Y_2 \in G'$ with $Y_1 \supseteq X_1 \cap X_0$ and $Y_2 \supseteq X_2 \cap X_0$ where $X_1, X_2 \in G$. Then $Y_1 \cap Y_2 \supseteq (X_1 \cap X_2) \cap X_0$ where $X_1 \cap X_2 \in G$, and so $Y_1 \cap Y_2 \in G'$.

Obviously $G' \supseteq G \supseteq F$ and $G' \neq G$ since $X_0 \in G'$ and $X_0 \notin G$. This contradicts the maximality of G in (P, \subseteq) . *qed*(*Claim*)

So take $X_1 \in G$ such that $X_1 \cap X_0 = \emptyset$ for some $X_1 \in G$. Then $X_1 \subseteq Z \setminus X_0$ and hence $Z \setminus X_0 \in G$. \square

Definition 78. *The axiom system ZFC consists of the ZF-axioms together with the axiom of choice AC.*

The system ZFC is usually taken as the foundation of mathematics. The ZF axioms have a good intuitive motivations. The axiom of choice is more controversial; AC has desirable consequences like Zorn's Lemma and its applications, but on the other hand AC has some paradoxical and problematic consequences. The status of AC within set theory can be compared to the parallel axiom in geometry. Similar to the situation in (non-)euclidean geometry one can show that if there is a model of the ZF axioms then there is a model of ZFC.

Exercise 25. Show that in the theory ZF the axiom of choice is equivalent to the *Hausdorff Maximality Principle* which says: for every partial order $(P, \leq) \in V$ there is an inclusion maximal chain X in (P, \leq) , i.e., if $Y \supseteq X$ is a chain in (P, \leq) then $Y = X$. [Hausdorff, Grundzüge der Mengenlehre, p. 141: *Wir haben damit für eine teilweise geordnete Menge A die Existenz größter geordneter Teilmengen B bewiesen; natürlich kann es deren verschiedene geben.*]

Whereas the axiom of choice has “desirable” consequences like the existence of bases and ultrafilters, there are also a number of counterintuitive and perhaps even undesirable consequences. We present one, concerning non-Lebesgue-measurable sets.

Recall that the 1-dimensional Lebesgue measure on \mathbb{R} is a function $\mu: \Omega \rightarrow \mathbb{R}^+ \cup \{0\} \cup \{\infty\}$ taking values in the extended non-negative real line with the properties:

- a) $\Omega \subseteq \mathcal{P}(\mathbb{R})$ contains all intervals and is closed under complements and countable unions;
- b) $\mu([0, 1]) = 1$;
- c) μ is *countably additive* (σ -additive): if $\{X_i \mid i < \omega\} \subseteq \Omega$ is a pairwise disjoint family then

$$\mu\left(\bigcup_{i < \omega} X_i\right) = \sum_{i < \omega} \mu(X_i);$$

- d) μ is *translation invariant*: if $X \in \Omega$ and $d \in \mathbb{R}$ then $X + d = \{x + d \mid x \in X\} \in \Omega$ and $\mu(X + d) = \mu(X)$.

Theorem 79. *There is a set $Z \subseteq [0, 1]$ which is not Lebesgue-measurable, i.e., $Z \notin \Omega$.*

Proof. Let

$$A = \{\mathbb{Q} + d \mid d \in \mathbb{R}\}.$$

(1) A consists of pairwise disjoint nonempty sets which intersect the interval $[0, 1]$.

Proof. Assume that $x \in (\mathbb{Q} + d) \cap (\mathbb{Q} + e)$. Take rational numbers $r, s \in \mathbb{Q}$ such that

$$x = r + d = s + e.$$

Then $d = (s - r) + e \in \mathbb{Q} + e$ and

$$\mathbb{Q} + d = \{t + d \mid t \in \mathbb{Q}\} = \{t + (s - r) + e \mid t \in \mathbb{Q}\} \subseteq \mathbb{Q} + e.$$

Similarly $\mathbb{Q} + e \subseteq \mathbb{Q} + d$ and so $\mathbb{Q} + e = \mathbb{Q} + d$.

Consider $d \in \mathbb{R}$. Take an integer $z \in \mathbb{Z}$ such that $z \leq d < z + 1$. Then

$$-z + d \in (\mathbb{Q} + d) \cap [0, 1)$$

qed(1)

By the axiom of choice let Z be a choice set for the set

$$\{(\mathbb{Q} + d) \cap [0, 1) \mid d \in \mathbb{R}\}.$$

(2) If $q, r \in \mathbb{Q}$ and $q \neq r$ then $(Z + q) \cap (Z + r) = \emptyset$.

Proof. Assume not, and take $z_0, z_1 \in Z$ such that $z_0 + q = z_1 + r$. Then $z_0 \in \mathbb{Q} + z_0$ and $z_1 \in \mathbb{Q} + z_0$. Since Z is a choice set, $z_0 = z_1$. But then $q = r$. Contradiction. *qed(2)*

(3) $[0, 1] \subseteq \bigcup_{q \in [-2, 2] \cap \mathbb{Q}} Z + q$.

Proof. Let $d \in [0, 1]$. Let $z \in Z \cap (\mathbb{Q} + d) \cap [0, 1)$. Take $q \in \mathbb{Q}$ such that $z = q + d$. Then $d = z + (-q)$ where $|q| \leq 2$. *qed(3)*

Assume now that $Z \in \Omega$. Since $Z \subseteq [0, 1]$ we have $\mu(Z) \leq 1$.

Case 1: $\mu(Z) = 0$. Then

$$1 = \mu([0, 1]) \leq \mu\left(\bigcup_{q \in [-2, 2] \cap \mathbb{Q}} Z + q\right) = \sum_{q \in [-2, 2] \cap \mathbb{Q}} \mu(Z + q) = \sum_{q \in [-2, 2] \cap \mathbb{Q}} \mu(Z) = \sum_{q \in [-2, 2] \cap \mathbb{Q}} 0 = 0,$$

contradiction.

Case 2: $\mu(Z) = \varepsilon > 0$. Then

$$\mu\left(\bigcup_{q \in [-2, 2] \cap \mathbb{Q}} Z + q\right) \leq \mu([-2, 3]) = 5$$

but on the other hand

$$\mu\left(\bigcup_{q \in [-2, 2] \cap \mathbb{Q}} Z + q\right) = \sum_{q \in [-2, 2] \cap \mathbb{Q}} \mu(Z + q) = \sum_{q \in [-2, 2] \cap \mathbb{Q}} \mu(Z) = \sum_{q \in [-2, 2] \cap \mathbb{Q}} \varepsilon = \infty.$$

□

A non-measurable set constructed like the set Z in this proof is called a Vitali-set. Starting from a certain model of ZFC, one can construct a model of the theory ZF in which every set of reals is Lebesgue-measurable.

8 Wellfounded Relations

The axiom schema of foundation yields an induction theorem for the \in -relation, and in the previous section we have seen a recursive law for the rank-function. We generalize these techniques to *wellfounded* relations. The results of this section do not require the axiom of choice.

Definition 80. *Let R be a relation on a domain D .*

a) *R is wellfounded, iff for all terms A*

$$\emptyset \neq A \wedge A \subseteq D \rightarrow \exists x \in A \ A \cap \{y \mid yRx\} = \emptyset.$$

b) *R is strongly wellfounded iff it is wellfounded and*

$$\forall x \in D \ \{y \in D \mid yRx\} \in V.$$

c) *R is a wellorder iff R is a wellfounded strict linear order.*

d) *R is a strong wellorder iff R is a strongly wellfounded wellorder.*

By the scheme of foundation, the \in -relation is strongly wellfounded. The ordinals are strongly wellordered by $<$. There are wellfounded relations which are *not* strongly wellfounded: e.g., let $R \subseteq \text{Ord} \times \text{Ord}$,

$$xRy \text{ iff } (x \neq 0 \wedge y \neq 0 \wedge x < y) \vee (y = 0 \wedge x \neq 0),$$

be a rearrangement of $(\text{Ord}, <)$ with 0 put on top of all the other ordinals.

For strongly wellfounded relations, every element is contained in a *set-sized* initial segment of the relation.

Lemma 81. *Let R be a strongly wellfounded relation on D . Then*

$$\forall x \subseteq D \exists z (z \subseteq D \wedge x \subseteq z \wedge \forall u \in z \forall v Ru \ v \in z).$$

Moreover for all $x \subseteq D$, the R -transitive closure

$$\text{TC}_R(x) = \bigcap \{z \mid z \subseteq D \wedge x \subseteq z \wedge \forall u \in z \forall v Ru \ v \in z\}$$

of x is a set. In case R is the \in -relation, we write $\text{TC}(x)$ instead of $\text{TC}_\in(x)$.

Proof. We prove by R -induction that

$$\forall x \in D \ \text{TC}_R(\{x\}) \in V.$$

So let $x \in D$ and $\forall y R x \text{TC}_R(\{y\}) \in V$. Then

$$z = \{x\} \cup \bigcup_{y R x} \text{TC}_R(\{y\}) \in V$$

by replacement. z is a subset of D and includes $\{x\}$. z is R -closed, i.e., closed with respect to R -predecessors: each $\text{TC}_R(\{y\})$ is R -closed, and if $y R x$ then $y \in \{y\} \subseteq \text{TC}_R(\{y\}) \subseteq z$. So $\text{TC}_R(\{x\})$ is the intersection of a non-empty class, hence a set.

Finally observe that we may set

$$\text{TC}_R(x) = \bigcup_{y \in x} \text{TC}_R(\{y\}).$$

□

Exercise 26. Show that for an ordinal α , $\text{TC}(\alpha) = \alpha$ and $\text{TC}(\{\alpha\}) = \alpha + 1$.

For *strongly* wellfounded relations, the following recursion theorem holds:

Theorem 82. Let R be a *strongly wellfounded* relation on D . Let $G: V \rightarrow V$. Then there is a canonical class term F , given by the subsequent proof, such that

$$F: D \rightarrow V \text{ and } \forall x \in D F(x) = G(F \upharpoonright \{y \mid y R x\}).$$

We then say that F is defined by R -recursion with the recursion rule G . F is unique in the sense that if another term F' satisfies

$$F': D \rightarrow V \text{ and } \forall \alpha \in D F'(x) = G(F' \upharpoonright \{y \mid y R x\})$$

then $F = F'$.

Proof. We proceed as in the ordinal recursion theorem. Let

$$\tilde{F} := \{f \mid \exists z \subseteq D (\forall x \in z \{y \mid y R x\} \subseteq z, f: z \rightarrow V \text{ and } \forall x \in z f(x) = G(f \upharpoonright \{y \mid y R x\}))\}$$

be the class of all *approximations* to the desired function F .

(1) Let $f, g \in \tilde{F}$. Then f, g are *compatible*, i.e., $\forall x \in \text{dom}(f) \cap \text{dom}(g) f(x) = g(x)$.

Proof. By induction on R . Let $x \in \text{dom}(f) \cap \text{dom}(g)$ and assume that $\forall y R x f(y) = g(y)$. Then $f \upharpoonright \{y \mid y R x\} = g \upharpoonright \{y \mid y R x\}$

$$f(x) = G(f \upharpoonright \{y \mid y R x\}) = G(g \upharpoonright \{y \mid y R x\}) = g(x).$$

qed(1)

By the compatibility of the approximation functions the union

$$F = \bigcup \tilde{F}$$

is a function defined on $\text{dom}(F) \subseteq D$. $\text{dom}(F)$ is R -closed since the domain of every approximation is R -closed.

(2) $\forall x \in \text{dom}(F) (\{y \mid y R x\} \subseteq \text{dom}(F) \wedge F(x) = G(F \upharpoonright \{y \mid y R x\}))$.

Proof. Let $x \in \text{dom}(F)$. Take some approximation $f \in \tilde{F}$ such that $x \in \text{dom}(f)$. Then $\{y \mid y R x\} \subseteq \text{dom}(f) \subseteq \text{dom}(F)$ and

$$F(x) = f(x) = G(f \upharpoonright \{y \mid y R x\}) = G(F \upharpoonright \{y \mid y R x\}).$$

qed(2)

(3) $D = \text{dom}(F)$.

Proof. We show by R -induction that $\forall x \in D x \in \text{dom}(F)$. Let $x \in D$ and assume that $\forall y R x y \in \text{dom}(F)$. $\text{TC}_R(\{y \mid y R x\}) \subseteq \text{dom}(F)$ since $\text{dom}(F)$ is R -closed. Then

$$f = (F \upharpoonright \text{TC}_R(\{y \mid y R x\})) \cup \{(x, G(F \upharpoonright \{y \mid y R x\}))\}$$

is an approximation with $x \in \text{dom}(f)$, and so $x \in \text{dom}(F)$. □

Exercise 27. Define set theoretic operations

$$x + y = x \cup \{x + z \mid z \in y\}$$

and

$$x \cdot y = \bigcup_{z \in y} (x \cdot z + x)$$

and study their arithmetic/algebraic properties. Show that they extend ordinal arithmetic.

Theorem 83. *Let R be a strongly wellfounded relation on D and suppose that R is extensional, i.e., $\forall x, y \in D (\forall u (uRx \leftrightarrow uRy) \rightarrow x = y)$. Then there is a transitive class \bar{D} and an isomorphism $\pi: (D, R) \leftrightarrow (\bar{D}, \in)$. \bar{D} and π are uniquely determined by R and D , they are called the MOSTOWSKI-collapse of R and D .*

Proof. Define $\pi: D \rightarrow V$ by R -recursion with

$$\pi(x) = \{\pi(y) \mid yRx\}.$$

Let $\bar{D} = \text{rng}(\pi)$.

(1) \bar{D} is transitive.

Proof. Let $\pi(x) \in \bar{D}$ and $u \in \pi(x) = \{\pi(y) \mid yRx\}$. Let $u = \pi(y)$, yRx . Then $u \in \text{rng}(\pi) = \bar{D}$. *qed(1)*

(2) π is injective.

Proof. We prove by \in -induction that every $z \in \bar{D}$ has exactly one preimage under π . So let $z \in \bar{D}$ and let this property be true for all elements of z . Assume that $x, y \in D$ and $\pi(x) = \pi(y) = z$. Let uRx . Then $\pi(u) \in \pi(x) = \pi(y) = \{\pi(v) \mid vRy\}$. Take vRy such that $\pi(u) = \pi(v)$. By the inductive assumption, $u = v$, and uRy . Thus $\forall u (uRx \rightarrow uRy)$. By symmetry, $\forall u (uRy \leftrightarrow uRx)$. Since R is extensional, $x = y$. So z has exactly one preimage under π . *qed(2)*

(3) π is an isomorphism, i.e., π is bijective and $\forall x, y \in D (xRy \leftrightarrow \pi(x) \in \pi(y))$.

Proof. Let $x, y \in D$. If xRy then $\pi(x) \in \{\pi(u) \mid uRy\} = \pi(y)$. Conversely, if $\pi(x) \in \{\pi(u) \mid uRy\} = \pi(y)$ then let $\pi(x) = \pi(u)$ for some uRy . Since π is injective, $x = u$ and xRy . *qed(3)*

Uniqueness of the collapse \bar{D} and π is given by the next theorem. \square

Theorem 84. *Let X and Y be transitive and let $\sigma: X \leftrightarrow Y$ be an \in - \in -isomorphism between X and Y , i.e., $\forall x, y \in X (x \in y \leftrightarrow \sigma(x) \in \sigma(y))$. Then $\sigma = \text{id} \upharpoonright X$ and $X = Y$.*

Proof. We show that $\sigma(x) = x$ by \in -induction over X . Let $x \in X$ and assume that $\forall y \in x \sigma(y) = y$.

Let $y \in x$. By induction assumption, $y = \sigma(y) \in \sigma(x)$. Thus $x \subseteq \sigma(x)$.

Conversely, let $v \in \sigma(x)$. Since $Y = \text{rng}(\sigma)$ is transitive take $u \in X$ such that $v = \sigma(u)$. Since σ is an isomorphism, $u \in x$. By induction assumption, $v = \sigma(u) = u \in x$. Thus $\sigma(x) \subseteq x$. \square

If R is a well-order on D then R is obviously extensional. We study the Mostowski collapse of strongly well-ordered relations.

Theorem 85. *Let R be a strongly well-ordered relation on D . Let $\pi: (D, R) \leftrightarrow (\bar{D}, \in)$ be the MOSTOWSKI-collapse of R and D . If D is a proper class then $\bar{D} = \text{Ord}$. If D is a set then \bar{D} is an ordinal which is called the ordertype of (D, R) . We then write $\bar{D} = \text{otp}(D, R)$.*

Proof. \bar{D} is transitive since it is a Mostowski collapse.

(1) Every element of \bar{D} is transitive.

Proof. Let $x \in y \in z \in \bar{D}$. Since \bar{D} is transitive, $x, y, z \in \bar{D}$ and there are $a, b, c \in D$ such that $x = \pi(a)$, $y = \pi(b)$, and $z = \pi(c)$. Since π is an order-isomorphism, $aRbRc$. Since R is a transitive relation, aRc . This implies $x \in z$. *qed(1)*

(2) Every element of \bar{D} is an ordinal.

Proof. Let $z \in \bar{D}$. z is transitive, and it remains to show that every element of z is transitive. Let $y \in z$. Then $y \in \bar{D}$ and so y is transitive by (1). *qed(2)*

Consider the case that D is a proper class. Then \bar{D} is a proper class of ordinals. \bar{D} must be unbounded in the ordinals, since it would be a set otherwise. By transitivity, every ordinal which is smaller than some element of \bar{D} is an element of \bar{D} . Hence $\bar{D} = \text{Ord}$.

If D is a set, then \bar{D} is a transitive set, and by (1), $\bar{D} \in \text{Ord}$. \square

By Lemma 84, any order-isomorphism $\sigma: (\alpha, <) \leftrightarrow (\beta, <)$ between ordinals must be the identity. So the ordertype of a set-sized well-order (D, R) is the *unique* ordinal, to which it is order-isomorphic.

Lemma 86. *Let $x \subseteq \alpha \in \text{Ord}$. Then $(x, <)$ is a well-order. Let $\pi: (x, <) \leftrightarrow (\text{otp}(x, <), <)$ be the Mostowski collapse of $(x, <)$. Then $\forall \xi \in x \ \xi \geq \pi(\xi)$ and $\text{otp}(x, <) \leq \alpha$.*

Proof. By induction on $\xi \in x$. Let $\delta \in \pi(\xi) = \{\pi(\zeta) \mid \zeta \in x \wedge \zeta < \xi\}$. Let $\delta = \pi(\zeta)$ with $\zeta \in x \wedge \zeta < \xi$. By induction $\delta = \pi(\zeta) \leq \zeta < \xi$. Thus $\pi(\xi) \subseteq \xi$ and $\pi(\xi) \leq \xi$. Similarly consider $\delta \in \text{otp}(x, <) = \{\pi(\zeta) \mid \zeta \in x\}$. Let $\delta = \pi(\zeta)$ with $\zeta \in x$. Then $\delta = \pi(\zeta) \leq \zeta < \alpha$. Thus $\text{otp}(x, <) \subseteq \alpha$. \square

9 Cardinalities

Apart from its foundational role, set theory is mainly concerned with the study of arbitrary infinite sets and in particular with the question of their size. Cantor's approach to infinite sizes follows naive intuitions familiar from finite sets of objects.

Definition 87.

- a) x and y are equipollent, or equipotent, or have the same cardinality, written $x \sim y$, if $\exists f f: x \leftrightarrow y$.
- b) x has cardinality at most that of y , written $x \preceq y$, if $\exists f f: x \rightarrow y$ is injective.
- c) We write $x \prec y$ for $x \preceq y$ and $x \approx y$.

These relations are easily shown to satisfy

Lemma 88. *Assume ZF. Then*

- a) \sim is an equivalence relation on V .
- b) $x \sim y \rightarrow x \preceq y \wedge y \preceq x$.
- c) $x \preceq x$.
- d) $x \preceq y \wedge y \preceq z \rightarrow x \preceq z$.
- e) $x \subseteq y \rightarrow x \preceq y$.

The converse of b) is also true and proved in an exercise.

Theorem 89. (Cantor - Bernstein) $x \preceq y \wedge y \preceq x \rightarrow x \sim y$.

Assuming the axiom of choice, every set is equipollent with an ordinal (Theorem 72 c). One can take the minimal such ordinal as the canonical representative of the equivalence class with respect to \sim .

Definition 90.

- a) $\text{card}(x) = \min \{\alpha \mid \exists f f: \alpha \leftrightarrow x\}$ is the cardinality of the set x . One also writes $\bar{x} = \text{card}(x)$.
- b) An ordinal κ is a cardinal iff it $\kappa = \text{card}(x)$ for some set x .
- c) Let $\text{Cd} = \{\kappa \in \text{Ord} \mid \kappa \text{ is a cardinal}\}$ be the class of all cardinals, and let $\text{Card} = \{\kappa \geq \omega \mid \kappa \text{ is a cardinal}\}$ be the class of infinite cardinals.

Let us assume AC until further notice. Then Cantor's two approaches to cardinality agree.

Theorem 91.

- a) $x \preceq y \leftrightarrow \text{card}(x) \leq \text{card}(y)$.

b) $x \sim y \leftrightarrow \text{card}(x) = \text{card}(y)$.

Proof. a) Let $x \preceq y$ and let $f: x \rightarrow y$ be injective. Further let $f_x: \text{card}(x) \leftrightarrow x$ and $f_y: \text{card}(y) \leftrightarrow y$. Then $f_y^{-1} \circ f \circ f_x: \text{card}(x) \rightarrow \text{card}(y)$ is injective. Let $z = f_y^{-1} \circ f \circ f_x[\text{card}(x)] \subseteq \text{card}(y)$. Then $\text{card}(x) = \text{card}(z) \leq \text{otp}(z) \leq \text{card}(y)$.

Conversely, let $\text{card}(x) \leq \text{card}(y)$ with $f_x: \text{card}(x) \leftrightarrow x$ and $f_y: \text{card}(y) \leftrightarrow y$ as above. Then $f_y \circ f_x^{-1}: x \rightarrow y$ is injective and $x \preceq y$.

b) is trivial. \square

As an immediate corollary we get the Cantor–Schröder–Bernstein theorem with AC.

Theorem 92. (ZFC) *Let $a \preceq b$ and $b \preceq a$. Then $a \sim b$.*

We shall now explore “small” cardinals. Below ω , the notions of natural number, ordinal number and cardinal number agree.

Theorem 93. *For all natural numbers $n < \omega$ holds*

a) $\text{card}(n) = n$;

b) $n \in \text{Cd}$.

Proof. a) By complete induction on n .

For $n = 0$, $\emptyset: 0 \leftrightarrow 0$ and hence $\text{card}(0) = 0$.

Assume that $\text{card}(n) = n$. We claim that $\text{card}(n + 1) = n + 1$. Obviously $\text{card}(n + 1) \leq n + 1$.

Assume for a contradiction that $m = \text{card}(n + 1) < n + 1$. Take $f: m \leftrightarrow n + 1$. Let $f(i_0) = n$.

Case 1: $i_0 = m - 1$. Then $f \upharpoonright (m - 1): (m - 1) \leftrightarrow n$ and $\text{card}(n) \leq m - 1 < n$, contradiction.

