Theory of Sets

Aristotelis Panagiotopoulos

Mathematics Department, Caltech, 1200 E. California Blvd, Pasadena, CA 91125

Email address: panagio@caltech.edu

URL: http://www.its.caltech.edu/~panagio/

Contents

1. Introduction

Set theory is the offspring of analysis and logic. It was first developed "naively"[1] in the late 19th century by George Cantor who was motivated by real analysis and the study of sets of uniqueness of Fourier series.

Definition 1. A set $D \subseteq \mathbb{R}$ is a **set of uniqueness**, if whenever the series

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big)$$

converges pointwise to 0 for all $x \in \mathbb{R} \setminus D$, then $a_n = 0 = b_n$ for all $n \in \mathbb{N}$.

Cantor proved that whenever D is "small enough" then it is a set of uniqueness. In the process of making sense of what "small enough" means, he came up with the notion of a "countable set" and he proved that any interval $[a, b] \subseteq \mathbb{R}$ with $a < b$ is not countable. This discovery, that infinite sets come into different sizes, lead Cantor to develop a theory of "cardinality," which may be viewed as the beginning of set theory. While part of the mathematical community viewed Cantor's theory of sets with skepticism it quickly gained popularity. Two factors were:

(1) Using Cantor's theory of cardinality the theorem of Liouville, that there exist real numbers which are not algebraic, is an easy exercise[2].
(2) Many constructions in analysis (such as the construction of the collection of all Borel functions) involve processes naturally indexed by the ordinal ω_1.

The naive treatment of set theory was soon proven to be problematic. Bertrand Russell discovered in 1901 that if we are too liberal in what we allow to be a set then we can provably produce contradictions. For example, in Frege's *Begriffsschrift* (1879) the convention was that given any property φ one may form the set

$$S_\varphi := \{x \mid \varphi(x) \text{ holds}\}, \tag{1}$$

which consists of all sets x that satisfy the property φ. Russell's paradox is the following observation: if φ is taken to be the property $x \notin x$ we get a contradiction:

$$S_\varphi \in S_\varphi \iff S_\varphi \notin S_\varphi,$$

It is therefore necessary to find an axiomatization of set theory which:

(A) is restrictive enough in what qualifies as a set, so that we avoid Russel's type of paradoxes;
(B) is flexible enough in allowing us to perform constructions which we find intuitive enough with respect to how sets "ought to behave".

With respect to (B) above, lets attempt to describe intuitively the **universe V of all sets**. The universe V is built bottom up in stages. At the 0-th stage we have

[1] meaning: non-axiomatically
[2] Indeed, see HW2

the smallest possible set: $V_0 := \emptyset$. Given V_n, the set V_{n+1} is simply the **powerset** $\mathcal{P}(V_n)$ of V_n, i.e., the set of all subsets of V_n. So,

$$V_1 = \mathcal{P}(\emptyset) = \{\emptyset\}, \quad V_