Case 2: $i_0 < m - 1$. Then define $g: (m - 1) \leftrightarrow n$ by

$$g(i) = \begin{cases} f(i), & \text{if } i \neq i_0; \\ f(m - 1), & \text{if } i = i_0. \end{cases}$$

Hence $\text{card}(n) \leq m - 1 < n$, contradiction.

b) follows immediately from a). \square

Theorem 94.

a) $\text{card}(\omega) = \omega$;

b) $\omega \in \text{Card}$.

Proof. Assume for a contradiction that $n = \text{card}(\omega) < \omega$. Let $f: n \leftrightarrow \omega$. Define $g: (n - 1) \rightarrow \omega$ by

$$g(i) = \begin{cases} f(i), & \text{if } f(i) < f(n - 1), \\ f(i) - 1, & \text{if } f(i) > f(n - 1). \end{cases}$$

(1) g is injective.

Proof. Let $i < j < n - 1$.

Case 1. $f(i), f(j) < f(n - 1)$. Then $g(i) = f(i) \neq f(j) = g(j)$.

Case 2. $f(i) < f(n - 1) < f(j)$. Then $g(i) = f(i) < f(n - 1) \leq f(j) - 1 = g(j)$.

Case 3. $f(j) < f(n - 1) < f(i)$. Then $g(j) = f(j) < f(n - 1) \leq f(i) - 1 = g(i)$.

Case 4. $f(n - 1) < f(i), f(j)$. Then $g(i) = f(i) - 1 \neq f(j) - 1 = g(j)$. *qed(1)*

(2) g is surjective.

Proof. Let $k \in \omega$.

Case 1. $k < f(n - 1)$. By the bijectivity of f take $i < n - 1$ such that $f(i) = k$. Then $g(i) = f(i) = k$.

Case 2. $k \geq f(n - 1)$. By the bijectivity of f take $i < n - 1$ such that $f(i) = k + 1$. Then $g(i) = f(i) - 1 = k$. *qed(2)*

But this is a contradiction to the supposed minimality of $n = \text{card}(\omega)$. \square

Lemma 95.

a) $\text{card}(\omega + 1) = \omega$.

b) $\text{card}(\omega + \omega) = \omega$.

c) $\text{card}(\omega \cdot \omega) = \omega$.

Proof. a) Define $f_a: \omega \leftrightarrow \omega + 1$ by

$$f(n) = \begin{cases} \omega, & \text{if } n = 0 \\ n - 1, & \text{else} \end{cases}$$

b) Define $f_b: \omega \leftrightarrow \omega + \omega$ by

$$f(n) = \begin{cases} m, & \text{if } n = 2 \cdot m \\ \omega + m, & \text{if } n = 2 \cdot m + 1 \end{cases}$$

c) Define $f_c: \omega \leftrightarrow \omega \cdot \omega$ by

$$f(n) = \omega \cdot k + l, \text{ if } n = 2^k \cdot (2 \cdot l + 1) - 1$$

□

10 Finite, countable, uncountable sets

Definition 96.

- a) x is finite if $\text{card}(x) < \omega$.
- b) x is infinite if x is not finite.
- c) x is countable if $\text{card}(x) \leq \omega$.
- d) x is countably infinite if $\text{card}(x) = \omega$.
- e) x is uncountable if x is not countable.

10.1 Finite sets

We have the following closure properties for finite sets:

Theorem 97. Let a, b finite, let $x \in V$.

- a) Every subset of a finite set is finite.
- b) $a \cup \{x\}$, $a \cup b$, $a \cap b$, $a \times b$, $a \setminus b$, and $\mathcal{P}(a)$ are finite. We have $\text{card}(\mathcal{P}(a)) = 2^{\text{card}(a)}$.
- c) If a_i is finite for $i \in b$ then $\bigcup_{i < b} a_i$ is finite.

Proof. Easy. □

Finite sets can be distinguished by dependencies between injective and surjective maps.

Theorem 98. Let a be finite. Then

- a) $\forall f \left(f: a \xrightarrow{\text{inj.}} a \text{ implies } f: a \xrightarrow{\text{surj.}} a \right)$
- b) $\forall f \left(f: a \xrightarrow{\text{surj.}} a \text{ implies } f: a \xrightarrow{\text{inj.}} a \right)$

Using the axiom of choice one can also show the converse.

Theorem 99. Let a be infinite. Then

- a) $\exists f: \omega \xrightarrow{\text{inj.}} a$.
- b) $\exists f \left(f: a \xrightarrow{\text{inj.}} a \text{ and } \neg f: a \xrightarrow{\text{surj.}} a \right)$
- c) $\exists f \left(f: a \xrightarrow{\text{surj.}} a \text{ and } \neg f: a \xrightarrow{\text{inj.}} a \right)$

This yields:

Theorem 100. For $a \in V$ the following statements are equivalent:

- a) a is finite;
- b) $\forall f \left(f: a \xrightarrow{\text{inj.}} a \text{ implies } f: a \xrightarrow{\text{surj.}} a \right)$;
- c) $\forall f \left(f: a \xrightarrow{\text{surj.}} a \text{ implies } f: a \xrightarrow{\text{inj.}} a \right)$.

If one does not assume the axiom of choice, one can use b) or c) to define the notion of finiteness.

10.2 Countable sets

We have the following closure properties for countable sets:

Theorem 101. Let a, b countable, let $x \in V$.

- a) Every subset of a countable set is countable
- b) $a \cup \{x\}$, $a \cup b$, $a \cap b$, $a \times b$, $a \setminus b$ are countable
- c) If a_n is countable for $n < \omega$ then $\bigcup_{n < \omega} a_n$ is countable

Proof. Countability will be shown by exhibiting injections into countable sets. Then a) is trivial.

b) Let $f_a: a \rightarrow \omega$ and $f_b: b \rightarrow \omega$ be injective. Then define injective maps:

$$f_0: a \cup \{x\} \rightarrow \omega, f_0(u) = \begin{cases} f_a(u) + 1, & \text{if } u \in a \\ 0, & \text{else} \end{cases}$$

$$f_1: a \cup b \rightarrow \omega, f_1(u) = \begin{cases} 2 \cdot f_a(u) + 1, & \text{if } u \in a \\ 2 \cdot f_b(u), & \text{else} \end{cases}$$

$$f_2: a \times b \rightarrow \omega, f_2(u, v) = 2^{f_a(u)} \cdot (2 \cdot f_b(v) + 1)$$

c) By the axiom of choice choose a sequence $(h_n | n < \omega)$ of injections $h_n: a_n \rightarrow \omega$. Define

$$f_3: \bigcup_{n < \omega} a_n \rightarrow \omega, f_3(u) = 2^n \cdot (2 \cdot h_n(u) + 1), \text{ where } n \text{ is minimal such that } u \in a_n.$$

□

10.3 Uncountable sets

Theorem 102. (Cantor) $x < \mathcal{P}(x)$

Proof. $\text{card}(x) \leq \text{card}(\mathcal{P}(x))$ is clear. Assume that $\text{card}(x) = \text{card}(\mathcal{P}(x))$ and let $f: x \leftrightarrow \mathcal{P}(x)$ be bijective. Define

$$a = \{u \in x | u \notin f(u)\} \subseteq x.$$

Let $a = f(u_0)$. Then

$$u_0 \in f(u_0) \leftrightarrow u_0 \in a \leftrightarrow u_0 \notin f(u_0).$$

Contradiction. Hence $\text{card}(x) < \text{card}(\mathcal{P}(x))$. □

Theorem 103. $\aleph := \text{card}(\mathcal{P}(\omega))$ is an uncountable cardinal.

Note that by previous exercises or lemmas we have

$$\text{card}(\mathcal{P}(\omega)) = \text{card}(\mathbb{R}) = \text{card}(2^\omega) = \text{card}({}^\omega\omega)$$

Cantor spent a lot of effort on determining the size of \aleph and postulated that \aleph is the smallest uncountable cardinal.

11 The Alefs

Theorem 104. $\forall \alpha \exists \kappa \in \text{Card } \kappa > \alpha$. Hence Card is a proper class of ordinals.

Proof. Let $\alpha \geq \omega$. Then $\kappa = \text{card}(\mathcal{P}(\alpha)) > \text{card}(\alpha)$. And $\kappa > \alpha$ since otherwise $\text{card}(\mathcal{P}(\alpha)) \leq \alpha$ and $\text{card}(\text{card}(\mathcal{P}(\alpha))) \leq \text{card}(\alpha)$. \square

Definition 105. For any ordinal δ let δ^+ be the smallest cardinal $> \delta$.

Theorem 106. Let $X \subseteq \text{Cd}$ be a set. Then $\bigcup X \in \text{Cd}$.

Proof. Set $\kappa = \bigcup X$. κ is an ordinal. Assume that $\text{card}(\kappa) < \kappa$. Take $\lambda \in X$ such that $\text{card}(\kappa) < \lambda$. Then $\lambda \leq \kappa$ and $\text{card}(\lambda) \leq \text{card}(\kappa) < \lambda$. But $\text{card}(\lambda) = \lambda$ because λ is a cardinal. \square

This allows the following

Definition 107. Define the alef sequence

$$(\aleph_\alpha | \alpha \in \text{Ord})$$

recursively by

$$\begin{aligned} \aleph_0 &= \omega \\ \aleph_{\alpha+1} &= \aleph_\alpha^+ \\ \aleph_\lambda &= \bigcup_{\alpha < \lambda} \aleph_\alpha \text{ for limit ordinals } \lambda \end{aligned}$$

Obviously

$$\text{Card} = \{\aleph_\alpha | \alpha \in \text{Ord}\}$$

is the class of all cardinals.

Definition 108. An infinite cardinal of the form $\aleph_{\alpha+1}$ is a successor cardinal. An infinite cardinal of the form \aleph_λ with λ a limit ordinal is a limit cardinal.

Exercise 28. There are cardinals κ such that $\kappa = \aleph_\kappa$.

12 Cardinal Arithmetic

For disjoint finite sets a and b natural addition and multiplication satisfies

$$\text{card}(a \cup b) = \text{card}(a) + \text{card}(b) \text{ and } \text{card}(a \times b) = \text{card}(a) \cdot \text{card}(b).$$

This motivates the following extension of natural arithmetic to all cardinals.

Definition 109. Let κ, λ finite or infinite cardinals. Then let

- a) $\kappa + \lambda = \text{card}(a \cup b)$, where a, b are disjoint sets with $\kappa = \text{card}(a)$ and $\lambda = \text{card}(b)$; $\kappa + \lambda$ is the (cardinal) sum of κ and λ .
- b) $\kappa \cdot \lambda = \text{card}(\kappa \times \lambda)$; $\kappa \cdot \lambda$ is the (cardinal) product of κ and λ .
- c) $\kappa^\lambda = \text{card}(\lambda^\kappa)$; κ^λ is the (cardinal) power of κ and λ .

Note that we are using the same notations as for ordinal arithmetic. It will usually be clear from the context whether ordinal or cardinal operations are intended.

The ‘‘arithmetic’’ properties of certain set operations yield usual arithmetic laws for cardinal arithmetic.

Lemma 110.

- a) Cardinal addition is associative and commutative with neutral element 0.

b) Cardinal multiplication is associative and commutative with neutral element 1.

c) $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$.

d) $\kappa^0 = 1$, $0^\kappa = 0$ for $\kappa \neq 0$, $\kappa^1 = \kappa$, $1^\kappa = 1$, $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$, $\kappa^{\lambda \cdot \mu} = (\kappa^\lambda)^\mu$.

Proof. c) Let a, b be disjoint sets with $\lambda = \text{card}(a)$ and $\mu = \text{card}(b)$. Then

$$\begin{aligned} \kappa \cdot (\lambda + \mu) &= \text{card}(\kappa \times (a \cup b)) \\ &= \text{card}((\kappa \times a) \cup (\kappa \times b)) \\ &= \text{card}((\kappa \times a)) + \text{card}((\kappa \times b)) \\ &= \kappa \cdot \lambda + \kappa \cdot \mu, \end{aligned}$$

using that $\kappa \times (a \cup b) = (\kappa \times a) \cup (\kappa \times b)$ and that $\kappa \times a$ and $\kappa \times b$ are disjoint.

d)

$$\kappa^0 = \text{card}({}^0\kappa) = \text{card}(\{\emptyset\}) = \text{card}(1) = 1.$$

In case $\kappa \neq 0$ we have that ${}^\kappa 0 = \{f \mid f: \kappa \rightarrow \emptyset\} = \emptyset$ and thus

$$0^\kappa = \text{card}({}^\kappa 0) = \text{card}(\emptyset) = 0.$$

For $\kappa^1 = \kappa$ consider the map $\kappa \leftrightarrow {}^1\kappa$ given by $\alpha \mapsto \{(0, \alpha)\}$.

For $1^\kappa = 1$ observe that ${}^\kappa 1 = \{\{(\alpha, 0) \mid \alpha < \kappa\}\}$ is a singleton set.

Let a, b be disjoint sets with $\lambda = \text{card}(a)$ and $\mu = \text{card}(b)$. Then

$$\begin{aligned} \kappa^{\lambda+\mu} &= \text{card}({}^{a \cup b}\kappa) \\ &= \text{card}({}^a\kappa \times {}^b\kappa) \\ &= \text{card}({}^a\kappa) \cdot \text{card}({}^b\kappa) \\ &= \kappa^\lambda \cdot \kappa^\mu, \end{aligned}$$

using that ${}^{a \cup b}\kappa \sim {}^a\kappa \times {}^b\kappa$ via the map $f \mapsto (f \upharpoonright a, f \upharpoonright b)$.

Finally,

$$\begin{aligned} \kappa^{\lambda \cdot \mu} &= \text{card}({}^{\lambda \times \mu}\kappa) \\ &= \text{card}({}^\mu({}^\lambda\kappa)) \\ &= \text{card}({}^\lambda\kappa)^\mu \\ &= (\kappa^\lambda)^\mu, \end{aligned}$$

using that ${}^{\lambda \times \mu}\kappa \sim {}^\mu({}^\lambda\kappa)$ via the map

$$f \mapsto (f_\xi \mid \xi < \mu)$$

where $f_\xi: \lambda \rightarrow \kappa$ with $f_\xi(\zeta) = f(\zeta, \xi)$, □

We determine the values of cardinal addition and multiplication for infinite cardinals.

Definition 111. Define the Gödel ordering $<^2$ of $\text{Ord} \times \text{Ord}$ by

$$\begin{aligned} (\alpha, \beta) <^2 (\alpha', \beta') \text{ iff } &\max(\alpha, \beta) < \max(\alpha', \beta'), \\ &\text{or } \max(\alpha, \beta) = \max(\alpha', \beta') \wedge \alpha < \alpha', \\ &\text{or } \max(\alpha, \beta) = \max(\alpha', \beta') \wedge \alpha = \alpha' \wedge \beta < \beta'. \end{aligned}$$

Lemma 112. $<^2$ is a wellordering of $\text{Ord} \times \text{Ord}$. Let $G: (\text{Ord} \times \text{Ord}, <^2) \leftrightarrow (\text{Ord}, <)$ be the Mostowski collapse of $(\text{Ord} \times \text{Ord}, <^2)$. G is the Gödel pairing function. Define inverse functions $G_1: \text{Ord} \rightarrow \text{Ord}$ and $G_2: \text{Ord} \rightarrow \text{Ord}$ such that

$$\forall \alpha \ G(G_1(\alpha), G_2(\alpha)) = \alpha.$$

Lemma 113. $G: \aleph_\alpha \times \aleph_\alpha \leftrightarrow \aleph_\alpha$.

Proof. By induction on α .

Case 1. $\alpha = 0$. By the definition of $<^2$, $\aleph_0 \times \aleph_0$ is an initial segment of $<^2$. Let

$$G[\aleph_0 \times \aleph_0] = \delta \in \text{Ord}.$$

We show that $\delta = \aleph_0$. Since $\aleph_0 \times \aleph_0$ is infinite, $\delta \geq \aleph_0$. Assume that $\delta > \aleph_0$. Take $m, n \in \omega$ such that $G(m, n) = \omega$. Then (m, n) has infinitely many predecessors in $<^2$. But on the other hand

$$\{(k, l) \mid (k, l) <^2 (m, n)\} \subseteq (\max(m, n) + 1) \times (\max(m, n) + 1)$$

is finite. Hence $G[\aleph_0 \times \aleph_0] = \aleph_0$.

Case 2. $\alpha > 0$ and the Lemma holds for $\beta < \alpha$. Let

$$G[\aleph_\alpha \times \aleph_\alpha] = \eta \in \text{Ord}.$$

We show that $\eta = \aleph_\alpha$. Since $\text{card}(\aleph_\alpha \times \aleph_\alpha) \geq \aleph_\alpha$ we have $\eta \geq \aleph_\alpha$. Assume that $\eta > \aleph_\alpha$. Take $(\xi, \zeta) \in \aleph_\alpha \times \aleph_\alpha$ such that $G(\xi, \zeta) = \aleph_\alpha$. Then G witnesses that

$$\{(\xi', \zeta') \mid (\xi', \zeta') <^2 (\xi, \zeta)\} \sim \aleph_\alpha.$$

On the other hand set $\aleph_\beta = \text{card}(\max(\xi, \zeta) + 1) < \aleph_\alpha$. Then, using the inductive hypothesis,

$$\begin{aligned} \text{card}(\{(\xi', \zeta') \mid (\xi', \zeta') <^2 (\xi, \zeta)\}) &\leq \text{card}((\max(\xi, \zeta) + 1) \times (\max(\xi, \zeta) + 1)) \\ &= \text{card}(\aleph_\beta \times \aleph_\beta) \\ &= \aleph_\beta < \aleph_\alpha, \end{aligned}$$

contradiction. Hence $G[\aleph_\alpha \times \aleph_\alpha] = \aleph_\alpha$. □

Theorem 114.

- a) If $\kappa \in \text{Card}$ then $\kappa \cdot \kappa = \kappa$.
- b) If $\kappa \in \text{Card}$ and $\lambda \in \text{Cd}$, $\lambda \neq 0$ then $\kappa \cdot \lambda = \max(\kappa, \lambda)$.
- c) If $\kappa \in \text{Card}$ and $\lambda \in \text{Cd}$ then $\kappa + \lambda = \max(\kappa, \lambda)$.

Proof. a) $\kappa \cdot \kappa = \text{card}(\kappa \times \kappa) = \kappa$, by the properties of the Gödel pairing function.

b) The map $i \mapsto (i, 0)$ injects κ into $\kappa \times \lambda$, and the map $j \mapsto (0, j)$ injects λ into $\kappa \times \lambda$. Hence $\kappa, \lambda \leq \kappa \cdot \lambda$. Thus

$$\max(\kappa, \lambda) \leq \kappa \cdot \lambda \leq \max(\kappa, \lambda) \cdot \max(\kappa, \lambda) \stackrel{(a)}{=} \max(\kappa, \lambda).$$

c) Obviously $\kappa \sim \{0\} \times \kappa$ and $\lambda \sim \{1\} \times \lambda$. The inclusion

$$(\{0\} \times \kappa) \cup (\{1\} \times \lambda) \subseteq \max(\kappa, \lambda) \times \max(\kappa, \lambda)$$

implies

$$\max(\kappa, \lambda) \leq \kappa + \lambda \leq \max(\kappa, \lambda) \cdot \max(\kappa, \lambda) \stackrel{(a)}{=} \max(\kappa, \lambda). \quad \square$$

For infinite cardinal *exponentiation* the situation is very different. Only a few values can be determined explicitly.

Lemma 115. For $\kappa \in \text{Card}$ and $1 \leq n < \omega$ we have $\kappa^n = \kappa$.

Proof. By complete induction. $\kappa^1 = \kappa$ was proved before. And

$$\kappa^{n+1} = (\kappa^n) \cdot \kappa^1 = \kappa \cdot \kappa = \kappa. \quad \square$$

The “next” exponential value 2^{\aleph_0} is however very undetermined. It is possible, in a sense to be made precise later, that 2^{\aleph_0} is any successor cardinal, like e.g. \aleph_{13} .

Cantor’s continuum hypothesis is equivalent to the cardinal arithmetic statement

$$2^{\aleph_0} = \aleph_1.$$

Lemma 116. For $\kappa \in \text{Card}$ and $2 \leq \lambda \leq 2^\kappa$ we have $\lambda^\kappa = 2^\kappa$.

Proof.

$$2^\kappa \leq \lambda^\kappa \leq (2^\kappa)^\kappa = 2^{\kappa \cdot \kappa} = 2^\kappa. \quad \square$$

13 Cofinality

To get some more information on cardinal exponentiation, we need to measure how “fast” a cardinal can be approximated using smaller cardinals.

Definition 117.

- a) A set $x \subseteq \lambda$ is cofinal in the limit ordinal λ if $\forall \alpha < \lambda \exists \xi \in x \alpha < \xi$.
- b) The cofinality of a limit ordinal λ is

$$\text{cof}(\lambda) = \min \{ \text{otp}(x) \mid x \subseteq \lambda \text{ is cofinal in } \lambda \}.$$

- c) A limit ordinal λ is regular if $\text{cof}(\lambda) = \lambda$; otherwise λ is singular.

These notions are due to Felix Hausdorff, who called them “konfinal” and “Konfinalität”. Please observe the “konfinal” in German.

Lemma 118.

- a) $\text{cof}(\lambda) = \min \{ \text{card}(x) \mid x \subseteq \lambda \text{ is cofinal in } \lambda \}$
- b) $\text{cof}(\aleph_0) = \aleph_0$, i.e., \aleph_0 is regular
- c) $\text{cof}(\lambda) \leq \text{card}(\lambda) \leq \lambda$
- d) $\text{cof}(\lambda) \in \text{Card}$
- e) $\text{cof}(\lambda)$ is regular, i.e., $\text{cof}(\text{cof}(\lambda)) = \text{cof}(\lambda)$
- f) If γ is a limit ordinal then $\text{cof}(\aleph_\gamma) = \text{cof}(\gamma)$
- g) $\text{cof}(\aleph_\omega) = \aleph_0$, i.e., \aleph_ω is a singular cardinal

Proof. a) \geq holds since $\text{otp}(x) \geq \text{card}(x)$. Conversely let x have minimal cardinality such that x is cofinal in λ and let $f: \text{card}(x) \leftrightarrow x$. Define a weakly increasing map $g: \text{card}(x) \rightarrow \lambda$ by

$$g(i) = \bigcup_{j < i} f(j).$$

g is welldefined by the minimality of x . $y = g[\text{card}(x)]$ is cofinal in λ . y is order-isomorphic to

$$\{ i < \text{card}(x) \mid \forall j < i \ g(j) < g(i) \} \subseteq \text{card}(x).$$

Hence

$$\text{otp}(y) = \text{otp}(\{ i < \text{card}(x) \mid \forall j < i \ g(j) < g(i) \}) \leq \text{card}(x).$$

Thus

$$\text{cof}(\lambda) \leq \text{otp}(y) = \text{card}(x) = \min \{ \text{card}(x) \mid x \subseteq \lambda \text{ is cofinal in } \lambda \}.$$

b) – d) follow from a).

e) Let $x \subseteq \lambda$ be cofinal in λ with $\text{otp}(x) = \text{cof}(\lambda)$ and order-isomorphism $f: \text{cof}(\lambda) \leftrightarrow x$. Let $y \subseteq \text{cof}(\lambda)$ be cofinal with $\text{otp}(y) = \text{cof}(\text{cof}(\lambda))$ and order-isomorphism $g: \text{cof}(\text{cof}(\lambda)) \leftrightarrow y$. Then $z = f \circ g[\text{cof}(\text{cof}(\lambda))]$ is cofinal in λ with $\text{otp}(z) = \text{cof}(\text{cof}(\lambda))$. Hence

$$\text{cof}(\lambda) \leq \text{otp}(z) = \text{cof}(\text{cof}(\lambda)).$$

The converse inequality follows from c).

f) (\leq) Let x be cofinal in γ with $\text{otp}(x) = \text{cof}(\gamma)$. Then $\{\aleph_i \mid i \in x\}$ is cofinal in \aleph_γ with

$$\text{otp}(\{\aleph_i \mid i \in x\}) = \text{otp}(x) = \text{cof}(\gamma).$$

Hence $\text{cof}(\aleph_\gamma) \leq \text{cof}(\gamma)$.

(\supseteq) Now let y be cofinal in \aleph_γ with $\text{otp}(y) = \text{cof}(\aleph_\gamma)$. Define $x = \{i < \gamma \mid \exists \delta \in y \aleph_i \leq \delta < \aleph_{i+1}\}$. Then x is cofinal in γ with $\text{card}(x) \leq \text{card}(y) = \text{cof}(\aleph_\gamma)$. Hence $\text{cof}(\gamma) \leq \text{cof}(\aleph_\gamma)$. \square

Theorem 119. *Every successor cardinal $\aleph_{\alpha+1}$ is regular.*

Proof. Assume that $\aleph_{\alpha+1}$ is singular. Let x have minimal cardinality such that x is cofinal in $\aleph_{\alpha+1}$. Then $\text{card}(x) \leq \aleph_\alpha$. Let $f: \aleph_\alpha \rightarrow x$ be surjective. Using the axiom of choice take a sequence $(g_i \mid 0 < i < \aleph_{\alpha+1})$ of surjective functions $g_i: \aleph_\alpha \rightarrow i$. Define function $h: \aleph_\alpha \times \aleph_\alpha \rightarrow \aleph_{\alpha+1}$ by

$$h(\xi, \zeta) = g_{f(\xi)}(\zeta).$$

(1) $h: \aleph_\alpha \times \aleph_\alpha \rightarrow \aleph_{\alpha+1}$ is surjective.

Proof. Let $\nu \in \aleph_{\alpha+1}$. Take $\xi < \aleph_\alpha$ such that $f(\xi) > \nu$. $g_{f(\xi)}: \aleph_\alpha \rightarrow f(\xi)$ is surjective. Take $\zeta < \aleph_\alpha$ such that $g_{f(\xi)}(\zeta) = \nu$. Thus $\nu = h(\xi, \zeta) \in \text{ran}(h)$. *qed*(1)

This implies

$$\aleph_{\alpha+1} = \text{card}(\aleph_{\alpha+1}) \leq \text{card}(\aleph_\alpha \times \aleph_\alpha) = \aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha.$$

Contradiction. \square

So $\aleph_0, \aleph_1, \aleph_2, \dots, \aleph_n, \dots$ are all regular.

Question 120. (Hausdorff) *Are there regular limit cardinals $> \aleph_0$?*

Definition 121. *For $(\kappa_i \mid i < \delta)$ a sequence of finite or infinite cardinals define the sum*

$$\sum_{i < \delta} \kappa_i = \text{card}\left(\bigcup_{i < \delta} \kappa_i \times \{i\}\right)$$

and the product

$$\prod_{i < \delta} \kappa_i = \text{card}\left(\prod_{i < \delta} \kappa_i\right)$$

where

$$\prod_{i < \delta} A_i = \{f \mid f: \delta \rightarrow V \wedge \forall i < \delta f(i) \in A_i\}.$$

Theorem 122. (König) *If $(\kappa_i \mid i < \delta)$ and $(\lambda_i \mid i < \delta)$ are sequences of cardinals such that $\forall i < \delta \kappa_i < \lambda_i$ then*

$$\sum_{i < \delta} \kappa_i < \prod_{i < \delta} \lambda_i$$

Proof. Assume for a contradiction that $\sum_{i < \delta} \kappa_i \geq \prod_{i < \delta} \lambda_i$ and that $G: \bigcup_{i < \delta} \kappa_i \times \{i\} \leftrightarrow \prod_{i < \delta} \lambda_i$ were a surjection. For $i < \delta$

$$\text{card}(\{G(\nu, i)(i) \mid \nu < \kappa_i\}) \leq \kappa_i < \lambda_i,$$

and one can choose $\nu_i \in \lambda_i \setminus \{G(\nu, i)(i) \mid \nu < \kappa_i\}$. Define $f \in \prod_{i < \delta} \lambda_i$ by

$$f(i) = \nu_i.$$

Since G is surjective, take $(\nu_0, i_0) \in \text{dom}(G)$ such that $G(\nu_0, i_0) = f$. Then

$$G(\nu_0, i_0)(i_0) = f(i_0) = \nu_{i_0} \neq G(\nu_0, i_0)(i_0)$$

for all $\nu < \kappa_{i_0}$. Contradiction. \square

Theorem 123. *If κ, λ are cardinals such that $\kappa \geq 2$ and $\lambda \geq \aleph_0$ then*

$$\text{cof}(\kappa^\lambda) > \lambda$$

Hence

$$\text{cof}(2^{\aleph_0}) \geq \aleph_1$$

and in particular

$$2^{\aleph_0} \neq \aleph_\omega.$$

Proof. Assume that $\text{cof}(\kappa^\lambda) \leq \lambda$. Then there is a function $f: \lambda \rightarrow \kappa^\lambda$ such that $\text{ran}(f)$ is cofinal in κ^λ . Then $\bigcup_{i < \lambda} f(i) = \kappa^\lambda$ and so

$$\kappa^\lambda = \text{card}\left(\bigcup_{i < \lambda} f(i)\right) \leq \text{card}\left(\bigcup_{i < \lambda} f(i) \times \{i\}\right) = \text{card}\left(\bigcup_{i < \lambda} \text{card}(f(i)) \times \{i\}\right) = \sum_{i < \lambda} \text{card}(f(i)).$$

But by König's Theorem,

$$\kappa^\lambda = \kappa^{\lambda \cdot \lambda} = (\kappa^\lambda)^\lambda = \prod_{i < \lambda} \kappa^\lambda > \sum_{i < \lambda} \text{card}(f(i)). \quad \square$$

Theorem 124. (The Hausdorff recursion formula)

$$\aleph_{\alpha+1}^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1}.$$

Proof. Distinguish two cases:

Case 1: $2^{\aleph_\beta} \leq \aleph_{\alpha+1}$. Then

$$\aleph_{\alpha+1}^{\aleph_\beta} = 2^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta} = \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1}.$$

Case 2: $2^{\aleph_\beta} < \aleph_{\alpha+1}$: Then $\aleph_\beta < \aleph_{\alpha+1}$. Using the regularity of $\aleph_{\alpha+1}$

$$\begin{aligned} & \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1} \\ & \leq \aleph_{\alpha+1}^{\aleph_\beta} \cdot \aleph_{\alpha+1} = \\ & = \aleph_{\alpha+1}^{\aleph_\beta} = \text{card}(\{f \mid f: \aleph_\beta \rightarrow \aleph_{\alpha+1}\}) \\ & = \text{card}\left(\bigcup_{\nu < \aleph_{\alpha+1}} \{f \mid f: \aleph_\beta \rightarrow \nu\}\right) \\ & \leq \sum_{\nu < \aleph_{\alpha+1}} \text{card}(\{f \mid f: \aleph_\beta \rightarrow \nu\}) \\ & = \sum_{\nu < \aleph_{\alpha+1}} \text{card}(\{f \mid f: \aleph_\beta \rightarrow \text{card}(\nu)\}) \\ & = \sum_{\nu < \aleph_{\alpha+1}} \text{card}(\nu)^{\aleph_\beta} \\ & \leq \sum_{\nu < \aleph_{\alpha+1}} \aleph_\alpha^{\aleph_\beta} \\ & = \aleph_\alpha^{\aleph_\beta} \cdot \aleph_{\alpha+1} \end{aligned}$$

□

14 Cardinal exponentiation and the Generalized Continuum Hypothesis

The function $\kappa \mapsto 2^\kappa$ is called the *continuum function*, due to the relations between 2^{\aleph_0} and the usual continuum of real numbers. The *beth numbers* are defined in analogy with the aleph function, using the 2^κ -operation instead of the cardinal successor function.

Definition 125. Define the sequence

$$(\beth_\alpha \mid \alpha \in \text{Ord})$$

of beth numbers recursively by

$$\begin{aligned} \beth_0 &= \aleph_0 \\ \beth_{\alpha+1} &= 2^{\beth_\alpha} \\ \beth_\lambda &= \bigcup_{\alpha < \lambda} \beth_\alpha \text{ for limit ordinals } \lambda \end{aligned}$$

Like every continuous ordinal function, there are fixed points $\beth_\alpha = \alpha$ of this sequence.

Definition 126.

a) An inaccessible cardinal κ is a regular fixed point of the \aleph_α -function:

$$\kappa = \aleph_\kappa \text{ and } \text{cof}(\kappa) = \kappa.$$

b) A strongly inaccessible cardinal κ is a regular fixed point of the \beth_α -sequence:

$$\kappa = \beth_\kappa \text{ and } \text{cof}(\kappa) = \kappa.$$

The existence of inaccessible and strongly inaccessible cardinals can not be shown in ZFC, provided the theory ZFC is consistent.

Definition 127. Define the gimel function $\beth: \text{Card} \rightarrow \text{Card}$ by $\beth(\kappa) = \kappa^{\text{cof}(\kappa)}$.

By König's theorem, $\beth(\kappa) > \kappa$. Note that \aleph (Alef), \beth (Beth) and \beth (Gimel) are the first three letters of the Hebrew alphabet. The gimel function determines all values of the continuum function.

Definition 128. For $\kappa \in \text{Cd}$ and $\lambda \in \text{Card}$ let

$$\kappa^{<\lambda} = \bigcup_{\nu < \lambda} \kappa^{\text{card}(\nu)}.$$

Theorem 129.

a) If κ is regular then $2^\kappa = \beth(\kappa)$.

b) If κ is a singular cardinal and the continuum function is eventually constant below κ , i.e.,

$$\exists \bar{\kappa} < \kappa \forall \lambda (\bar{\kappa} \leq \lambda < \kappa \rightarrow 2^{\bar{\kappa}} = 2^\lambda),$$

then $2^\kappa = 2^{<\kappa}$.

c) If κ is a singular cardinal and the continuum function is not eventually constant below κ then $2^\kappa = \beth(2^{<\kappa})$.

Proof. a) If κ is regular then

$$2^\kappa = \kappa^\kappa = \kappa^{\text{cof}(\kappa)} = \beth(\kappa).$$

Now let κ be singular and let the sequence $(\kappa_i | i < \text{cof}(\kappa))$ be strictly increasing and cofinal in κ . For $i < \text{cof}(\kappa)$ choose (AC) an injection $f_i: \mathcal{P}(\kappa_i) \rightarrow 2^{<\kappa}$. Define

$$G: \mathcal{P}(\kappa) \rightarrow {}^{\text{cof}(\kappa)}(2^{<\kappa})$$

by

$$x \mapsto (f_i(x \cap \kappa_i) | i < \text{cof}(\kappa)).$$

We argue that G is injective: let $x, y \in \mathcal{P}(\kappa)$, $x \neq y$. Then take $i < \text{cof}(\kappa)$ such that $x \cap \kappa_i \neq y \cap \kappa_i$. Since f_i is injective: $f_i(x \cap \kappa_i) \neq f_i(y \cap \kappa_i)$. Then $G(x) \neq G(y)$ because

$$G(x)(i) = f_i(x \cap \kappa_i) \neq f_i(y \cap \kappa_i) = G(y)(i).$$

By the injectivity of G

$$2^\kappa \leq (2^{<\kappa})^{\text{cof}(\kappa)}.$$

b) Let $2^{<\kappa} = 2^{\bar{\kappa}}$ be the eventually constant value of the continuum function below κ . Then

$$2^{<\kappa} \leq 2^\kappa \leq (2^{<\kappa})^{\text{cof}(\kappa)} = (2^{\bar{\kappa}})^{\text{cof}(\kappa)} = 2^{\bar{\kappa} \cdot \text{cof}(\kappa)} = 2^{\max(\bar{\kappa}, \text{cof}(\kappa))} = 2^{\bar{\kappa}} = 2^{<\kappa}.$$

c) In this case we show that $\text{cof}(\kappa) = \text{cof}(2^{<\kappa})$. The function

$$i \mapsto 2^{\kappa_i}$$

is not eventually constant and thus maps $\text{cof}(\kappa)$ cofinally into $2^{<\kappa}$. Hence $\text{cof}(2^{<\kappa}) \leq \text{cof}(\kappa)$. Assume that $\text{cof}(2^{<\kappa}) < \text{cof}(\kappa)$. Let $z \subseteq 2^{<\kappa}$ be cofinal such that $\text{card}(z) < \text{cof}(\kappa)$. Then

$$\bar{z} = \{i \mid \exists \delta \in z \ 2^{\kappa i} \leq \delta < 2^{\kappa(i+1)}\} \subseteq \text{cof}(\kappa)$$

is cofinal in $\text{cof}(\kappa)$ and $\text{card}(\bar{z}) \leq \text{card}(z) < \text{cof}(\kappa)$, contradiction.

So we obtain

$$\beth(2^{<\kappa}) = (2^{<\kappa})^{\text{cof}(2^{<\kappa})} = (2^{<\kappa})^{\text{cof}(\kappa)} \leq (2^\kappa)^{\text{cof}(\kappa)} \leq 2^\kappa \leq (2^{<\kappa})^{\text{cof}(\kappa)} = (2^{<\kappa})^{\text{cof}(2^{<\kappa})} = \beth(2^{<\kappa}).$$

□

The following theorem shows that κ^λ is uniquely determined by the gimel function.

Theorem 130. *Let $\lambda \in \text{Card}$. Then κ^λ is determined by the previous theorem and by recursion on κ :*

- a) $0^\lambda = 0, 1^\lambda = 1$.
- b) For $2 \leq \kappa \leq \lambda$ we have $\kappa^\lambda = 2^\lambda$.
- c) If $\kappa > \lambda$ and $\xi < \kappa$ such that $\xi^\lambda \geq \kappa$ then $\kappa^\lambda = \xi^\lambda$.
- d) If $\kappa > \lambda, \forall \xi < \kappa \ \xi^\lambda < \kappa$, and $\text{cof}(\kappa) > \lambda$ then $\kappa^\lambda = \kappa$.
- e) If $\kappa > \lambda, \forall \xi < \kappa \ \xi^\lambda < \kappa$, and $\text{cof}(\kappa) \leq \lambda$ then $\kappa^\lambda = \beth(\kappa)$.

Proof. a) and b) follow immediately from earlier results.

c) $\xi^\lambda \leq \kappa^\lambda \leq (\xi^\lambda)^\lambda = \xi^\lambda$.

d) $\text{cof}(\kappa) > \lambda$ implies that every function from λ into κ is bounded by some ordinal $\nu < \kappa$. Hence

$$\begin{aligned} \kappa \leq \kappa^\lambda &= \text{card}\{f \mid f: \lambda \rightarrow \kappa\} \\ &= \text{card}\left(\bigcup_{\nu < \kappa} \{f \mid f: \lambda \rightarrow \nu\}\right) \\ &\leq \sum_{\nu < \kappa} \text{card}(\{f \mid f: \lambda \rightarrow \nu\}) \\ &= \sum_{\nu < \kappa} \text{card}(\{f \mid f: \lambda \rightarrow \text{card}(\nu)\}) \\ &= \sum_{\nu < \kappa} \text{card}(\nu)^\lambda \\ &\leq \sum_{\nu < \kappa} \kappa \\ &= \kappa. \end{aligned}$$

e) Let $(\xi_i \mid i < \text{cof}(\kappa))$ be a strictly increasing sequence which is cofinal in κ . Define a function

$$G: {}^\lambda \kappa \rightarrow \prod_{i < \text{cof}(\kappa)} {}^\lambda \xi_i$$

by

$$f \mapsto (f_i \mid i < \text{cof}(\kappa))$$

where

$$f_i(\alpha) = \begin{cases} f(\alpha), & \text{if } f(\alpha) < \xi_i \\ 0, & \text{else} \end{cases}$$

Then G is injective: Let $f, g \in {}^\lambda \kappa, f \neq g$. Take α such that $f(\alpha) \neq g(\alpha)$ and take i such that $f(\alpha), g(\alpha) < \xi_i$. Then $f_i(\alpha) = f(\alpha) \neq g(\alpha) = g_i(\alpha), f_i \neq g_i$, and hence $G(f) \neq G(g)$.

Using G we get

$$\begin{aligned} \beth(\kappa) = \kappa^{\text{cof}(\kappa)} &\leq \kappa^\lambda \leq \prod_{i < \text{cof}(\kappa)} \text{card}(\xi_i)^\lambda \\ &\leq \prod_{i < \text{cof}(\kappa)} \kappa \\ &= \kappa^{\text{cof}(\kappa)} = \beth(\kappa) \end{aligned}$$

□

Definition 131. (Hausdorff) *The generalized continuum hypothesis (GCH) is the statement*

$$\forall \kappa \in \text{Card } 2^\kappa = \kappa^+.$$

This is the “minimal” hypothesis in view of Cantor’s $2^\kappa \geq \kappa^+$. The GCH generalizes Cantor’s continuum hypothesis CH and also the hypothesis $2^{\aleph_1} = \aleph_2$ also expressed by Cantor. Since CH is independent of the axioms of set theory, GCH is independent as well. Indeed the continuum function is hardly determined by the axioms of ZFC and one can for example have

$$2^{\aleph_0} = \aleph_{73}, 2^{\aleph_1} = \aleph_{2015}, \dots$$

Obviously

Lemma 132. GCH implies that $\forall \kappa \in \text{Card } \beth(\kappa) = \kappa^+$.

Thus GCH also determines all values of the κ^λ function. Axiomatic set theory proves that one can assume GCH without the danger of adding inconsistencies to the system ZFC: a model of the ZFC axioms can be modified into a model of ZFC + GCH. The consequences of GCH for cardinal exponentiation can be readily described.

Theorem 133. Assume GCH. Then for $\kappa, \lambda \in \text{Card}$ κ^λ is determined as follows:

- a) For $\lambda < \text{cof}(\kappa)$: $\kappa^\lambda = \kappa$.
- b) For $\text{cof}(\kappa) \leq \lambda \leq \kappa$: $\kappa^\lambda = \kappa^+$.
- c) For $\lambda > \kappa$: $\kappa^\lambda = \lambda^+$.

Proof. a)

$$\begin{aligned} \kappa &\leq \kappa^\lambda \\ &= \text{card}\{f \mid f: \lambda \rightarrow \kappa\} \\ &= \text{card} \bigcup \{f \mid f: \lambda \rightarrow \nu\} \\ &\leq \sum_{\nu < \kappa} \text{card}(\nu)^\lambda \\ &\leq \sum_{\nu < \kappa} \kappa \\ &= \kappa. \end{aligned}$$

b) By GCH and König’s theorem, $\kappa^{\text{cof}(\kappa)} = \kappa^+$. Thus

$$\kappa^+ = \kappa^{\text{cof}(\kappa)} \leq \kappa^\lambda \leq \kappa^\kappa = 2^\kappa = \kappa^+.$$

c)

$$\lambda^+ = 2^\lambda \leq \kappa^\lambda \leq \lambda^\lambda = 2^\lambda = \lambda^+. \quad \square$$

Question 134. Is every (infinite) cardinal product $\prod_{i < \alpha} \kappa_i$ also determined by GCH?

15 Closed unbounded and stationary sets

The continuum function $\kappa \mapsto 2^\kappa$ satisfies the laws

- a) $\kappa \leq \lambda \rightarrow 2^\kappa \leq 2^\lambda$
- b) $\text{cof}(2^\kappa) > \kappa$

Axiomatic set theory shows that for *regular* cardinals κ these are the only laws deducible from ZFC: for (adequate) functions $F: \text{Card} \rightarrow \text{Card}$ satisfying a) and b) for regular cardinals there is a model of set theory in which

$$\kappa \text{ regular} \rightarrow 2^\kappa = F(\kappa).$$

So there remains the consideration of 2^κ for *singular* cardinals κ . Indeed singular cardinal exponentiation satisfies some interesting further laws and is an area of present research. To prove a few of these laws we have to extend the apparatus of uncountable combinatorics.

Definition 135. Let $\kappa \in \text{Card}$ and $C \subseteq \kappa$. C is unbounded in κ if

$$\forall \alpha < \kappa \exists \beta \in C \alpha < \beta.$$

C is closed in κ if

$$\forall \lambda < \kappa (\lambda \text{ is a limit ordinal} \wedge C \cap \lambda \text{ is unbounded in } \lambda \rightarrow \lambda \in C).$$

Thus C contains its limit points $< \kappa$.

C is closed unbounded, or cub in κ , if C is unbounded in κ and closed in κ .

Exercise 29. Define a topology on κ such that the closed sets of the topology are exactly the closed sets in the sense of the previous definition.

Lemma 136. Let $\kappa \in \text{Card}$, $\text{cof}(\kappa) \geq \omega_1$ and C, D be closed unbounded in κ . Then $C \cap D$ is cub in κ .

Proof. $C \cap D$ is closed in κ : Let $\lambda < \kappa$ be a limit ordinal and a limit point of $C \cap D$. Then λ is a limit point of C and $\lambda \in C$. Similarly $\lambda \in D$ and together $\lambda \in C \cap D$.

$C \cap D$ is unbounded in κ : Let $\alpha < \kappa$. Define a sequence $(\beta_n | n < \omega)$ by recursion:

$$\beta_n = \begin{cases} \text{the least element of } C \text{ which is larger than } \alpha, \beta_0, \dots, \beta_{n-1} \text{ in case } n \text{ is even} \\ \text{the least element of } D \text{ which is larger than } \beta_0, \dots, \beta_{n-1} \text{ in case } n \text{ is odd} \end{cases}$$

Let $\beta = \bigcup_{n < \omega} \beta_n$. β is a limit ordinal $> \alpha$. $\beta < \kappa$ since $\text{cof}(\kappa) \geq \omega_1$. By construction, β is a limit point of C and of D . Hence $\beta \in C \cap D$. \square

Exercise 30. Let $\kappa \in \text{Card}$, $\text{cof}(\kappa) \geq \omega_1$. Let $(C_i | i < \gamma)$ be a sequence of sets C_i which are closed unbounded in κ and let $\gamma < \text{cof}(\kappa)$. Then $\bigcap_{i < \gamma} C_i$ is cub in κ .

Definition 137. Let $\kappa \in \text{Card}$, $\text{cof}(\kappa) \geq \omega_1$. The closed unbounded filter on κ is

$$\mathcal{C}_\kappa = \{X \subseteq \kappa \mid \text{there is a set } C \subseteq X \text{ which is closed unbounded in } \kappa\}.$$

Lemma 138. \mathcal{C}_κ is a $\text{cof}(\kappa)$ -complete filter on κ , i.e.

- a) $\emptyset \neq \mathcal{C}_\kappa \subseteq \mathcal{P}(\kappa)$
- b) $\emptyset \notin \mathcal{C}_\kappa$
- c) $X \in \mathcal{C}_\kappa \wedge X \subseteq Y \subseteq \kappa \rightarrow Y \in \mathcal{C}_\kappa$
- d) $X \in \mathcal{C}_\kappa \wedge Y \in \mathcal{C}_\kappa \rightarrow X \cap Y \in \mathcal{C}_\kappa$
- e) $\gamma < \text{cof}(\kappa) \wedge \{X_i \mid i < \gamma\} \subseteq \mathcal{C}_\kappa \rightarrow \bigcap_{i < \gamma} X_i \in \mathcal{C}_\kappa$

Proof. c) and d) follow from Lemma 136 and Exercise 30. \square

A filter captures a notion of “large set”. Even intersections of large sets are large, so that certain constructions can be continued on large sets. Largeness also yields notions of “small” and of “not small”, called “non-stationary” and “stationary”.

Definition 139. Let $\kappa \in \text{Card}$, $\text{cof}(\kappa) \geq \omega_1$.

- a) $X \subseteq \kappa$ is non-stationary in κ if $\kappa \setminus X \in \mathcal{C}_\kappa$. We call

$$\text{NS}_\kappa = \{X \mid \kappa \setminus X \in \mathcal{C}_\kappa\}$$

the non-stationary ideal on κ .

- b) $X \subseteq \kappa$ is stationary in κ if $X \notin \text{NS}_\kappa$.

Lemma 140. $X \subseteq \kappa$ is stationary in κ iff $X \cap C \neq \emptyset$ for every cub $C \subseteq \kappa$.

Proof. X is stationary iff $X \notin \text{NS}_\kappa$ iff $\kappa \setminus X \notin \mathcal{C}_\kappa$ iff there is no $C \subseteq \kappa$ cub such that $C \subseteq \kappa \setminus X$ iff for every cub $C \subseteq \kappa$ $C \not\subseteq \kappa \setminus X$ iff for every cub $C \subseteq \kappa$ $X \cap C \neq \emptyset$. \square

Lemma 141. Every set in \mathcal{C}_κ is stationary.

Lemma 142. NS_κ is a $\text{cof}(\kappa)$ -complete ideal on κ , i.e.

- a) $\emptyset \neq \text{NS}_\kappa \subseteq \mathcal{P}(\kappa)$
- b) $\kappa \notin \text{NS}_\kappa$
- c) $X \in \text{NS}_\kappa \wedge Y \subseteq X \subseteq \kappa \rightarrow Y \in \text{NS}_\kappa$
- d) $X \in \text{NS}_\kappa \wedge Y \in \text{NS}_\kappa \rightarrow X \cap Y \in \text{NS}_\kappa$
- e) $\gamma < \text{cof}(\kappa) \wedge \{X_i \mid i < \gamma\} \subseteq \text{NS}_\kappa \rightarrow \bigcup_{i < \gamma} X_i \in \text{NS}_\kappa$

Proof. e) Let $\gamma < \text{cof}(\kappa) \wedge \{X_i \mid i < \gamma\} \subseteq \text{NS}_\kappa$. Then $\{\kappa \setminus X_i \mid i < \gamma\} \subseteq \mathcal{C}_\kappa$. By Lemma 136, $\bigcap_{i < \gamma} (\kappa \setminus X_i) \in \mathcal{C}_\kappa$. Hence

$$\bigcup_{i < \gamma} X_i = \kappa \setminus \bigcap_{i < \gamma} (\kappa \setminus X_i) \in \text{NS}_\kappa.$$

\square

For regular uncountable κ these filters and ideals have even better completeness properties.

Definition 143. Let κ be a regular uncountable cardinal. For a sequence $(X_i)_{i < \kappa}$ of subsets of κ define

- a) the diagonal intersection

$$\bigtriangleup_{i < \kappa} X_i = \{\beta < \kappa \mid \forall i < \beta \beta \in X_i\},$$

- b) the diagonal union

$$\bigtriangledown_{i < \kappa} X_i = \{\beta < \kappa \mid \exists i < \beta \beta \in X_i\}.$$

Lemma 144. Let κ be a regular uncountable cardinal. Then

- a) \mathcal{C}_κ is closed under diagonal intersections, i.e.,

$$\{X_i \mid i < \kappa\} \subseteq \mathcal{C}_\kappa \rightarrow \bigtriangleup_{i < \kappa} X_i \in \mathcal{C}_\kappa,$$

- b) NS_κ is closed under diagonal unions, i.e.,

$$\{X_i \mid i < \kappa\} \subseteq \text{NS}_\kappa \rightarrow \bigtriangledown_{i < \kappa} X_i \in \text{NS}_\kappa.$$

Proof. a) Let $\{X_i \mid i < \kappa\} \subseteq \mathcal{C}_\kappa$. For $i < \kappa$ choose $C_i \in \mathcal{C}_\kappa$ such that $C_i \subseteq X_i$. Then

$$\bigtriangleup_{i < \kappa} C_i \subseteq \bigtriangleup_{i < \kappa} X_i$$

and it suffices to show that $\bigtriangleup_{i < \kappa} C_i$ is cub in κ .

$\bigtriangleup_{i < \kappa} C_i$ is closed in κ : Let $\lambda < \kappa$ be a limit ordinal and a limit point of $\bigtriangleup_{i < \kappa} C_i$. Consider $j < \lambda$. By the definition of the diagonal intersection

$$\left(\bigtriangleup_{i < \kappa} C_i \right) \setminus (j+1) \subseteq C_j.$$

Hence λ is a limit point of C_j and $\lambda \in C_j$ by the closure of C_j . Thus $\forall j < \lambda \lambda \in C_j$ and thus $\lambda \in \bigtriangleup_{i < \kappa} C_i$.

$\Delta_{i < \kappa} C_i$ is unbounded in κ : Let $\alpha < \kappa$. Define a sequence $(\beta_n | n < \omega)$ by recursion: set $\beta_0 = \alpha$ and

$$\beta_{n+1} = \text{the least element of } \left(\bigcap_{i < \beta_n} C_i \right) \setminus (\beta_n + 1).$$

Let $\beta = \bigcup_{n < \omega} \beta_n$. β is a limit ordinal $> \alpha$. $\beta < \kappa$ since $\text{cof}(\kappa) \geq \omega_1$. We show that $\beta \in \Delta_{i < \kappa} C_i$. Consider $j < \beta$. Take $n < \omega$ such that $j < \beta_n$. Then

$$\{\beta_k | n < k < \omega\} \subseteq C_j$$

and β is a limit point of C_j . $\beta \in C_j$ by the closure of C_j . Hence $\forall j < \beta \beta \in C_j$. \square

Sets in an ideal behave similar to sets of (Lebesgue-)measure 0. Then sets not in the ideal have “positive measure”. So stationary sets are *positive* with respect to the non-stationary ideal. Closure under diagonal intersections corresponds to a surprising canonization property of certain functions.

Definition 145. A function $f: Z \rightarrow \text{Ord}$ where $Z \subseteq \text{Ord}$ is regressive if

$$\alpha \in Z \setminus \{0\} \rightarrow f(\alpha) < \alpha.$$

Exercise 31. If $\gamma \geq \omega$ then there is no regressive *injective* function $f: \gamma \rightarrow \gamma$.

Theorem 146. (Fodor’s Lemma) Let κ be a regular uncountable cardinal and let $f: S \rightarrow \kappa$ be regressive, where S is stationary in κ . Then there is a stationary $T \subseteq S$ such that $f \upharpoonright T$ is constant.

Proof. Assume that for every $i < \kappa$ $f^{-1}[\{i\}]$ is not stationary. So for every $i < \kappa$ choose a C_i cub such that $\forall j \in T \cap C_i f(j) \neq i$. The set

$$C = \Delta_{i < \kappa} C_i$$

is cub in κ , and so there is $\alpha \in C \cap T$, $\alpha > 0$. But then for all $i < \alpha$ $\alpha \in T \cap C_i$ and $f(\alpha) \neq i$. But then $f(\alpha) \geq \alpha$, contradicting the regressivity of f . \square

We now give examples of closed unbounded and stationary sets.

Lemma 147. Let κ be an uncountable regular cardinal and let C be cub in κ . The derivation C' of C is defined as

$$C' = \{\alpha \in C \mid \alpha \text{ is a limit point of } C\}.$$

Then C' is cub in κ .

One can form iterated derivations $C^{(i)}$ for $i < \kappa$ by

$$\begin{aligned} C^{(0)} &= C \\ C^{(i+1)} &= (C^{(i)})' \\ C^{(\lambda)} &= \bigcap_{i < \lambda} C^{(i)} \text{ for limit ordinals } \lambda < \kappa \end{aligned}$$

Every $C^{(i)}$ is cub in κ .

The lemma implies immediately:

Lemma 148. Let Lim be the class of limit ordinals. Let κ be an uncountable regular cardinal. Then $\text{Lim} \cap \kappa$ is cub in κ .

Topologically these derivation correspond to the process of omitting isolated points. Such iterated derivations were first studied by Cantor.

Example 149. For κ an uncountable regular cardinal let $(\kappa)^{<\omega}$ be the set of all finite sequences from κ , i.e., $(\kappa)^{<\omega} = \{u \mid \exists n < \omega \ u: n \rightarrow \kappa\}$. For $h: (\kappa)^{<\omega} \rightarrow \kappa$ let

$$C_h = \{\beta < \kappa \mid h[(\beta)^{<\omega}] \subseteq \beta\}$$

be the set of ordinals $< \kappa$ which are closed under h . Then C_h is cub in κ . Given $\alpha < \kappa$, a closed ordinal $\beta > \alpha$ can be found as $\beta = \bigcup_{n < \omega} \beta_n$ where $\beta_0 = \alpha$ and

$$\beta_{n+1} = \left(\bigcup h[(\beta_n)^{<\omega}] \right) + 1 < \kappa.$$

Conversely, if C is cub in κ one can define $g: \kappa \rightarrow \kappa$ by

$$g(\alpha) = \text{the smallest element of } C \text{ which is } > \alpha.$$

If we define C_g as above then

$$C_g = \{0\} \cup C'.$$

Lemma 150. Let $\mu < \kappa$ be uncountable regular cardinals. Then

$$E_\mu^\kappa = \{\alpha < \kappa \mid \text{cof}(\alpha) = \mu\}$$

is stationary in κ .

Proof. Let C be cub in κ . Define a strictly increasing $\mu + 1$ -sequence $(\alpha_i)_{i \leq \mu}$ of elements of C by

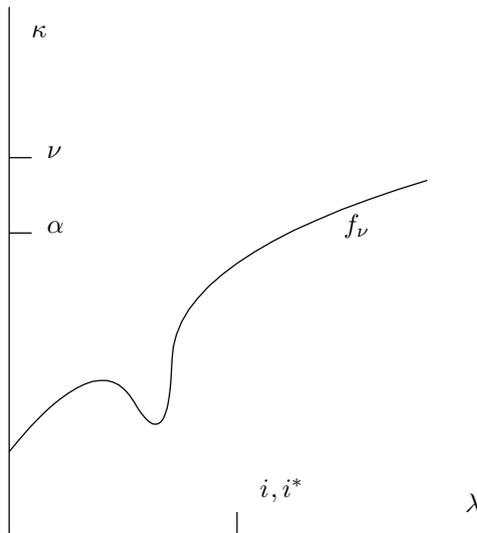
$$\alpha_i = \text{the smallest element } \alpha \text{ of } C \text{ such that } \forall j < i \ \alpha_j < \alpha.$$

Then $\alpha_\mu < \kappa$ and $\text{cof}(\alpha_\mu) = \mu$. Hence $C \cap E_\mu^\kappa \neq \emptyset$. \square

So $E_\omega^{\aleph_2}$ and $E_{\omega_1}^{\aleph_2}$ are disjoint stationary subsets of \aleph_2 . Actually one can find a lot of disjoint stationary sets, using Fodor's lemma.

Theorem 151. Let κ be a successor cardinal and let $S \subseteq \kappa$ be stationary. Then there is a family $(S_i \mid i < \kappa)$ of pairwise disjoint stationary subsets of S .

Proof. Let $\lambda \in \text{Card}$ such that $\kappa = \lambda^+$. For each $\nu < \kappa$, $\nu \neq 0$ choose a surjective function $f_\nu: \lambda \rightarrow \nu$.



(1) For every $\alpha < \kappa$ there is some $i < \lambda$ such that $\{\nu \in S \mid f_\nu(i) \geq \alpha\}$ is stationary in κ .

Proof. Assume for a contradiction that there is $\alpha < \kappa$ such that for any $i < \lambda$ the set $\{\nu \in S \mid f_\nu(i) \geq \alpha\}$ is non-stationary in κ . Choose cub sets C_i such that

$$\{\nu \in S \mid f_\nu(i) \geq \alpha\} \cap C_i = \emptyset.$$

The set $\bigcap_{i < \lambda} C_i$ is cub in κ . Let $\nu \in S \cap \bigcap_{i < \lambda} C_i$ and $\nu > \alpha$. Then $f_\nu(i) < \alpha$ for all $i < \lambda$, which contradicts the surjectivity of $f_\nu: \lambda \rightarrow \nu$. *qed*(1)

(2) There is some $i_* < \lambda$ such that for every $\alpha < \kappa$ the set $\{\nu \in S \mid f_\nu(i_*) \geq \alpha\}$ is stationary in κ .

Proof. By (1), we can find for every $\alpha < \kappa$ some $i_\alpha < \lambda$ such that $\{\nu \in S \mid f_\nu(i_\alpha) \geq \alpha\}$ is stationary in κ . By the pidgeon principle there is an unbounded subset $Z \subseteq \kappa$ and an $i_* < \lambda$ such that $\forall \alpha \in Z \ i_\alpha = i_*$. So for every $\alpha \in Z$ the set $\{\nu \in S \mid f_\nu(i_*) \geq \alpha\}$ is stationary in κ , which proves the claim. *qed*(2)

For $\beta < \kappa$ set $S_\beta = \{\nu \in S \mid f_\nu(i_*) = \beta\}$.

(3) The set of $\beta < \kappa$, where S_β is stationary in κ , is unbounded in κ .

Proof. Assume not and let $\alpha < \kappa$ such that S_β is stationary implies $\beta < \alpha$. By (2), $T = \{\nu \in S \mid f_\nu(i_*) \geq \alpha\}$ is stationary in κ . The function $\nu \mapsto f_\nu(i_*) < \nu$ is regressive on T . By Fodor's Theorem the function is constant on a stationary subset of T . Let $\beta, \alpha \leq \beta < \kappa$ be the constant value. Then $S_\beta = \{\nu \in S \mid f_\nu(i_*) = \beta\}$ is stationary in κ , contradiction. *qed*(3)

So there are κ -many $\beta < \kappa$ such that S_β is stationary. Note that these S_β are pairwise disjoint subsets of S . \square

Abstractly this means that every NS_κ -positive set can be split into κ -many NS_κ -positive sets. Consider the property: there are μ -many NS_κ -positive sets $(S_i \mid i < \mu)$ which are *almost disjoint* with respect to NS_κ : $i \neq j \rightarrow S_i \cap S_j \in \text{NS}_\kappa$. If this property is false, we say that the ideal NS_κ is μ -saturated. The property that NS_{\aleph_1} is \aleph_2 -saturated is not decided by ZFC. That property has many consequences and is central in modern set theoretic research.

16 Silver's Theorem

The value of 2^κ for regular cardinals κ is hardly determined by the value of 2^λ at other cardinals. The situation at singular cardinals is different, the first result in this area was proved by JACK SILVER. We shall use the notion of *almost disjoint* functions.

Definition 152. Let λ be a limit ordinal. Two functions $f, g: \lambda \rightarrow V$ are almost disjoint if there is $\alpha < \lambda$ such that $\forall \beta (\alpha < \beta < \lambda \rightarrow f(\beta) \neq g(\beta))$. A set $\mathcal{F} \subseteq {}^\lambda V$ of functions is almost disjoint if f and g are almost disjoint for any $f, g \in \mathcal{F}$, $f \neq g$.

Lemma 153. a) There is no almost disjoint family $\mathcal{F} \subseteq {}^\omega 2$ of size 3.

b) There is an almost disjoint family $\mathcal{F} \subseteq {}^\omega \omega$ of size 2^{\aleph_0} .

Proof. a) is obvious. b) Take $h: (\omega)^{<\omega} \leftrightarrow \omega$. For $a: \omega \rightarrow 2$ define $f_a: \omega \rightarrow \omega$ by

$$f_a(n) = h(a \upharpoonright n).$$

Consider functions $a, b: \omega \rightarrow 2$, $a \neq b$. We show that f_a and f_b are almost disjoint. Take $n < \omega$ such that $a \upharpoonright n \neq b \upharpoonright n$. Then for $n \leq m < \omega$ we have

$$f_a(m) = h(a \upharpoonright m) \neq h(b \upharpoonright m) = f_b(m).$$

Thus $\{f_a \mid a \in {}^\omega 2\}$ is an almost disjoint family of size 2^{\aleph_0} . \square

Theorem 154. (Silver) Let $\omega < \lambda = \text{cof}(\kappa) < \kappa \in \text{Card}$. Let $2^\mu = \mu^+$ for all $\omega \leq \mu \in \kappa \cap \text{Card}$. Then $2^\kappa = \kappa^+$.

So let us assume that $\omega < \lambda = \text{cof}(\kappa) < \kappa \in \text{Card}$ and $2^\mu = \mu^+$ for all $\omega \leq \mu \in \kappa \cap \text{Card}$. Fix a strictly increasing sequence $(\kappa_\alpha \mid \alpha < \lambda)$ which is cofinal in κ and continuous, i.e., for any limit ordinal $\delta < \lambda$ we have $\kappa_\delta = \bigcup_{\alpha < \delta} \kappa_\alpha$.

Lemma 155. *Assume that $\mu^\lambda < \kappa$ for all $\mu < \kappa$. Let $\mathcal{F} \subseteq \prod_{\alpha < \lambda} A_\alpha$ be almost disjoint, where $S_0 = \{\alpha < \lambda \mid \text{card}(A_\alpha) \leq \kappa_\alpha\}$ is stationary in λ . Then $\text{card}(\mathcal{F}) \leq \kappa$.*

Proof. Assume w.l.o.g. that $A_\alpha \subseteq \kappa_\alpha$ for $\alpha \in S$. For $f \in \mathcal{F}$ define $h_f: S_0 \rightarrow \lambda$ by

$$h_f(\alpha) := \text{the least } \beta \text{ such that } f(\alpha) \in \kappa_\beta.$$

Then $h \upharpoonright (S_0 \cap \text{Lim})$ is regressive. So, by Fodor's lemma, one can choose a stationary $S_f \subseteq S_0 \cap \text{Lim}$ such that h is constant on S_f . Hence f is, on S_f , bounded in κ . If $f \upharpoonright S_f = g \upharpoonright S_g$, then $f = g$, since \mathcal{F} is almost disjoint. So $f \mapsto f \upharpoonright S_f$ is one-to-one. For a fixed $T \subseteq \lambda$, the set of functions on T that are bounded in κ has cardinality $\sup\{\mu^\lambda \mid \mu < \kappa\} = \kappa$ by the cardinality assumption. Also there are $< \kappa$ such T , since $\text{card}(\wp(\lambda)) = 2^\lambda < \kappa$. Hence $\text{card}(\mathcal{F}) \leq \kappa \cdot \kappa = \kappa$. \square

Lemma 156. *Let $\omega < \lambda = \text{cof}(\kappa) < \kappa \in \text{Card}$, and assume that $\mu^\lambda < \kappa$ for all $\mu < \kappa$. Let $\mathcal{F} \subseteq \prod_{\alpha < \lambda} A_\alpha$ be almost disjoint, $\text{card}(A_\alpha) \leq \kappa_\alpha^+$. Then $\text{card}(\mathcal{F}) \leq \kappa^+$.*

Proof. Assume w.l.o.g. that $A_\alpha \subseteq \kappa_\alpha^+$. Let $S \subseteq \lambda$ be stationary and $f \in \mathcal{F}$. Let

$$\mathcal{F}_{f,S} = \{g \in \mathcal{F} \mid (\forall \alpha \in S)(g(\alpha) \leq f(\alpha))\}.$$

The map $g \mapsto g \upharpoonright S$ injects $\mathcal{F}_{f,S}$ into $\prod_{\alpha \in S} (f(\alpha) + 1)$ where $\text{card}(f(\alpha) + 1) \leq \kappa_\alpha$. By the previous lemma, $\text{card}(\mathcal{F}_{f,S}) \leq \kappa$.

Define

$$\mathcal{F}_f = \bigcup \{\mathcal{F}_{f,S} \mid S \subseteq \lambda \text{ is stationary}\}.$$

Since there are $< \kappa$ stationary subsets of λ ,

(1) $\text{card}(\mathcal{F}_f) \leq \kappa$.

We construct a sequence $(f_\xi \mid \xi < \delta)$ of functions in \mathcal{F} by induction such that $\mathcal{F} = \bigcup \{\mathcal{F}_{f_\xi} \mid \xi < \delta\}$.

Take an arbitrary $f_0 \in \mathcal{F}$.

If $(f_\nu \mid \nu < \xi)$ is already defined, choose $f_\xi \notin \bigcup \{\mathcal{F}_{f_\nu} \mid \nu < \xi\}$ if possible. If there is no such f_ξ , set $\delta = \xi$ and stop.

(2) $\delta \leq \kappa^+$.

Proof. Assume that f_{κ^+} is defined. If $\nu < \kappa^+$ then $f_{\kappa^+} \notin \mathcal{F}_{f_\nu}$. So $\{\alpha \mid f_{\kappa^+}(\alpha) \leq f_\nu(\alpha)\}$ is non-stationary and $\{\alpha \mid f_\nu(\alpha) \leq f_{\kappa^+}(\alpha)\} \in \mathcal{C}_\lambda$. Thus $f_\nu \in \mathcal{F}_{\kappa^+}$ for all $\nu < \xi$, and $\text{card}(\mathcal{F}_{f_{\kappa^+}}) \geq \kappa^+$. This contradicts (1). *qed*(2)

Altogether

$$\text{card}(\mathcal{F}) = \text{card}\left(\bigcup \{\mathcal{F}_{f_\xi} \mid \xi < \delta\}\right) \leq \sum_{\xi < \delta} \text{card}(\mathcal{F}_{f_\xi}) \leq \sum_{\xi < \delta} \kappa \leq \kappa \cdot \kappa^+ = \kappa^+.$$

\square

Proof. (Silver's Theorem) Define a map from $\mathcal{P}(\kappa)$ into $\prod_{\alpha < \lambda} \wp(\kappa_\alpha)$ by

$$X \mapsto f_X = (X \cap \kappa_\alpha \mid \alpha < \lambda).$$

If $X \neq Y$ then f_X and f_Y are almost disjoint. So

$$\mathcal{F} = \{f_X \mid X \in \mathcal{P}(\kappa)\} \subseteq \prod_{\alpha < \lambda} \wp(\kappa_\alpha)$$

is an almost disjoint family of functions. The GCH below κ implies that $\text{card}(\wp(\kappa_\alpha)) = 2^{\kappa_\alpha} = \kappa_\alpha^+$. Moreover $\mu^\lambda \leq \max(\mu, \lambda)^{\max(\mu, \lambda)} = \max(\mu, \lambda)^+ < \kappa$ for all $\mu < \kappa$. By Lemma 156, $\text{card}(\mathcal{F}) \leq \kappa^+$. Hence

$$\kappa^+ \leq 2^\kappa = \text{card}(\mathcal{P}(\kappa)) \leq \text{card}(\mathcal{F}) \leq \kappa^+$$

\square

Exercise 32. Use the methods of the proof of Silver's Theorem to show

- a) Let $\mathcal{F} \subseteq \prod_{\alpha < \lambda} A_\alpha$ be almost disjoint, $\text{card}(A_\alpha) \leq \kappa_\alpha^{++}$. Then $\text{card}(\mathcal{F}) \leq \kappa^{++}$.

b) Let $\omega < \lambda = \text{cof}(\kappa) < \kappa \in \text{Card}$. Let $2^\mu = \mu^{++}$ for all $\omega \leq \mu \in \kappa \cap \text{Card}$. Then $2^\kappa \leq \kappa^{++}$.

17 Ranks of functions

The previous exercise indicates the possibility that one may generalize Silver's theorem by a kind of induction on the height of the continuum function below κ . This idea will lead to the Galvin-Hajnal theorem. Let $\kappa = \aleph_\alpha$ be a singular strong limit cardinal (i.e., $\mu < \aleph_\alpha \rightarrow 2^\mu < \aleph_\alpha$) with $\omega < \lambda = \text{cof}(\aleph_\alpha) < \aleph_\alpha$. Let $(\kappa_\delta \mid \delta < \lambda)$ be a strictly increasing continuous sequence cofinal in \aleph_α . The continuum function below \aleph_α determines a function $\varphi_0: \lambda \rightarrow \alpha$ by

$$2^{\kappa_\delta} = \kappa_\delta^{+\varphi_0(\delta)},$$

where $\kappa_\delta^{+\varphi_0(\delta)} = \aleph_{\eta+\varphi_0(\delta)}$ if $\kappa_\delta = \aleph_\eta$.

Let us now study functions $\varphi: \lambda \rightarrow \alpha$ along appropriate wellfounded relations. Fix a regular uncountable cardinal λ and some limit ordinal α .

Definition 157. For a stationary set $S \subseteq \lambda$ define a relation $<_S$ on functions $\varphi, \psi: \lambda \rightarrow \alpha$ by

$$\varphi <_S \psi \text{ iff there is a cub } C \subseteq \lambda \text{ such that } \forall \delta \in S \cap C \varphi(\delta) < \psi(\delta).$$

Equivalently one can say that the set of δ where φ and ψ behaves differently is very small:

$$\varphi <_S \psi \text{ iff } \{\delta \in S \mid \varphi(\delta) \geq \psi(\delta)\} \in \text{NS}_\lambda.$$

Lemma 158. $<_S$ is a strongly wellfounded relation on ${}^\lambda\alpha$.

Proof. Assume not. Then, using AC, there is a strictly descending ω -sequence

$$\psi_0 >_S \psi_1 >_S \psi_2 >_S \dots$$

Choose cub sets $C_0, C_1, \dots \subseteq \lambda$ such that $\forall \delta \in S \cap C_n \psi_n(\delta) > \psi_{n+1}(\delta)$. Since $\bigcap_{n < \omega} C_n$ is cub in λ one can take $\delta \in S \cap \bigcap_{n < \omega} C_n$. Then

$$\psi_0(\delta) > \psi_1(\delta) > \psi_2(\delta) > \dots$$

is a descending ω -sequence of ordinals. Contradiction. \square

Definition 159. Define the $<_S$ -rank $\|\psi\|_S \in \text{Ord}$ of $\psi \in {}^\lambda\alpha$ by recursion on $<_S$:

$$\|\psi\|_S = \bigcup \{\|\varphi\|_S + 1 \mid \varphi <_S \psi\}.$$

We also write $\|\psi\|$ instead of $\|\psi\|_\lambda$.

We prove some properties of this rank.

Lemma 160.

- a) If $S \subseteq T$ are stationary in λ then $\varphi <_T \psi$ implies $\varphi <_S \psi$.
- b) If $S \subseteq T$ are stationary in λ then $\|\psi\|_T \leq \|\psi\|_S$.
- c) If S, T are stationary in λ then

$$\|\psi\|_{S \cup T} = \min(\|\psi\|_S, \|\psi\|_T).$$

- d) If S is stationary and N is nonstationary then $\varphi <_{S \cup N} \psi$ iff $\varphi <_S \psi$.
- e) If S is stationary and N is nonstationary then

$$\|\psi\|_{S \cup N} = \|\psi\|_S.$$

Proof. a) Let $\varphi <_T \psi$. Take $C \subseteq \lambda$ cub such that $\forall \delta \in T \cap C \varphi(\delta) < \psi(\delta)$. Then $\forall \delta \in S \cap C \varphi(\delta) < \psi(\delta)$ and so $\varphi <_S \psi$.

b) By induction on $<_T$.

$$\begin{aligned} \|\psi\|_T &= \bigcup \{ \|\varphi\|_T + 1 \mid \varphi <_T \psi \} \\ &\leq \bigcup \{ \|\varphi\|_S + 1 \mid \varphi <_T \psi \} \text{ by the inductive assumption} \\ &\leq \bigcup \{ \|\varphi\|_S + 1 \mid \varphi <_S \psi \} \text{ by a)} \\ &= \|\psi\|_S. \end{aligned}$$

c) By b) $\|\psi\|_{S \cup T} \leq \|\psi\|_S, \|\psi\|_T$ and so

$$\|\psi\|_{S \cup T} \leq \min(\|\psi\|_S, \|\psi\|_T).$$

Assume that the equality is false and consider $\psi <_{S \cup T}$ -minimal such that

$$\|\psi\|_{S \cup T} < \min(\|\psi\|_S, \|\psi\|_T).$$

Since $\|\psi\|_{S \cup T} \in \|\psi\|_S = \bigcup \{ \|\varphi\|_S + 1 \mid \varphi <_S \psi \}$ take $\psi_S <_S \psi$ such that $\|\psi\|_{S \cup T} < \|\psi_S\|_S + 1$, i.e., $\|\psi\|_{S \cup T} \leq \|\psi_S\|_S$; take $C_S \subseteq \lambda$ cub such that $\forall \delta \in S \cup C_S \psi_S(\delta) < \psi(\delta)$. Similarly take $\psi_T <_T \psi$ such that $\|\psi\|_{S \cup T} \leq \|\psi_T\|_T$ and some $C_T \subseteq \lambda$ cub such that $\forall \delta \in T \cup C_T \psi_T(\delta) < \psi(\delta)$. Define $\varphi: \lambda \rightarrow \alpha$,

$$\varphi(\delta) = \begin{cases} \psi_S(\delta) & \text{if } \delta \in S \setminus T \\ \psi_T(\delta) & \text{if } \delta \in T \setminus S \\ \max(\psi_S(\delta), \psi_T(\delta)) & \text{if } \delta \in S \cap T \\ 0 & \text{else} \end{cases}$$

For $\delta \in (S \cup T) \cap (C_S \cap C_T)$ holds $\varphi(\delta) < \psi(\delta)$, thus $\varphi <_{S \cup T} \psi$. Since $\forall \delta \in S \psi_S(\delta) \leq \varphi(\delta)$ we have $\|\psi_S\|_S \leq \|\varphi\|_S$. Similarly $\|\psi_T\|_T \leq \|\varphi\|_T$. By the $<_{S \cup T}$ -minimality of ψ we have

$$\|\varphi\|_{S \cup T} = \min(\|\varphi\|_S, \|\varphi\|_T) \geq \min(\|\psi_S\|_S, \|\psi_T\|_T) \geq \|\psi\|_{S \cup T}$$

contradicting $\varphi <_{S \cup T} \psi$.

d) $\varphi <_{S \cup N} \psi$ iff $\{\delta \in S \cup N \mid \varphi(\delta) \geq \psi(\delta)\} \in \text{NS}_\lambda$ iff $\{\delta \in S \mid \varphi(\delta) \geq \psi(\delta)\} \in \text{NS}_\lambda$ iff $\varphi <_S \psi$.

e) follows directly from d). □

Note that $\|\psi\| = \|\psi\|_\lambda \leq \|\psi\|_S$. This motivates to exclude stationary sets S which do not compute the “correct” rank of ψ .

Definition 161. Let

$$I_\psi = \text{NS}_\lambda \cup \{S \mid S \text{ is stationary and } \|\psi\| < \|\psi\|_S\}.$$

Lemma 162. I_ψ is an ideal on λ .

Proof. I_ψ is a non-empty family of subsets of λ and $\lambda \notin I_\psi$.

I_ψ is closed under subsets: let $A \in I_\psi$ and $B \subseteq A$. If B is nonstationary then $B \in I_\psi$. If B is stationary then $\|\psi\|_B \geq \|\psi\|_A > \|\psi\|$ and $B \in I_\psi$.

I_ψ is closed under unions: let $A \in I_\psi$ and $B \in I_\psi$. If A and B are nonstationary then $A \cup B \in \text{NS}_\lambda \subseteq I_\psi$. If A is stationary and B is nonstationary then by e) of Lemma 163 $\|\psi\|_{A \cup B} = \|\psi\|_A > \|\psi\|$ and $A \cup B \in I_\psi$. If A is stationary and B is stationary then by c) of Lemma 163

$$\|\psi\|_{A \cup B} = \min(\|\psi\|_A, \|\psi\|_B) > \|\psi\|$$

and $A \cup B \in I_\psi$. □

Lemma 163.

a) If $\|\psi\| = 0$ then $\{\delta < \lambda \mid \psi(\delta) = 0\}$ is stationary in λ .

b) If $\|\psi\|$ is a successor ordinal then

$$\{\delta < \lambda \mid \psi(\delta) \text{ is a successor ordinal}\} \notin I_\psi.$$

c) If $\|\psi\|$ is a limit ordinal then

$$\{\delta < \lambda \mid \psi(\delta) \text{ is a limit ordinal}\} \notin I_\psi.$$

Proof. a) Let $\|\psi\| = 0$. If $\{\delta < \lambda \mid \psi(\delta) = 0\}$ is nonstationary there is a cub $C \subseteq \lambda$ such that $\forall \delta \in C \psi(\delta) > 0$. But then $\psi \succ_\lambda \text{const}_0$ and $\|\psi\| \geq 1$. Contradiction.

b) Let $\|\psi\|$ be a successor ordinal but assume that

$$\{\delta < \lambda \mid \psi(\delta) \text{ is a successor ordinal}\} \in I_\psi.$$

Then

$$S = \{\delta < \lambda \mid \psi(\delta) \text{ is a limit ordinal}\} \notin I_\psi.$$

By the definition of I_ψ we get that $\|\psi\|_S = \|\psi\|$ is also a successor ordinal. By the definition of $\|\psi\|_S$ take $\varphi <_S \psi$ such that $\|\psi\|_S = \|\varphi\|_S + 1$. Take a cub set $C \subseteq \lambda$ such that $\forall \delta \in S \cap C \varphi(\delta) < \psi(\delta)$. Define $\varphi^+ : \lambda \rightarrow \alpha$ by $\varphi^+(\delta) = \varphi(\delta) + 1$. Since $\psi(\delta)$ is a limit ordinal for $\delta \in S$:

$$\forall \delta \in S \cap C \varphi(\delta) < \varphi^+(\delta) < \psi(\delta).$$

Then $\varphi <_S \varphi^+ <_S \psi$ and $\|\varphi\|_S < \|\varphi^+\|_S < \|\psi\|_S$, contradicting $\|\psi\|_S = \|\varphi\|_S + 1$.

c) Let $\|\psi\|$ be a limit ordinal but assume that

$$\{\delta < \lambda \mid \psi(\delta) \text{ is a limit ordinal}\} \in I_\psi.$$

Then

$$S = \{\delta < \lambda \mid \psi(\delta) \text{ is a successor ordinal}\} \notin I_\psi.$$

By the definition of I_ψ we get that $\|\psi\|_S = \|\psi\|$ is also a limit ordinal. Define $\psi^- : \lambda \rightarrow \alpha$ by

$$\psi^-(\delta) = \begin{cases} \psi(\delta) - 1, & \text{if } \delta \in S \\ 0, & \text{else} \end{cases}$$

Consider a function $\varphi <_S \psi$. Take a cub set $C \subseteq \lambda$ such that $\forall \delta \in S \cap C \varphi(\delta) < \psi(\delta)$ and thus $\forall \delta \in S \cap C \varphi(\delta) \leq \psi^-(\delta)$. Then

$$\|\varphi\|_S = \|\varphi\|_{S \cap C} \leq \|\psi^-\|_{S \cap C} = \|\psi^-\|_S < \|\psi\|_S$$

This means that

$$\begin{aligned} \|\psi\|_S &= \bigcup \{\|\varphi\|_S + 1 \mid \varphi <_S \psi\} \\ &\leq \|\psi^-\|_S + 1 \end{aligned}$$

Since $\|\psi^-\|_S < \|\psi\|_S$ this implies

$$\|\psi\| = \|\psi\|_S = \|\psi^-\|_S + 1$$

is a successor ordinal, contradiction. □

18 The Galvin Hajnal Theorem

Theorem 164. Let $\kappa = \aleph_\alpha$ be a singular strong limit cardinal (i.e., $\mu < \aleph_\alpha \rightarrow 2^\mu < \aleph_\alpha$) with $\omega < \lambda = \text{cof}(\aleph_\alpha) < \aleph_\alpha$. Then

$$2^{\aleph_\alpha} < \aleph_\gamma$$

where $\gamma = (2^{\text{card}(\alpha)})^+$.

Note that if $\alpha = \aleph_\alpha$ the theorem claims that

$$2^\alpha < \aleph_{(2^\alpha)^+}$$

which is obviously true.

Lemma 165. Let $\psi : \lambda \rightarrow \alpha$ and let $\mathcal{F} \subseteq \prod_{\delta < \lambda} A_\delta$ be an almost disjoint family of functions such that

$$\text{card}(A_\delta) \leq \kappa_\delta^{+\psi(\delta)}$$

for $\delta < \lambda$. Then $\text{card}(\mathcal{F}) \leq \aleph_{\alpha + \|\psi\|}$.

Proof. By induction on $\|\psi\| \in \text{Ord}$.

$\|\psi\| = 0$: Then by Lemma 163(a) ψ vanishes on a stationary set $S \subseteq \lambda$ and Lemma 155 proves the case.

$\|\psi\|$ is the successor ordinal $\gamma + 1$: by Lemma 163(b)

$$S_0 = \{\delta < \lambda \mid \psi(\delta) \text{ is a successor ordinal}\} \notin I_\psi.$$

We may assume that $A_\delta \subseteq \kappa_\delta^{+\psi(\delta)}$ for $\delta < \lambda$.

(1) Let $f \in \mathcal{F}$ and $S \subseteq S_0$, $S \notin I_\psi$. Then the set

$$\mathcal{F}_{f,S} = \{g \in \mathcal{F} \mid \forall \delta \in S \ g(\delta) \leq f(\delta)\}$$

has cardinality $\leq \aleph_{\alpha + \gamma}$.

Proof. Define $\varphi: \lambda \rightarrow \alpha$ by

$$\varphi(\delta) = \begin{cases} \psi(\delta) - 1, & \text{if } \delta \in S \\ \psi(\delta), & \text{else} \end{cases}$$

Since $S \notin I_\psi$,

$$\|\varphi\| \leq \|\varphi\|_S < \|\psi\|_S = \|\psi\| = \gamma + 1$$

and $\|\varphi\| \leq \gamma$. Now

$$\mathcal{F}_{f,S} \subseteq \prod_{\delta < \lambda} f(\delta),$$

and since $f(\delta) < \kappa_\delta^{+\psi(\delta)}$ we have

$$\text{card}(f(\delta) + 1) \leq \kappa_\delta^{+\varphi(\delta)}.$$

By the induction hypothesis, $\text{card}(\mathcal{F}_{f,S}) \leq \aleph_{\alpha + \|\varphi\|} \leq \aleph_{\alpha + \gamma}$. *qed(1)*

(2) Let $f \in \mathcal{F}$. Then the set

$$\mathcal{F}_f = \{g \in \mathcal{F} \mid \exists S \subseteq S_0 (S \notin I_\psi \wedge \forall \delta \in S \ g(\delta) \leq f(\delta))\}$$

has cardinality $\leq \aleph_{\alpha + \gamma}$.

Proof. Since \aleph_α is a strong limit cardinal, there are at most $2^\lambda < \aleph_\alpha$ many $S \subseteq S_0$. Hence

$$\mathcal{F}_f = \bigcup_{S \subseteq S_0, S \notin I_\psi} \mathcal{F}_{f,S}$$

is a union of less than \aleph_α many sets of size $\leq \aleph_{\alpha + \gamma}$. *qed(2)*

Like in the proof of Lemma 156 we construct a sequence $(f_\xi \mid \xi < \zeta)$ of functions in \mathcal{F} by induction such that $\mathcal{F} = \bigcup \{\mathcal{F}_{f_\xi} \mid \xi < \zeta\}$.

Take an arbitrary $f_0 \in \mathcal{F}$.

If $(f_\nu \mid \nu < \xi)$ is already defined, choose $f_\xi \notin \bigcup \{\mathcal{F}_{f_\nu} \mid \nu < \xi\}$ if possible. If there is no such f_ξ , set $\zeta = \xi$ and stop.

(3) $\zeta \leq \aleph_{\alpha + \gamma + 1}$.

Proof. Assume that $f_{\aleph_{\alpha + \gamma + 1}}$ is defined. Set $\eta = \aleph_{\alpha + \gamma + 1}$. Consider $\nu < \eta$. $f_\eta \notin \mathcal{F}_{f_\nu}$ and so there is no $S \subseteq S_0$, $S \notin I_\psi$ such that $\forall \delta \in S \ f_\eta(\delta) \leq f_\nu(\delta)$. This means that

$$\{\delta \in S \mid f_\eta(\delta) \leq f_\nu(\delta)\} \in I_\psi$$

and

$$\{\delta \in S \mid f_\nu(\delta) < f_\eta(\delta)\} \notin I_\psi.$$

This implies that $f_\nu \in \mathcal{F}_{f_\eta}$ and thus $\{f_\nu \mid \nu < \eta\} \subseteq \mathcal{F}_{f_\eta}$. Hence $\text{card}(\mathcal{F}_{f_\eta}) \geq \eta = \aleph_{\alpha + \gamma + 1}$, which contradicts (2). *qed(3)*

Now

$$\text{card}(\mathcal{F}) = \text{card}\left(\bigcup \{\mathcal{F}_{f_\xi} \mid \xi < \zeta\}\right) \leq \sum_{\xi < \zeta} \text{card}(\mathcal{F}_{f_\xi}) \leq \sum_{\xi < \zeta} \aleph_{\alpha + \gamma} \leq \sum_{\xi < \aleph_{\alpha + \gamma + 1}} \aleph_{\alpha + \gamma} = \aleph_{\alpha + \gamma + 1}.$$

Finally consider the case that $\|\psi\|$ is a limit cardinal. By Lemma 163(c),

$$S = \{\delta < \lambda \mid \psi(\delta) \text{ is a limit ordinal}\} \notin I_\psi.$$

Again we may assume that $A_\delta \subseteq \kappa_\delta^{+\psi(\delta)}$ for $\delta < \lambda$.

For functions $\varphi: \lambda \rightarrow \alpha$ define

$$\mathcal{F}_\varphi = \{f \in \mathcal{F} \mid \forall \delta < \lambda \ f(\delta) \in \kappa_\delta^{+\varphi(\delta)}\}.$$

Consider $f \in \mathcal{F}$. Define $\varphi: \lambda \rightarrow \alpha$ by taking $\varphi(\delta)$ minimal such that

$$f(\delta) \in \kappa_\delta^{+\varphi(\delta)}.$$

Obviously $f \in \mathcal{F}_\varphi$. Moreover $\varphi <_S \psi$. Since $S \notin I_\psi$ we have

$$\|\varphi\| \leq \|\varphi\|_S < \|\psi\|_S = \|\psi\|.$$

By the induction hypothesis

$$\text{card}(\mathcal{F}_\varphi) \leq \aleph_{\alpha+\|\varphi\|} < \aleph_{\alpha+\|\psi\|}.$$

Thus

$$\mathcal{F} \subseteq \bigcup \{\mathcal{F}_\varphi \mid \exists \varphi (\varphi: \lambda \rightarrow \alpha \wedge \text{card}(\mathcal{F}_\varphi) < \aleph_{\alpha+\|\psi\|})\}$$

and

$$\text{card}(\mathcal{F}) \leq \sum_{\varphi: \lambda \rightarrow \alpha} \aleph_{\alpha+\|\varphi\|} \leq \text{card}({}^\lambda \alpha) \cdot \aleph_{\alpha+\|\psi\|} \leq \aleph_\alpha \cdot \aleph_{\alpha+\|\psi\|} = \aleph_{\alpha+\|\psi\|}.$$

□

We are now able to prove

Theorem. Let $\kappa = \aleph_\alpha$ be a singular strong limit cardinal (i.e., $\mu < \aleph_\alpha \rightarrow 2^\mu < \aleph_\alpha$) with $\omega < \lambda = \text{cof}(\aleph_\alpha) < \aleph_\alpha$. Then

$$2^{\aleph_\alpha} < \aleph_\gamma$$

where $\gamma = (2^{\text{card}(\alpha)})^+$.

Proof. Define an injective map from $\mathcal{P}(\kappa)$ into $\prod_{\delta < \lambda} \wp(\kappa_\delta)$ by

$$X \mapsto f_X = (X \cap \kappa_\delta \mid \delta < \lambda).$$

If $X \neq Y$ then f_X and f_Y are almost disjoint. So

$$\mathcal{F} = \{f_X \mid X \in \mathcal{P}(\kappa)\} \subseteq \prod_{\delta < \lambda} \wp(\kappa_\delta)$$

is an almost disjoint family of functions. Since κ is a strong limit cardinal, there is a function $\psi: \lambda \rightarrow \alpha$ such that

$$\text{card}(\wp(\kappa_\delta)) \leq \kappa_\delta^{+\psi(\delta)}$$

for $\delta < \lambda$. By the previous lemma

$$\text{card}(\mathcal{P}(\kappa)) \leq \text{card}(\mathcal{F}) \leq \aleph_{\alpha+\|\psi\|}.$$

Since $\text{card}({}^\lambda \alpha) \leq \text{card}({}^\alpha \alpha) = 2^{\text{card}(\alpha)}$, the rank function $\|\cdot\|$ on ${}^\lambda \alpha$ is bounded by $\gamma = (2^{\text{card}(\alpha)})^+$. Hence $\|\psi\| < \gamma$ and

$$2^{\aleph_\alpha} = \text{card}(\mathcal{P}(\kappa)) \leq \aleph_{\alpha+\|\psi\|} < \aleph_{\alpha+\gamma} = \aleph_\gamma.$$

□

If, e.g., \aleph_{ω_1} is a strong limit cardinal then by this theorem

$$2^{\aleph_{\omega_1}} < \aleph_{(2^{\aleph_{\omega_1}})^+} < \aleph_{\aleph_{\omega_1}}.$$

So the continuum function at singular cardinals can be influenced by the behaviour below that cardinal. In particular instances the bounds for the continuum function can be improved. With considerably more effort one can also deal with singular cardinals of *countable* cofinality and prove, e.g.: if \aleph_ω is a strong limit cardinal then (Shelah)

$$2^{\aleph_\omega} < \aleph_{\aleph_4}.$$

19 Measurable cardinals

The results of the previous section used filters and ideals to express that certain sets are large or small respectively. There are also intermediate notions of size: a set X is of “positive measure” if it is not in the ideal under consideration. One may imagine that the measure of X is some positive real number. This poses the question, which kinds of “measures” do or can exist. Ideally every set should be given some non-negative number as a measure.

This approach is also motivated by the classical theory of Lebesgue measure on the real line and related spaces. Recall that the 1-dimensional Lebesgue measure on \mathbb{R} is a function $l: \Omega \rightarrow \mathbb{R} \cup \{\infty\}$ taking values in the extended real line with the properties:

- a) $\Omega \subseteq \mathcal{P}(\mathbb{R})$ contains all intervals and is closed under complements and countable unions;
- b) $l([0, 1]) = 1$;
- c) l is *countably additive* (σ -*additive*): if $\{X_i \mid i < \omega\} \subseteq \Omega$ is a pairwise disjoint family then

$$l\left(\bigcup_{i < \omega} X_i\right) = \sum_{i < \omega} l(X_i);$$

- d) l is *translation invariant*: if $X \in \Omega$ and $d \in \mathbb{R}$ then $X + d = \{x + d \mid x \in X\} \in \Omega$ and

$$l(X + d) = l(X).$$

Theorem 166. *There is a set $Z \subseteq [0, 1]$ which is not Lebesgue-measurable, i.e., $Z \notin \Omega$.*

Proof. Let

$$A = \{\mathbb{Q} + d \mid d \in \mathbb{R}\}.$$

(1) A consists of pairwise disjoint nonempty sets which intersect the interval $[0, 1]$.

Proof. Assume that $x \in (\mathbb{Q} + d) \cap (\mathbb{Q} + e)$. Take rational numbers $r, s \in \mathbb{Q}$ such that

$$x = r + d = s + e.$$

Then $d = (s - r) + e \in \mathbb{Q} + e$ and

$$\mathbb{Q} + d = \{t + d \mid t \in \mathbb{Q}\} = \{t + (s - r) + e \mid t \in \mathbb{Q}\} \subseteq \mathbb{Q} + e.$$

Similarly $\mathbb{Q} + e \subseteq \mathbb{Q} + d$ and so $\mathbb{Q} + e = \mathbb{Q} + d$.

Consider $d \in \mathbb{R}$. Take an integer $z \in \mathbb{Z}$ such that $z \leq d < z + 1$. Then

$$-z + d \in (\mathbb{Q} + d) \cap [0, 1]$$

qed(1)

By the axiom of choice let Z be a choice set for the set

$$\{(\mathbb{Q} + d) \cap [0, 1] \mid d \in \mathbb{R}\}.$$

(2) If $q, r \in \mathbb{Q}$ and $q \neq r$ then $(Z + q) \cap (Z + r) = \emptyset$.

Proof. Assume not, and take $z_0, z_1 \in Z$ such that $z_0 + q = z_1 + r$. Then $z_0 \in \mathbb{Q} + z_0$ and $z_1 \in \mathbb{Q} + z_0$. Since Z is a choice set, $z_0 = z_1$. But then $q = r$. Contradiction. *qed(2)*

(3) $[0, 1] \subseteq \bigcup_{q \in [-1, 1] \cap \mathbb{Q}} Z + q$.

Proof. Let $d \in [0, 1]$. Let $z \in Z \cap (\mathbb{Q} + d) \cap [0, 1]$. Take $q \in \mathbb{Q}$ such that $z = q + d$. Then $d = z + (-q)$ where $|q| \leq 1$. *qed(3)*

Assume now that $Z \in \Omega$. Since $Z \subseteq [0, 1]$ we have $l(Z) \leq 1$.

Case 1: $l(Z) = 0$. Then

$$1 = l([0, 1]) \leq l\left(\bigcup_{q \in [-1, 1] \cap \mathbb{Q}} Z + q\right) = \sum_{q \in [-1, 1] \cap \mathbb{Q}} l(Z + q) = \sum_{q \in [-1, 1] \cap \mathbb{Q}} l(Z) = \sum_{q \in [-1, 1] \cap \mathbb{Q}} 0 = 0,$$

contradiction.

Case 2: $l(Z) = \varepsilon > 0$. Then

$$l\left(\bigcup_{q \in [-1, 1] \cap \mathbb{Q}} Z + q\right) \leq l([0, 2]) = 2$$

but on the other hand

$$l\left(\bigcup_{q \in [-1,1] \cap \mathbb{Q}} Z + q\right) = \sum_{q \in [-1,1] \cap \mathbb{Q}} l(Z + q) = \sum_{q \in [-1,1] \cap \mathbb{Q}} l(Z) = \sum_{q \in [-1,1] \cap \mathbb{Q}} \varepsilon = \infty.$$

□

We shall now consider measures which are defined on *all* subsets of a given set, but we do not require a geometric structure on the set and in particular no translation invariance. We also restrict our consideration to *finite* measures.

Definition 167. A measure on a set X is a function $\mu: \mathcal{P}(X) \rightarrow [0, 1]$ such that

- a) $\mu(\emptyset) = 0$ and $\mu(X) = 1$;
- b) l is countably additive (σ -additive): if $\{X_i \mid i < \omega\} \subseteq \mathcal{P}(X)$ is a pairwise disjoint family then

$$\mu\left(\bigcup_{i < \omega} X_i\right) = \sum_{i < \omega} \mu(X_i).$$

μ is called non-trivial if $\mu(\{x\}) = 0$ for every $x \in X$. μ is 2-valued if $\text{ran}(\mu) = \{0, 1\}$, otherwise μ is real-valued.

Note that if $f: \kappa \leftrightarrow X$ is a bijection then a measure μ on X immediately induces a measure μ' on κ by

$$\mu'(A) = \mu(f[A]).$$

So we can focus our attention on measures on *cardinals*.

Lemma 168. Every 2-valued measure μ on \mathbb{R} is trivial, i.e., there is some $x \in \mathbb{R}$ such that

$$\mu(A) = 1 \text{ iff } x_0 \in A.$$

Proof. Identify \mathbb{R} with ${}^\omega 2$. Define $x_0: \omega \rightarrow 2$ by

$$x_0(n) = 1 \text{ iff } \mu(\{x \in {}^\omega 2 \mid x(n) = 1\}) = 1.$$

By the σ -additivity of μ ,

$$\mu({}^\omega 2 \setminus \{x_0\}) = \mu\left(\bigcup_{n < \omega} \{x \in {}^\omega 2 \mid x(n) \neq x_0(n)\}\right) \leq \sum_{n < \omega} \mu(\{x \in {}^\omega 2 \mid x(n) \neq x_0(n)\}) = \sum_{n < \omega} 0 = 0.$$

So $\mu(\{x_0\}) = 1$. □

Assume that κ is the smallest cardinal which has a non-trivial measure μ . A set $A \subseteq \kappa$ with $\mu(A) > 0$ *splits* if there is a *partition* $A_1, A_2 \subseteq A$ such that $A_1 \cup A_2 = A$, $A_1 \cap A_2 = \emptyset$, $0 < \mu(A_1) < \mu(A)$ and $0 < \mu(A_2) < \mu(A)$.

Case 1. There is a set $A_0 \subseteq \kappa$ with $\mu(A_0) > 0$ which does *not* split.

Then define $\nu: \mathcal{P}(A_0) \rightarrow 2$ by

$$\nu(A) = 1 \text{ iff } \mu(A) = \mu(A_0).$$

(1) ν is a 2-valued non-trivial measure on A_0 .

Proof. We have to check σ -additivity. Let $\{X_i \mid i < \omega\} \subseteq \mathcal{P}(A_0)$ be a pairwise disjoint family. By the σ -additivity of μ

$$\mu\left(\bigcup_{i < \omega} X_i\right) = \sum_{i < \omega} \mu(X_i).$$

Then

$$\nu\left(\bigcup_{i < \omega} X_i\right) = 1 \text{ iff } \mu\left(\bigcup_{i < \omega} X_i\right) = \mu(A_0) \text{ iff } \exists i < \omega \mu(X_i) = \mu(A_0) \text{ iff } \exists i < \omega \nu(X_i) = 1.$$

Thus

$$\nu\left(\bigcup_{i<\omega} X_i\right) = \sum_{i<\omega} \nu(X_i).$$

qed(1)

By the minimality of κ we have $\text{card}(A_0) = \kappa$.

Case 2. Every set $A \subseteq \kappa$ with $\mu(A) > 0$ splits. In this case we call the measure μ *atomless*.

We first show that indeed A splits into relatively large subsets:

(2) Every set $A \subseteq \kappa$ with $\mu(A) > 0$ possesses a subset $B \subseteq A$ such that $\frac{1}{3}\mu(A) \leq \mu(B) \leq \frac{2}{3}\mu(A)$.

Proof. Assume not. Then

$$\delta = \sup \left\{ \mu(B) \mid B \subseteq A \wedge \mu(B) \leq \frac{1}{2}\mu(A) \right\} \leq \frac{1}{3}\mu(A).$$

For $n \in \omega \setminus \{0\}$ choose $B_n \subseteq A$ such that $\delta - \frac{1}{n} < \mu(B_n) \leq \delta$.

We show by induction on n that

$$\mu(B_1 \cup B_2 \cup \dots \cup B_n) \leq \delta.$$

Assume that $\mu(B_1 \cup B_2 \cup \dots \cup B_n) \leq \delta$.

Assume for a contradiction that $\mu(B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1}) > \delta$. Then

$$\mu(B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1}) \leq \mu(B_1 \cup B_2 \cup \dots \cup B_n) + \mu(B_{n+1}) \leq \delta + \delta \leq \frac{2}{3}\mu(A).$$

By the initial assumption we cannot have $\mu(B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1}) = \frac{2}{3}\mu(A)$. Hence

$$\delta < \mu(B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1}) < \frac{2}{3}\mu(A)$$

and

$$\mu(A \setminus (B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1})) > \frac{1}{3}\mu(A) \geq \delta.$$

But then $B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1}$ or its relative complement $A \setminus (B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1})$ would be a counterexample to the definition of δ . Thus $\mu(B_1 \cup B_2 \cup \dots \cup B_n \cup B_{n+1}) \leq \delta$. The σ -additivity of μ implies

$$\begin{aligned} \mu\left(\bigcup_{1 \leq n < \omega} B_n\right) &= \mu\left(\bigcup_{1 \leq n < \omega} (B_n \setminus (B_1 \cup \dots \cup B_{n-1}))\right) \\ &= \sum_{1 \leq n < \omega} \mu(B_n \setminus (B_1 \cup \dots \cup B_{n-1})) \\ &= \lim_{m \rightarrow \infty} \sum_{n=1}^m \mu(B_n \setminus (B_1 \cup \dots \cup B_{n-1})) \\ &= \lim_{m \rightarrow \infty} \mu\left(\bigcup_{n=1}^m (B_n \setminus (B_1 \cup \dots \cup B_{n-1}))\right) \\ &= \lim_{m \rightarrow \infty} \mu(B_1 \cup \dots \cup B_m) \\ &= \delta. \end{aligned}$$

Set $B_* = \bigcup_{1 \leq n < \omega} B_n$. Then $\mu(B_*) = \delta$. $A \setminus B_*$ splits, so take a partition $A \setminus B_* = C \dot{\cup} D$ such that $0 < \mu(C) \leq \mu(D) < \mu(A \setminus B_*) = \mu(A) - \delta$. By the initial assumption we have $\mu(C) < \frac{1}{3}\mu(A)$ or $\mu(C) > \frac{2}{3}\mu(A)$, and $\mu(D) < \frac{1}{3}\mu(A)$ or $\mu(D) > \frac{2}{3}\mu(A)$.

If $\mu(D) < \frac{1}{3}\mu(A)$ then

$$\mu(A) = \mu(B_* \cup C \cup D) = \mu(B_*) + \mu(C) + \mu(D) < \frac{1}{3}\mu(A) + \frac{1}{3}\mu(A) + \frac{1}{3}\mu(A) = \mu(A),$$

contradiction. Hence $\mu(C) < \frac{1}{3}\mu(A)$ and $\mu(D) > \frac{2}{3}\mu(A)$. But then

$$\delta < \mu(B_* \cup C) = \mu(A \setminus D) = \mu(A) - \mu(D) < \mu(A) - \frac{2}{3}\mu(A) < \frac{1}{2}\mu(A),$$

contradicting the definition of δ . *qed(2)*

Recall the binary tree

$${}^{<\omega}2 = \{s \mid \exists n < \omega \ s : n \rightarrow 2\}.$$

We construct a *binary splitting* $A: {}^{<\omega}2 \rightarrow \mathcal{P}(\kappa)$ of the underlying set κ by recursion on the length of the binary sequences. Put $A(\emptyset) = A_0 = \kappa$. If $A(s) = A_s \subseteq \kappa$ is constructed, use (2) to choose a splitting $A_s = A_{s0} \dot{\cup} A_{s1}$ of A_s such that $\frac{1}{3} \mu(A_s) \leq \mu(A_{s0}) \leq \mu(A_{s1}) \leq \frac{2}{3} \mu(A_s)$.

We shall pull the measure μ back to a measure ν on the reals. For $X \subseteq \mathbb{R}$ define

$$X' = \bigcup_{x \in X} \bigcap_{n < \omega} A_{x \upharpoonright n}.$$

Define $\nu: \mathcal{P}({}^{\omega}2) \rightarrow [0, 1]$ by

$$\nu(X) = \mu(X').$$

First we show that the assignment $X \mapsto X'$ preserves some set theoretic properties.

(3) $\emptyset' = \emptyset$.

(4) $({}^{\omega}2)' = \kappa$.

Proof. Let $\alpha \in \kappa$. Define $x: \omega \rightarrow 2$ recursively by

$$x(n) = 1 \text{ iff } \alpha \in A_{(x \upharpoonright n)1}.$$

Then $\alpha \in \bigcap_{n < \omega} A_{x \upharpoonright n}$ and

$$\kappa = \bigcup_{x \in {}^{\omega}2} \bigcap_{n < \omega} A_{x \upharpoonright n}.$$

qed(4)

(5) $(X \cap Y)' = X' \cap Y'$.

Proof. Let $\alpha \in (X \cap Y)'$. Take $x \in X \cap Y$ such that $\alpha \in \bigcap_{n < \omega} A_{x \upharpoonright n}$. Then

$$\alpha \in \bigcup_{x \in X} \bigcap_{n < \omega} A_{x \upharpoonright n} = X'$$

and also $\alpha \in Y'$.

Conversely consider $\alpha \in X' \cap Y'$. Take $x \in X$ such that $\alpha \in \bigcap_{n < \omega} A_{x \upharpoonright n}$ and take $y \in Y$ such that $\alpha \in \bigcap_{n < \omega} A_{y \upharpoonright n}$. Assume for a contradiction that $x \neq y$. Take $n \in \omega$ such that $x \upharpoonright n = y \upharpoonright n$ and $x(n) \neq y(n)$. Then $\alpha \in A_{x \upharpoonright (n+1)} \cap A_{y \upharpoonright (n+1)}$ although $A_{x \upharpoonright (n+1)} \cap A_{y \upharpoonright (n+1)} = \emptyset$ by construction.

Thus $x = y \in X \cap Y$ and

$$\alpha \in \bigcup_{x \in X \cap Y} \bigcap_{n < \omega} A_{x \upharpoonright n} = (X \cap Y)'.$$

qed(5)

(6) $(\bigcup_{i \in I} X_i)' = \bigcup_{i \in I} X_i'$.

Proof.

$$\begin{aligned} \left(\bigcup_{i \in I} X_i \right)' &= \bigcup_{x \in \bigcup_{i \in I} X_i} \bigcap_{n < \omega} A_{x \upharpoonright n} \\ &= \bigcup_{i \in I} \bigcup_{x \in X_i} \bigcap_{n < \omega} A_{x \upharpoonright n} \\ &= \bigcup_{i \in I} X_i'. \end{aligned}$$

qed(6)

(7) ν is a non-trivial measure on ${}^{\omega}2$.

Proof. $\nu(\emptyset) = \mu(\emptyset') = \mu(\emptyset) = 0$ and $\nu({}^{\omega}2) = \mu(({}^{\omega}2)') = \mu(\kappa) = 1$.

To check σ -additivity consider a pairwise disjoint family $\{X_i \mid i < \omega\} \subseteq \mathcal{P}({}^{\omega}2)$. Then $\{X_i' \mid i < \omega\} \subseteq \mathcal{P}(\kappa)$ is pairwise disjoint by (5) and (3). By the σ -additivity of μ and by (6),

$$\nu \left(\bigcup_{i < \omega} X_i \right) = \mu \left(\bigcup_{i < \omega} X_i' \right) = \sum_{i < \omega} \mu(X_i') = \sum_{i < \omega} \nu(X_i).$$

To check non-triviality consider $x \in {}^{\omega}2$. Then

$$\{x\}' = \bigcap_{n < \omega} A_{x \upharpoonright n}.$$

One can show inductively that

$$\mu(A_{x \upharpoonright n}) \leq \left(\frac{2}{3}\right)^n$$

for all $n < \omega$. Thus

$$\mu(\{x\}') \leq \mu(A_{x \upharpoonright n}) \leq \left(\frac{2}{3}\right)^n$$

for all n and so $\mu(\{x\}') = 0$. This implies

$$\nu(\{x\}) = \mu(\{x\}') = 0.$$

qed(7)

Since κ was assumed to be minimal carrying a non-trivial measure, we have $\kappa \leq 2^{\aleph_0}$ in *Case 2*.

In both cases we can show a strong additivity property:

(8) The measure μ on κ is κ -additive, i.e., for every pairwise disjoint family $\{X_i \mid i < \gamma\} \subseteq \mathcal{P}(\kappa)$ with $\gamma < \kappa$ we have

$$\mu\left(\bigcup_{i < \gamma} X_i\right) = \sum_{i < \gamma} \mu(X_i),$$

where the right hand side is defined as

$$\sum_{i < \gamma} \mu(X_i) = \sup \left\{ \sum_{i \in I_0} \mu(X_i) \mid I_0 \text{ is a finite subset of } \gamma \right\}.$$

Proof. Assume that μ is not κ -additive and let $\gamma < \kappa$ be least such that there is a family $\{X_i \mid i < \gamma\} \subseteq \mathcal{P}(\kappa)$ with

$$\mu\left(\bigcup_{i < \gamma} X_i\right) \neq \sum_{i < \gamma} \mu(X_i).$$

Then γ is an uncountable cardinal and

$$\mu\left(\bigcup_{i < \gamma} X_i\right) > \sum_{i < \gamma} \mu(X_i).$$

Let

$$J = \{i < \gamma \mid \mu(X_i) > 0\} \subseteq \gamma.$$

We claim that J is at most countable. If J were uncountable, there must be some rational number $\frac{1}{n}$ such that $\mu(X_i) > \frac{1}{n}$ for uncountably many $i \in J$. But then

$$\mu\left(\bigcup_{i < \gamma} X_i\right) = \infty,$$

contradiction.

The σ -additivity of μ entails

$$\begin{aligned} \mu\left(\bigcup_{i \in \gamma \setminus J} X_i\right) &= \mu\left(\bigcup_{i < \gamma} X_i\right) - \mu\left(\bigcup_{i \in J} X_i\right) \\ &> \sum_{i < \gamma} \mu(X_i) - \sum_{i \in J} \mu(X_i) \\ &= \sum_{i < \gamma} \mu(X_i) - \sum_{i < \gamma} \mu(X_i) \\ &= 0. \end{aligned}$$

So we may assume that *all* elements of the disjoint family $\{X_i \mid i < \gamma\} \subseteq \mathcal{P}(\kappa)$ have μ -measure 0.

We shall pull the measure μ back to a measure ν on γ . Let $\mu_0 = \mu(\bigcup_{i \in \gamma} X_i)$. Define $\nu: \mathcal{P}(\gamma) \rightarrow [0, 1]$ by

$$\nu(Y) = \frac{1}{\mu_0} \mu\left(\bigcup_{i \in Y} X_i\right)$$

We show that ν is a non-trivial measure on γ .

$$\nu(\emptyset) = \frac{1}{\mu_0} \mu(\emptyset) = 0 \text{ and } \nu(\gamma) = \frac{1}{\mu_0} \mu\left(\bigcup_{i \in \gamma} X_i\right) = 1.$$

To check σ -additivity consider a pairwise disjoint family $\{Y_j \mid j < \omega\} \subseteq \mathcal{P}(\gamma)$. Then

$$\left\{ \bigcup_{i \in Y_j} X_i \mid j < \omega \right\} \subseteq \mathcal{P}(\kappa)$$

is pairwise disjoint. By the σ -additivity of μ

$$\begin{aligned} \nu\left(\bigcup_{j < \omega} Y_j\right) &= \frac{1}{\mu_0} \mu\left(\bigcup_{i \in \bigcup_{j < \omega} Y_j} X_i\right) \\ &= \frac{1}{\mu_0} \mu\left(\bigcup_{j < \omega} \bigcup_{i \in Y_j} X_i\right) \\ &= \frac{1}{\mu_0} \sum_{j < \omega} \mu\left(\bigcup_{i \in Y_j} X_i\right) \\ &= \sum_{j < \omega} \frac{1}{\mu_0} \mu\left(\bigcup_{i \in Y_j} X_i\right) \\ &= \sum_{j < \omega} \nu(Y_j). \end{aligned}$$

Finally, ν is non-trivial since for every $i < \gamma$

$$\nu(\{i\}) = \frac{1}{\mu_0} \mu\left(\bigcup_{j \in \{i\}} X_j\right) = \frac{1}{\mu_0} \mu(X_i) = 0.$$

But the existence of ν contradicts the minimality of κ . *qed*(8)

We can now draw conclusions from the previous arguments.

Definition 169. A cardinal κ is measurable if κ is uncountable and possesses a 2-valued κ -additive non-trivial measure. A cardinal κ is real-valued measurable if κ is uncountable and possesses a non-atomic κ -additive non-trivial measure.

Recall that μ is non-atomic, if every sets $A \subseteq \kappa$ with $\mu(A) > 0$ splits, i.e., there is a partition $A_1, A_2 \subseteq A$ such that $A_1 \cup A_2 = A$, $A_1 \cap A_2 = \emptyset$, $0 < \mu(A_1) < \mu(A)$ and $0 < \mu(A_2) < \mu(A)$.

Theorem 170. Let κ be minimal such that κ carries a non-trivial measure. Then either κ is a measurable cardinal, or $\kappa \leq 2^{\aleph_0}$ and κ is a real-valued measurable cardinal.

Proof. If we are in *Case 1* above, then κ is measurable. In *Case 2*, $\kappa \leq 2^{\aleph_0}$. By Lemma 168 there is no 2-valued non-trivial measure on κ . Hence κ is real-valued measurable. \square

Lemma 171. Let κ be measurable or real-valued measurable. Then κ is regular.

Proof. Let μ be a non-trivial κ -additive measure on κ . Assume for a contradiction that $\text{cof}(\kappa) = \gamma < \kappa$. Let $(\kappa_i \mid i < \gamma) \subseteq \kappa$ be cofinal in κ . For $i < \gamma$ we have

$$\mu(\kappa_i) = \mu\left(\bigcup_{\alpha < \kappa_i} \{\alpha\}\right) = \sum_{\alpha < \kappa_i} \mu(\{\alpha\}) = \sum_{\alpha < \kappa_i} 0 = 0,$$

using non-triviality and κ -additivity. Similarly

$$\mu(\kappa) = \mu\left(\bigcup_{i < \gamma} \kappa_i\right) \leq \sum_{i < \gamma} \mu(\kappa_i) = \sum_{i < \gamma} 0 = 0,$$

contradiction. \square

Lemma 172. *Let κ be a measurable cardinal. Then κ is strongly inaccessible.*

Proof. Let μ be a non-trivial κ -additive 2-valued measure on κ . Assume for a contradiction that $\lambda < \kappa$ but $2^\lambda \geq \kappa$. We may assume that μ is a measure on ${}^\lambda 2$. Define $x_0: \lambda \rightarrow 2$ by

$$x_0(i) = 1 \text{ iff } \mu(\{y \in {}^\lambda 2 \mid y(i) = 1\}) = 1.$$

The κ -additivity of μ implies

$$\begin{aligned} \mu({}^\lambda 2 \setminus \{x_0\}) &= \mu\left(\bigcup_{i < \lambda} \{y \in {}^\lambda 2 \mid y(i) = 1 - x_0(i)\}\right) \\ &\leq \sum_{i < \lambda} \mu(\{y \in {}^\lambda 2 \mid y(i) = 1 - x_0(i)\}) \\ &= \sum_{i < \lambda} 0 \\ &= 0. \end{aligned}$$

But then $\mu(\{x_0\}) = 1$ which contradicts the non-triviality of μ . □

Thus measurable cardinals are *large cardinals*. Large cardinals are central notions in set theory. They can be viewed as “ideal points” in the cardinal hierarchy with respect to certain properties. A strongly inaccessible cardinal κ is an ideal point of cardinal arithmetic. κ cannot be reached by the formation of cardinal powers or even infinitary sums and products from smaller parameters.

One could “cut off” the universe of sets at κ and restrict consideration to V_κ as a “subuniverse”. We shall see in later courses that V_κ is a model of the ZFC-axioms and one could restrict mathematics to working inside V_κ . On the other hand the assumption of inaccessible and stronger large cardinals like measurable cardinals greatly enriches set theoretic combinatorics. There are also isolated instances, when the assumption of large cardinals influences the behaviour of smaller sets like the set of real numbers.

20 Normal measures and ultrafilters

We shall now combine techniques from filters and ideals with measures on measurable cardinals.

Lemma 173. *Let κ be an uncountable cardinal. Then κ is a measurable cardinal iff there is non-principal, κ -complete ultrafilter U on κ , i.e., $\forall i < \kappa \{i\} \notin U$ (non-principality) and for every family $\{X_i \mid i < \gamma\} \subseteq U$ with $\gamma < \kappa$ we have*

$$\bigcap_{i < \gamma} X_i \in U.$$

Proof. If μ is a 2-valued κ -additive non-trivial measure μ on κ then

$$U = \{X \subseteq \kappa \mid \mu(X) = 1\}$$

is an ultrafilter with the desired properties. Conversely, if \tilde{U} is a non-principal, κ -complete ultrafilter on κ then $\tilde{\mu}: \mathcal{P}(\kappa) \rightarrow 2$ defined by

$$\tilde{\mu}(X) = 1 \text{ iff } X \in \tilde{U}$$

is a 2-valued κ -additive non-trivial measure. □

Let us fix a measurable cardinal κ and a non-principal, κ -complete ultrafilter U on κ .

Definition 174. *Define a relation $<_U$ on functions $f, g: \kappa \rightarrow \kappa$ by*

$$f <_U g \text{ iff } \{i < \kappa \mid f(i) < g(i)\} \in U.$$

Recall that we defined in some previous arguments

$$\varphi <_S \psi \text{ iff } \{\delta \in S \mid \varphi(\delta) \geq \psi(\delta)\} \in \text{NS}_\lambda.$$

Lemma 175. $<_U$ is a strongly wellfounded on ${}^\kappa\kappa$.

Proof. Assume not and assume that $A \subseteq {}^\kappa\kappa$, $A \neq \emptyset$ does not have a $<_U$ -minimal element. Then define recursively a sequence f_n for $n < \omega$: choose $f_0 \in A$ arbitrary; if $f_n \in A$ is defined, it is not $<_U$ -minimal in A , and so we can choose (AC) $f_{n+1} \in A$ such that $f_{n+1} <_U f_n$. For $n < \omega$ set

$$X_n = \{i < \kappa \mid f_{n+1}(i) < f_n(i)\} \in U.$$

Since U is κ -complete,

$$\bigcap_{n < \omega} X_n \in U.$$

So one can take $i_0 \in \bigcap_{n < \omega} X_n$. So for each $n < \omega$

$$f_{n+1}(i_0) < f_n(i_0),$$

and $(f_n(i_0))_{n < \omega}$ is a strictly decreasing ω -chain in the ordinals. Contradiction. \square

Definition 176. Define the $<_U$ -rank $\|f\|_U \in \text{Ord}$ of $f \in {}^\kappa\kappa$ by recursion on $<_U$:

$$\|f\|_U = \bigcup \{\|g\|_U + 1 \mid g <_U f\}.$$

Lemma 177.

a) For $\alpha < \kappa$ and for every $f \in {}^\kappa\kappa$

$$\|f\|_U = \alpha \text{ iff } \{i < \kappa \mid f(i) = \alpha\} \in U.$$

Hence $\|c_\alpha\|_U = \alpha$ where $c_\alpha: \kappa \rightarrow \kappa$ is the constant function $c_\alpha(i) = \alpha$.

b) There exists $f \in {}^\kappa\kappa$ such that $\|f\|_U \geq \kappa$

Proof. a) By induction on $\alpha < \kappa$. Assume a) holds for $\beta < \alpha$. Let $f \in {}^\kappa\kappa$.

Assume that $\|f\|_U = \alpha$ but $X_\alpha = \{i < \kappa \mid f(i) = \alpha\} \notin U$. By inductive hypothesis, also $X_\beta = \{i < \kappa \mid f(i) = \beta\} \notin U$ for all $\beta < \alpha$. Since U is κ -complete, any union of less than κ many sets which are not in the ultrafilter is not in the ultrafilter. Thus

$$\{i < \kappa \mid f(i) \leq \alpha\} = \bigcup_{\beta \leq \alpha} X_\beta \notin U$$

and

$$\{i < \kappa \mid f(i) > \alpha\} \in U.$$

This implies that

$$c_\alpha <_U f.$$

By inductive hypothesis, $\|c_\alpha\|_U \neq \beta$ for all $\beta < \alpha$. Hence $\|c_\alpha\|_U \geq \alpha$. But then

$$\|f\|_U > \|c_\alpha\|_U \geq \alpha,$$

contradiction.

Conversely assume that $A = \{i < \kappa \mid f(i) = \alpha\} \in U$. Then $\{i < \kappa \mid f(i) = \beta\} \notin U$ for all $\beta < \alpha$ and by the inductive hypothesis $\|f\|_U \neq \beta$ for every $\beta < \alpha$. Hence $\|f\|_U \geq \alpha$. Assume for a contradiction that $\|f\|_U > \alpha$. Then there exists $g \in {}^\kappa\kappa$ such that $g <_U f$ and $\|g\|_U \geq \alpha$. By the inductive assumption, the set

$$X_\beta = \{i < \kappa \mid g(i) = \beta\} \notin U$$

for every $\beta < \alpha$. By the κ -completeness of U

$$\{i < \kappa \mid g(i) < \alpha\} = \bigcup_{\beta < \alpha} X_\beta \notin U$$

and

$$B = \{i < \kappa \mid g(i) \geq \alpha\} \in U.$$

Since $g <_U f$, the set $C = \{i < \kappa \mid g(i) < f(i)\} \in U$. Now take $i \in A \cap B \cap C$. Then

$$\alpha \leq g(i) < f(i) = \alpha,$$

contradiction.

b) Let $d: \kappa \rightarrow \kappa$ be the *diagonal function* $d(i) = i$. For $\alpha < \kappa$

$$\{i < \kappa \mid c_\alpha(i) < d(i)\} = \{i < \kappa \mid \alpha < i\} = \bigcap_{\beta \leq \alpha} (\kappa \setminus \{\beta\}) \in U$$

by the non-principality and κ -completeness of U . Hence $c_\alpha <_U d$ and $\|d\|_U \geq \|c_\alpha\|_U = \alpha$. Thus $\|d\|_U \geq \kappa$. \square

Consider a function $f: \kappa \rightarrow \kappa$ with $\|f\|_U = \kappa$. Then a function $g: \kappa \rightarrow \kappa$ with $g <_U f$ has rank $\|g\|_U = \alpha < \kappa$, i.e.,

$$\{i < \kappa \mid g(i) = \alpha\} \in U.$$

This resembles Fodor's theorem by which *regressive* functions are constant on stationary sets. Indeed one can modify the filter U to come even closer to Fodor's theorem.

Define $\tilde{U} \subseteq \mathcal{P}(\kappa)$ by

$$X \in \tilde{U} \text{ iff } f^{-1}(X) = \{i < \kappa \mid f(i) \in X\} \in U.$$

(1) \tilde{U} is an ultrafilter on κ .

Proof. $\emptyset \notin \tilde{U}$ since $f^{-1}(\emptyset) = \emptyset \notin U$.

If $X \in \tilde{U}$ and $X \subseteq Y \subseteq \kappa$, then $f^{-1}(X) \in U$ and $f^{-1}(X) \subseteq f^{-1}(Y)$, and so $f^{-1}(Y) \in U$ and $Y \in \tilde{U}$.

If $X, Y \in \tilde{U}$, then $f^{-1}(X), f^{-1}(Y) \in U$ and $f^{-1}(X \cap Y) = f^{-1}(X) \cap f^{-1}(Y) \in U$, thus $X \cap Y \in \tilde{U}$.

If $X \subseteq \kappa$ and $X \notin \tilde{U}$, then $f^{-1}(X) \notin U$ and $\kappa \setminus f^{-1}(X) \in U$. Hence $f^{-1}(\kappa \setminus X) = \kappa \setminus f^{-1}(X) \in U$ and $\kappa \setminus X \in \tilde{U}$. *qed*(1)

Note that the argument is based on the preservation of the set theoretic notions $\emptyset, \subseteq, \cap, \setminus$ by the $f^{-1}(\cdot)$ -operation.

(2) \tilde{U} is κ -complete.

Proof. Let $\{X_i \mid i < \delta\} \subseteq \tilde{U}$ where $\delta < \kappa$. Then $\{f^{-1}(X_i) \mid i < \delta\} \subseteq U$ and $\bigcap_{i < \delta} f^{-1}(X_i) \in U$ by the κ -completeness of U . Since

$$f^{-1}\left(\bigcap_{i < \delta} X_i\right) = \bigcap_{i < \delta} f^{-1}(X_i)$$

we get $\bigcap_{i < \delta} X_i \in \tilde{U}$. *qed*(2)

(3) \tilde{U} is non-principal.

Proof. Let $\alpha < \kappa$. By Lemma 177(a) $\{i < \kappa \mid f(i) = \alpha\} \notin U$, so

$$f^{-1}(\{\alpha\}) = \{i < \kappa \mid f(i) = \alpha\} \notin U$$

and $\{\alpha\} \notin \tilde{U}$. *qed*(3)

Definition 178. An ultrafilter U on κ is normal if every function $h: \kappa \rightarrow \kappa$ which is regressive on a set in U , i.e., $\{i < \kappa \mid h(i) < i\} \in U$, is constant on a set in U , i.e., there is a $\delta < \kappa$ such that $\{i < \kappa \mid h(i) = \delta\} \in U$.

(4) \tilde{U} is normal.

Proof. Let $h: \kappa \rightarrow \kappa$ and $\{i < \kappa \mid h(i) < i\} \in \tilde{U}$. Then

$$f^{-1}(\{i < \kappa \mid h(i) < i\}) = \{j < \kappa \mid h(f(j)) < f(j)\} \in U$$

and $h \circ f \in_U f$. Since $\|f\|_U = \kappa$, $\|h \circ f\|_U = \delta$ for some $\delta < \kappa$. By Lemma 177(a) this is equivalent to

$$\{j < \kappa \mid h(f(j)) = \delta\} \in U.$$

So

$$f^{-1}(\{i < \kappa \mid h(i) = \delta\}) = \{j < \kappa \mid h(f(j)) = \delta\} \in U$$

and

$$\{i < \kappa \mid h(i) = \delta\} \in \tilde{U}.$$

qed(4)

Hence we have shown

Theorem 179. κ is a measurable cardinal iff there is normal non-principal, κ -complete ultrafilter of κ .

Normal ultrafilter have better combinatorial properties like

Lemma 180. Let κ be measurable carrying a normal non-principal, κ -complete ultrafilter U . Then

- a) $\mathcal{C}_\kappa \subseteq U$ where \mathcal{C}_κ is the closed unbounded filter on κ .
- b) U is closed unter diagonal intersections, i.e., if $\{X_i \mid i < \kappa\} \subseteq U$ then

$$\bigtriangleup_{i < \kappa} X_i \in U.$$

Proof. a) Let $C \subseteq \kappa$ be closed unbounded in κ but assume that $C \notin U$. Define $h: \kappa \rightarrow \kappa$ by

$$h(i) = \max(C \cap i).$$

Then h is regressive on $\kappa \setminus C \in U$. Since U is normal, take $\delta < \kappa$ such that

$$Y = \{i < \kappa \mid h(i) = \delta\} \in U.$$

Since Y is unbounded in κ , $\max(C \cap \kappa)$ must be equal to δ . But this contradicts the unboundedness of C in κ .

b) Let $\{X_i \mid i < \kappa\} \subseteq U$ but assume that $\bigtriangleup_{i < \kappa} X_i \notin U$. Define $h: \kappa \rightarrow \kappa$ by

$$h(j) = \min \{i < j \mid j \notin X_i\} \text{ if this exists and } h(j) = 0 \text{ else.}$$

Then h is regressive on $\kappa \setminus \bigtriangleup_{i < \kappa} X_i \in U$. Since U is normal, take $\delta < \kappa$ such that

$$Y = \{j < \kappa \mid h(j) = \delta\} \in U.$$

Then

$$X_\delta \cap Y \cap \left(\kappa \setminus \bigtriangleup_{i < \kappa} X_i \right) \subseteq X_\delta \cap \{j < \kappa \mid j \notin X_\delta\} = \emptyset.$$

This contradicts the fact that U is an ultrafilter for which the intersection of three elements has to also lie in the ultrafilter. \square

21 Measurable cardinals and the GCH

We use the methods of Silver's theorem in the context of measurable cardinals. Let κ be a measurable cardinal and fix a normal non-principal κ -complete ultrafilter U on κ .

Lemma 181. $\|d\|_U = \kappa$ where $d: \kappa \rightarrow \kappa$ is the diagonal function $d(i) = i$.

Proof. $\|d\|_U \geq \kappa$ was shown in the proof of 177(b).

For the converse consider $f <_U d$. Then

$$\{i < \kappa \mid f(i) < i\} \in U$$

means that f is regressive on a set in U . By the normality of U there is a constant $\delta < \kappa$ such that

$$\{i < \kappa \mid f(i) = \delta\} \in U.$$

By 177(a), $\|f\|_U = \delta$. Thus

$$\|d\|_U = \bigcup \{\|f\|_U + 1 \mid f <_U d\} \leq \bigcup \kappa = \kappa.$$

□

Definition 182. Two functions $f, g: \kappa \rightarrow V$ are U -almost disjoint if

$$\{i < \kappa \mid f(i) \neq g(i)\} \in U.$$

A family $\mathcal{F} \subseteq {}^\kappa V$ of functions is U -almost disjoint if f and g are U -almost disjoint for any $f, g \in \mathcal{F}$, $f \neq g$.

Lemma 183. Let \mathcal{F} be a U -almost disjoint family where $f <_U d$ for every $f \in \mathcal{F}$. Then $\text{card}(\mathcal{F}) \leq \kappa$.

Proof. By the previous Lemma, $f \mapsto \|f\|_U$ maps \mathcal{F} into κ . Moreover, this assignment is injective: if $\|f\|_U = \|g\|_U$ then

$$\{i < \kappa \mid f(i) = \|f\|_U\} \cap \{i < \kappa \mid g(i) = \|g\|_U\} \in U$$

and since \mathcal{F} is a U -almost disjoint family, $f = g$. Thus $\text{card}(\mathcal{F}) \leq \kappa$. □

Lemma 184. Let $d^+: \kappa \rightarrow \kappa$ be the cardinal successor function $d^+(i) = i^+$. Let \mathcal{F} be a U -almost disjoint family where $f <_U d^+$ for every $f \in \mathcal{F}$. Then $\text{card}(\mathcal{F}) \leq \kappa^+$.

Proof. Let $f \in \mathcal{F}$. Let

$$\mathcal{F}_f = \{g \in \mathcal{F} \mid g <_U f\} \cup \{f\}.$$

(1) $\text{card}(\mathcal{F}_f) \leq \kappa$.

Proof. For $i < \kappa$ with $f(i) < i^+$ choose an injection $h_i: f(i) \rightarrow i$. For $g \in \mathcal{F}_f$ define $g': \kappa \rightarrow V$ by

$$g'(i) = \begin{cases} h_i(g(i)), & \text{if } g(i) < f(i) < i^+ \\ g(i), & \text{else} \end{cases}$$

Then $g \mapsto g'$ is an injective map from \mathcal{F}_f into $\{g' \mid g \in \mathcal{F}_f\}$. $\{g' \mid g \in \mathcal{F}_f\}$ is an U -almost disjoint family where $g' <_U d$ for every $g \in \mathcal{F}_f$. By the previous lemma, $\text{card}(\{g' \mid g \in \mathcal{F}_f\}) \leq \kappa$ and so $\text{card}(\mathcal{F}_f) \leq \kappa$. *qed*(1)

We construct a sequence $(f_\xi \mid \xi < \zeta)$ of functions in \mathcal{F} by induction such that $\mathcal{F} = \bigcup \{\mathcal{F}_{f_\xi} \mid \xi < \zeta\}$.

Take an arbitrary $f_0 \in \mathcal{F}$ (wlog. $\mathcal{F} \neq \emptyset$).

If $(f_\nu \mid \nu < \xi)$ is already defined, choose $f_\xi \notin \bigcup \{\mathcal{F}_{f_\nu} \mid \nu < \xi\}$ if possible. If there is no such f_ξ , set $\zeta = \xi$ and stop.

(2) $\zeta \leq \kappa^+$.

Proof. Assume that f_{κ^+} is defined. If $\nu < \kappa^+$ then $f_{\kappa^+} \notin \mathcal{F}_{f_\nu}$. So $f_{\kappa^+} \not<_U f_\nu$. By the various assumptions on \mathcal{F} : $f_\nu <_U f_{\kappa^+}$. Thus $f_\nu \in \mathcal{F}_{\kappa^+}$ for all $\nu < \xi$, and $\text{card}(\mathcal{F}_{f_{\kappa^+}}) \geq \kappa^+$. This contradicts (1). *qed*(2)

Altogether

$$\text{card}(\mathcal{F}) = \text{card}\left(\bigcup \{\mathcal{F}_{f_\xi} \mid \xi < \zeta\}\right) \leq \sum_{\xi < \zeta} \text{card}(\mathcal{F}_{f_\xi}) \leq \sum_{\xi < \zeta} \kappa \leq \kappa \cdot \kappa^+ = \kappa^+.$$

□

Theorem 185. Let κ be a measurable cardinal and assume that $2^\lambda = \lambda$ for $\lambda < \kappa$. Then

$$2^\kappa = \kappa^+.$$

Proof. Fix a normal non-principal κ -complete ultrafilter U on κ .

For $i < \kappa$ choose a map $h_i: \mathcal{P}(i) \rightarrow \kappa$ which injects $\mathcal{P}(i)$ into i^+ for infinite i . Define a map from $\mathcal{P}(\kappa)$ into ${}^\kappa\kappa$ by

$$X \mapsto f_X = (h_i(X \cap i) \mid i < \kappa) \in {}^\kappa\kappa.$$

If $X \neq Y$ then f_X and f_Y are almost disjoint and hence U -almost disjoint. Also $f_X <_U d^+$. So by the previous Lemma

$$\text{card}(\mathcal{P}(\kappa)) \leq \text{card}(\{f_X \mid X \in \mathcal{P}(\kappa)\}) \leq \kappa^+.$$

□

22 Partition properties and partition cardinals

Definition 186. Let X be a term X . For a natural number $n < \omega$ let

$$[X]^n = \{a \subseteq X \mid \text{card}(a) = n\}$$

be the collection of n -element subsets of X . Further let

$$[X]^{<\omega} = \{a \subseteq X \mid \text{card}(a) < \omega\}$$

be the collection of finite subsets of X .

A function $f: [X]^n \rightarrow V$ or $f: [X]^{<\omega} \rightarrow V$ is called a partition of $[X]^n$ or $[X]^{<\omega}$ respectively.

Theorem 187. (Ramsey) Let $n < \omega$ and let $f: [\omega]^n \rightarrow 2$. Then there is an infinite $X \subseteq \omega$ such that $f \upharpoonright [X]^n$ is constant. The set is called homogeneous for f .

Proof. By induction on n . The claim is trivial for $n = 0$, and it is an easy instance of an infinitary pigeon principle in case $n = 1$. So assume the claim for n and let $f: [\omega]^{n+1} \rightarrow 2$ be given. We shall find an “ $n + 1$ -dimensional” homogeneous set for f by taking “ n -dimensional” homogeneous sets for “sections” f_x of f . For $x \in \omega$ define $f_x: [\omega \setminus (x+1)]^n \rightarrow 2$ by

$$f_x(s) = f(\{x\} \cup s).$$

We define sequences $x_0, x_1, \dots, c_0, c_1, \dots$ and X_0, X_1, \dots by simultaneous recursion such that

- a) $x_0 < x_1 < \dots < \omega$
- b) $x_i < \min(X_i), x_{i+1} \in X_i$
- c) $c_0, c_1, \dots \in \{0, 1\}$
- d) $\omega \supseteq X_0 \supseteq X_1 \supseteq \dots$ are infinite homogeneous sets for f_{x_0}, f_{x_1}, \dots respectively such that $\forall s \in [X_i]^n f_{x_i}(s) = c_i$

Set $x_0 = 0$. By the inductive assumption take X_0 to be an infinite homogeneous set for f_{x_0} and take $c_0 \in \{0, 1\}$ such that $\forall s \in [X_0]^n f_{x_0}(s) = c_0$.

If x_i and X_i are defined, take $x_{i+1} \in X_i$. Then use the inductive assumption and take $X_{i+1} \subseteq X_i \setminus (x_{i+1} + 1)$ to be an infinite homogeneous set for the function $f_{x_{i+1}} \upharpoonright [X_i \setminus (x_{i+1} + 1)]^n$. Take c_{i+1} such that $\forall s \in [X_{i+1}]^n f_{x_{i+1}}(s) = c_{i+1}$.

By the pigeon principle there is an infinite set $X \subseteq \{x_0, x_1, \dots\}$ and a $c \in \{0, 1\}$ such that $\forall x_i \in X c_i = c$.

We show that X is homogeneous for f with constant value c . Let $t \in [X]^{n+1}$. Let $x_i = \min(t)$ and $s = t \setminus \{x_i\}$. Then $s \in [X_i]^n$ and

$$f(t) = f_{x_i}(s) = c_i = c.$$

□

Then X is homogeneous for $f: \bigcup_{n \leq i < \omega} [\lambda]^i \rightarrow 2$ if for every $i \in [n, \omega)$ and every $x, y \in [X]^i$

$$f(x) = f(y)$$

Theorem 188. *Rowbottom's Theorem*

Definition 189. *Partition calculus: the notation $\beta \rightarrow (\alpha)_\lambda^n$ and $\beta \rightarrow (\alpha)_\lambda^{<\omega}$.*

Theorem 190. *Infinite exponent partitions do not have infinite homogeneous sets.*

Definition 191. *Weakly compact, Erdős cardinals $\kappa(\alpha)$, Ramsey cardinals*

Theorem 192. *Measurable cardinals are Ramsey*

Lemma 193. *Let $\alpha \geq \omega$ and $n < \omega$. Then $\beta \rightarrow (\alpha)^{<\omega}$ iff there is a function $g: [\lambda]^{n \leq i < \omega} = \bigcup_{n \leq i < \omega} [\lambda]^i \rightarrow 2$ without a homogeneous subset of order type α , i.e., there is no $X \subseteq \lambda$ with $\text{otp}(X) = \alpha$ such that for all $i \in [n, \omega)$ and $x, y \in [X]^i$ we have $g(x) = g(y)$.*

Proof. The implication from right to left is trivial. Conversely take $f: [\lambda]^{<\omega} \rightarrow 2$ which does not have a homogeneous subset of order α . Define $g: \bigcup_{n \leq i < \omega} [\lambda]^i \rightarrow 2$ by

$$g(x) = f(x \upharpoonright n),$$

where denotes the initial segment of x of order type n . Assume that g has a homogeneous subset X of order type α . Let $Y = X \setminus X \upharpoonright n$. Since α is infinite, $\text{otp}(Y) = \alpha$. We show for a contradiction that Y is homogeneous for f . Let $i < \omega$ and $x, y \in [Y]^i$. Then by the homogeneity of X

$$f(x) = g(X \upharpoonright n \cup x) = g(X \upharpoonright n \cup y) = f(y). \quad \square$$

Lemma 194. *Assume $\lambda \rightarrow (\alpha)^{<\omega}$ with α infinite. Then $2^\lambda \rightarrow (\alpha + 1)^{<\omega}$. In case λ is a limit ordinal also $2^\lambda \rightarrow (\alpha)^{<\omega}$.*

Proof. Take $f: [\lambda]^{3 \leq n < \omega} \rightarrow 2$ without a homogeneous subset of order type α . We then define $g: [{}^\lambda 2]^{2 \leq n < \omega} \rightarrow 2$. Let $<$ well-order ${}^\lambda 2$. Let \prec denote the lexicographic ordering of ${}^\lambda 2$. We want to ensure that on a homogeneous set \prec agrees with $<$ or with the converse ordering $>$. For $\{x, y\} \in [{}^\lambda 2]^2$ with $x < y$ set

$$g(\{x, y\}) = 0 \text{ iff } x \prec y.$$

The lexicographic order is defined via first differences. For $x, y \in {}^\lambda 2$ with $x \neq y$ let

$$\gamma(x, y) = \min \{ \xi \mid x(\xi) \neq y(\xi) \};$$

then $x \prec y$ iff $0 = x(\gamma(x, y)) < y(\gamma(x, y)) = 1$.

Now we define another auxiliary function which will turn a homogeneous set in ${}^\lambda 2$ into a homogeneous set in λ . Consider $\{x, y, z\} \in [{}^\lambda 2]^3$ with $x \prec y \prec z$. Since the functions are 2-valued we cannot have $\gamma(x, y) = \gamma(y, z)$.

If $\gamma(x, y) < \gamma(y, z)$ then y and z agree up to and including $\gamma(x, y)$. So the first differences $\gamma(x, y)$ and $\gamma(x, z)$ agree and

$$\gamma(x, y) = \gamma(x, z) < \gamma(y, z).$$

If $\gamma(y, z) < \gamma(x, y)$ then x and y agree up to and including $\gamma(y, z)$. So the first differences $\gamma(y, z)$ and $\gamma(x, z)$ agree and

$$\gamma(y, z) = \gamma(x, z) < \gamma(x, y)$$

According to these cases we define for $x \prec y \prec z$

$$g(\{x, y, z\}) = \begin{cases} 0 & \text{if } \gamma(x, y) = \gamma(x, z) < \gamma(y, z) \\ 1 & \text{if } \gamma(y, z) = \gamma(x, z) < \gamma(x, y) \end{cases}$$

So for $x \prec y$ from a g -homogeneous set, the ordinal $\gamma(x, y)$ will only depend on the first argument in case $g \equiv 1$ or, resp., on the second in case $g \equiv 0$.

So for $s \in [{}^\lambda 2]^{4 \leq i < \omega}$ from a g -homogeneous set the set

$$\gamma(s) = \{ \gamma(x, y) \mid x, y \in s, x \prec y \}$$

will have $\text{card}(s) - 1$ many elements.

Define $g: [{}^\lambda 2]^{4 \leq n < \omega} \rightarrow 2$ by

$$g(s) = \begin{cases} f(\gamma(s)) & \text{if } \text{card}(\gamma(s)) \geq 3 \\ 0 & \text{else} \end{cases}$$

Set $\beta = \alpha + 1$ if α is a successor ordinal, and $\beta = \alpha$ if α is a limit ordinal. We show that g witnesses $2^\lambda \not\rightarrow (\beta)^{<\omega}$.

So assume for a contradiction that $X \subseteq {}^\lambda 2$ has order type β with respect to $<$ and that X is homogeneous for g .

Case 1. $g[[X]^2] = \{0\}$. This means that $<$ and \prec agree on X so that X has order type β with respect to \prec . If $g[[X]^3] = \{1\}$, we have $\gamma(x, y) > \gamma(x, z)$ for $x \prec y \prec z$ from X . Taking an increasing ω -sequence $x_0 \prec x_1 \prec_2 \dots$ from X this yields a contradiction in the form of a descending ω -sequence of ordinals:

$$\gamma(x_0, x_1) > \gamma(x_0, x_2) > \gamma(x_0, x_3) \dots$$

Therefore $g[[X]^3] = \{0\}$. Then $\gamma(x, y)$ for $x, y \in X, x \prec y$ only depends on x and $\gamma(X) = \{\gamma(x, y) \mid x, y \in X, x \prec y\} \subseteq \text{Ord}$ has order type α . We obtain a contradiction by showing that $\gamma(X)$ is homogeneous for f : let $3 \leq n < \omega$ and consider two ascending n -sequences from $\gamma(X)$:

$$\gamma(x_0, y_0) < \dots < \gamma(x_{n-1}, y_{n-1}) \text{ and } \gamma(x'_0, y'_0) < \dots < \gamma(x'_{n-1}, y'_{n-1}).$$

Then $x_0 \prec x_1 \prec \dots \prec x_{n-1} \prec y_{n-1}$ and $x'_0 \prec x'_1 \prec \dots \prec x'_{n-1} \prec y'_{n-1}$ in X and

$$\begin{aligned} f(\{\gamma(x_0, y_0), \dots, \gamma(x_{n-1}, y_{n-1})\}) &= f(\{\gamma(x_0, y_{n-1}), \dots, \gamma(x_{n-1}, y_{n-1})\}) \\ &= f(\gamma(\{x_0, x_1, \dots, x_{n-1}, y_{n-1}\})) \\ &= g(\{x_0, x_1, \dots, x_{n-1}, y_{n-1}\}) \\ &= g(\{x'_0, x'_1, \dots, x'_{n-1}, y'_{n-1}\}) \\ &= f(\gamma(\{x'_0, x'_1, \dots, x'_{n-1}, y'_{n-1}\})) \\ &= f(\{\gamma(x'_0, y'_{n-1}), \dots, \gamma(x'_{n-1}, y'_{n-1})\}) \\ &= f(\{\gamma(x'_0, y'_0), \dots, \gamma(x'_{n-1}, y'_{n-1})\}). \end{aligned}$$

Case 2. $g[[X]^2] = \{1\}$. This means that $<$ and \succ agree on X so that X has order type β with respect to \succ . If $g[[X]^3] = \{0\}$, we have $\gamma(x, z) < \gamma(y, z)$ for $x \prec y \prec z$ from X . Taking a decreasing ω -sequence $x_0 \succ x_1 \succ_2 \dots$ from X this yields a contradiction in the form of a descending ω -sequence of ordinals:

$$\gamma(x_1, x_0) > \gamma(x_2, x_0) > \gamma(x_3, x_0) \dots$$

Therefore $g[[X]^3] = \{1\}$. Then $\gamma(x, y)$ for $x, y \in X, x \prec y$ only depends on y and $\gamma(X) = \{\gamma(x, y) \mid x, y \in X, x \prec y\} \subseteq \text{Ord}$ has order type β . We obtain a contradiction by showing that $\gamma(X)$ is homogeneous for f : let $3 \leq n < \omega$ and consider two ascending n -sequences from $\gamma(X)$:

$$\gamma(x_0, y_0) < \dots < \gamma(x_{n-1}, y_{n-1}) \text{ and } \gamma(x'_0, y'_0) < \dots < \gamma(x'_{n-1}, y'_{n-1}).$$

Then $x_0 \prec y_0 \prec \dots \prec y_{n-1}$ and $x'_0 \prec y'_0 \prec \dots \prec y'_{n-1}$ in X and

$$\begin{aligned} f(\{\gamma(x_0, y_0), \dots, \gamma(x_{n-1}, y_{n-1})\}) &= f(\{\gamma(x_0, y_0), \dots, \gamma(x_0, y_{n-1})\}) \\ &= f(\gamma(\{x_0, y_0, \dots, y_{n-1}\})) \\ &= g(\{x_0, y_0, \dots, y_{n-1}\}) \\ &= g(\{x'_0, y'_0, \dots, y'_{n-1}\}) \\ &= f(\gamma(\{x'_0, y'_0, \dots, y'_{n-1}\})) \\ &= f(\{\gamma(x'_0, y'_0), \dots, \gamma(x'_0, y'_{n-1})\}) \\ &= f(\{\gamma(x'_0, y'_0), \dots, \gamma(x'_{n-1}, y'_{n-1})\}). \end{aligned}$$

for $x \prec y \prec z$

$$g(\{x, y, z\}) = \begin{cases} 0 & \text{if } \gamma(x, y) = \gamma(x, z) < \gamma(y, z) \\ 1 & \text{if } \gamma(y, z) = \gamma(x, z) < \gamma(x, y) \end{cases}$$

□

Theorem 195. *If $\alpha \in \text{Lim}$ then $\kappa(\alpha)$ is strongly inaccessible.*

Theorem 196. *If κ is a measurable cardinal with normal measure U . Then the set*

$$\{\lambda < \kappa \mid \lambda \text{ is a Ramsey cardinal}\} \in U.$$

Proof. Otherwise $X = \{\lambda < \kappa \mid \lambda \text{ is not a Ramsey cardinal}\} \in U$. We can assume that every $\lambda \in X$ is an infinite cardinal. For every $\lambda \in X$ choose a function $f_\lambda: [\lambda]^{<\omega} \rightarrow 2$ which does not have a homogeneous subset of order type λ . Define $f: [\kappa]^{<\omega} \rightarrow 2$ by

$$f(x) = \begin{cases} f_{\max(x)}(x \setminus \max(x)) & \text{if } x \neq \emptyset \\ 0 & \text{else} \end{cases}$$

Let $Y \in U$ be homogeneous for f .

(1) $\{\lambda \in Y \mid \text{otp}(Y \cap \lambda) = \lambda\} \in U$.

Proof. Otherwise one can define a regressive function $\lambda \mapsto \text{otp}(Y \cap \lambda)$ on a measure one set in U which contradicts that $\text{otp}(Y) = \kappa$. *qed*(1)

So let $\lambda \in Y$ with $\text{otp}(Y \cap \lambda) = \lambda$. It is easy to see that $Y \cap \lambda$ is homogeneous for f_λ : if $n < \omega$ and $x, y \in [Y \cap \lambda]^{<\omega}$ then

$$f_\lambda(x) = f(x \cup \{\lambda\}) = f(y \cup \{\lambda\}) = f_\lambda(y).$$

Contradiction. □

So we obtain the following impression of the large cardinal *hierarchy*

$$\begin{array}{c} \text{the smallest measurable cardinal} \\ \vee \\ \text{the smallest Ramsey cardinal} \\ \vee \\ \vdots \\ \vee \\ \kappa(\beta) \\ \vee \\ \kappa(\alpha) \\ \vee \\ \text{the smallest inaccessible cardinal} \end{array}$$

23 The Mitchell order

Lemma 197. *Let U be a κ -complete, non-principal normal ultrafilter on κ such that*

$$A = \{\alpha < \kappa \mid \alpha \text{ is a measurable cardinal}\} \in U.$$

For $\alpha \in A$ choose an α -complete, non-principal normal ultrafilter U_α on α . Then

$$\bar{U} = \{X \subseteq \kappa \mid \{\alpha < A \mid X \cap \alpha \in U_\alpha\} \in U\}$$

is a κ -complete, non-principal normal ultrafilter on κ .

Proof. (1) \bar{U} is an ultrafilter on κ .

Proof. $\{\alpha < A \mid \emptyset \cap \alpha \in U_\alpha\} = \emptyset \notin U$ and so $\emptyset \notin \bar{U}$.

$\{\alpha < A \mid \kappa \cap \alpha \in U_\alpha\} = A \in U$ and so $\kappa \in \bar{U}$.

Let $X \in \bar{U}$ and $X \subseteq Y \subseteq \kappa$. Then $\{\alpha < A \mid X \cap \alpha \in U_\alpha\} \in U$. Since $X \cap \alpha \subseteq Y \cap \alpha$

$$\{\alpha < A \mid X \cap \alpha \in U_\alpha\} \subseteq \{\alpha < A \mid Y \cap \alpha \in U_\alpha\}$$

and so $Y \in \bar{U}$.

Let $X, Y \in \bar{U}$. Then $\{\alpha < A \mid X \cap \alpha \in U_\alpha\} \in U$ and $\{\alpha < A \mid Y \cap \alpha \in U_\alpha\} \in U$.

$$\{\alpha < A \mid (X \cap Y) \cap \alpha \in U_\alpha\} = \{\alpha < A \mid X \cap \alpha \in U_\alpha\} \cap \{\alpha < A \mid Y \cap \alpha \in U_\alpha\} \in U.$$

Hence $X \cap Y \in \bar{U}$. *qed*(1)

(2) \bar{U} is non-principal.

Proof. Consider the singleton set $\{\nu\}$ with $\nu < \kappa$. Then

$$\{\alpha < A \mid \{\nu\} \cap \alpha \in U_\alpha\} = \emptyset \notin U$$

and $\{\nu\} \notin \bar{U}$. *qed*(2)

(3) \bar{U} is κ -complete.

Proof. Let $\gamma < \kappa$ and $\{X_i \mid i < \gamma\} \subseteq \bar{U}$. For $i < \gamma$ let $A_i = \{\alpha < A \mid X_i \cap \alpha \in U_\alpha\} \in U$. Let $X = \bigcap_{i < \gamma} X_i$. Let $B = (\bigcap_{i < \gamma} A_i) \setminus (\gamma + 1) \in U$. Then by the closure of the filters U_α under intersections of γ -sequences for $\gamma < \alpha$ we have

$$B \subseteq \left\{ \alpha < A \mid \left(\bigcap_{i < \gamma} X_i \right) \cap \alpha \in U_\alpha \right\} \in U.$$

Thus $\bigcap_{i < \gamma} X_i \in \bar{U}$. *qed*(3)

(4) \bar{U} is normal.

Proof. Let $\{X_i \mid i < \kappa\} \subseteq \bar{U}$. For $i < \kappa$ let $A_i = \{\alpha < A \mid X_i \cap \alpha \in U_\alpha\} \in U$. Let $X = \bigtriangleup_{i < \kappa} X_i$ and $B = A \cap \bigtriangleup_{i < \kappa} A_i \in U$. Let $\alpha \in B$. Then $\forall i < \alpha \alpha \in A_i$ and $\forall i < \alpha X_i \cap \alpha \in U_\alpha$. Since U_α is normal,

$$\left(\bigtriangleup_{i < \kappa} X_i \right) \cap \alpha = \bigtriangleup_{i < \alpha} (X_i \cap \alpha) \in U_\alpha.$$

So

$$B \subseteq \left\{ \alpha < A \mid \left(\bigtriangleup_{i < \kappa} X_i \right) \cap \alpha \in U_\alpha \right\} \in U$$

and $\bigtriangleup_{i < \kappa} X_i \in \bar{U}$. □

In the situation of this lemma, we consider \bar{U} to be weaker than U since U is able to define \bar{U} as a ‘‘limit’’ of ultrafilters of smaller size. This is captured by

Definition 198. Let \bar{U}, U be κ -complete, non-principal normal ultrafilters on κ . We say that \bar{U} is of smaller Mitchell order than U , $\bar{U} \prec_\kappa U$, if there is a sequence $(U_\alpha \mid \alpha \in A)$ of α -complete, non-principal normal ultrafilters U_α on α with $A \in U$ such that

$$\bar{U} = \{X \subseteq \kappa \mid \{\alpha < A \mid X \cap \alpha \in U_\alpha\} \in U\}.$$

We shall show that \prec is well-founded so that one can define a rank on such ultrafilters and on measurable cardinals. We use the $<_U$ -rank of functions in ${}^\kappa\kappa$.

Lemma 199. Let \bar{U}, U be κ -complete, non-principal normal ultrafilters on κ such that $\bar{U} \prec_\kappa U$. Then there is a function $H: {}^\kappa\kappa \rightarrow {}^\kappa\kappa$ such that

$$f <_{\bar{U}} g \text{ implies } H(f) <_U H(g).$$

Proof. Fix a sequence $(U_\alpha \mid \alpha \in A)$ of α -complete, non-principal normal ultrafilters U_α on α with $A \in U$ such that

$$\bar{U} = \{X \subseteq \kappa \mid \{\alpha < A \mid X \cap \alpha \in U_\alpha\} \in U\}.$$

For $f: \kappa \rightarrow \kappa$ define $H(f): \kappa \rightarrow \kappa$ by

$$H(f)(\alpha) = \begin{cases} \|f \upharpoonright \alpha\|_{U_\alpha} & \text{if } \alpha \in A \text{ and } f \upharpoonright \alpha: \alpha \rightarrow \alpha \\ 0 & \text{else} \end{cases}$$

Note that in this definition $\|f \upharpoonright \alpha\|_{U_\alpha} < \text{card}({}^\alpha\alpha)^+ = (2^\alpha)^+ < \kappa$ since κ is strongly inaccessible. Consider functions $f, g: \kappa \rightarrow \kappa$ such that

$$f <_{\bar{U}} g.$$

Then $\{i < \kappa \mid f(i) < g(i)\} \in \bar{U}$. By definition of \bar{U} from U this means

$$B := \{\alpha < A \mid \{i < \alpha \mid f(i) < g(i)\} \in U_\alpha\} \in U.$$

By standard closure arguments

$$C := \{\alpha < \kappa \mid f[\alpha] \subseteq \alpha \text{ and } g[\alpha] \subseteq \alpha\} \in \mathcal{C}_\kappa \subseteq U.$$

For $\alpha \in B \cap C \in U$ we have $f \upharpoonright \alpha <_{U_\alpha} g \upharpoonright \alpha$ and

$$H(f)(\alpha) = \|f \upharpoonright \alpha\|_{U_\alpha} < \|g \upharpoonright \alpha\|_{U_\alpha} = H(g)(\alpha).$$

Thus

$$H(f) <_U H(g). \quad \square$$

Lemma 200. *In the situation of the previous Lemma we have*

- a) for all $f: \kappa \rightarrow \kappa$ $\|f\|_{\bar{U}} \leq \|H(f)\|_U$;
- b) there is a function $h: \kappa \rightarrow \kappa$ such that $H(f) <_U h$ for all $f: \kappa \rightarrow \kappa$;
- c) $\sup \{\|f\|_{\bar{U}} \mid f \in {}^\kappa\kappa\} < \sup \{\|f\|_U \mid f \in {}^\kappa\kappa\}$.

Proof. a) by induction on the well-founded relation $<_{\bar{U}}$. Assume that the claim holds for all $g <_{\bar{U}} f$. Then

$$\begin{aligned} \|H(f)\|_U &= \bigcup \{\|h\|_U + 1 \mid h <_U H(f)\} \\ &\geq \bigcup \{\|H(g)\|_U + 1 \mid g <_{\bar{U}} f\} \\ &\geq \bigcup \{\|g\|_{\bar{U}} + 1 \mid g <_{\bar{U}} f\} \\ &= \|f\|_{\bar{U}} \end{aligned}$$

b) Define $h: \kappa \rightarrow \kappa$ by

$$h(\alpha) = (2^\alpha)^+ = \text{card}({}^\alpha\alpha)^+ > \|f \upharpoonright \alpha\|_{U_\alpha} = H(f)(\alpha).$$

c)

$$\begin{aligned} \sup \{\|f\|_U \mid f \in {}^\kappa\kappa\} &> \|h\|_U \\ &\geq \sup \{\|H(f)\|_U \mid f \in {}^\kappa\kappa\} \\ &\geq \sup \{\|f\|_{\bar{U}} \mid f \in {}^\kappa\kappa\}. \end{aligned}$$

□

Clause c) implies the following

Theorem 201. *The relation \prec_κ is strongly well-founded. The Mitchell order of a κ -complete, non-principal normal ultrafilter U on κ is the rank of U in \prec_κ :*

$$o_\kappa(U) = \bigcup \{o_\kappa(\bar{U}) + 1 \mid \bar{U} \prec_\kappa U\}.$$

The Mitchell order of a cardinal κ is defined as

$$o(\kappa) = \bigcup \{o_\kappa(U) + 1 \mid U \text{ is a } \kappa\text{-complete, non-principal normal ultrafilter on } \kappa\}.$$

A cardinal is measurable iff $o(\kappa) \geq 1$.

The existence of a measurable cardinal of Mitchell order α is a large cardinal axiom. Working with such cardinals requires the assumption that such cardinals exist. We have extended the large cardinal hierarchy beyond measurable cardinals:

$0 = 1$
 \vdots
 the smallest measurable cardinal with Mitchell order ∞ ?
 \vdots
 the smallest measurable cardinal with Mitchell order κ^{++}
 \vdots
 the smallest measurable cardinal with Mitchell order κ
 \vdots
 the smallest measurable cardinal with Mitchell order 2
 \vee
 the smallest measurable cardinal
 \vee
 the smallest Ramsey cardinal
 \vee
 \vdots
 \vee
 $\kappa(\beta)$
 \vee
 $\kappa(\alpha)$
 \vee
 the smallest inaccessible cardinal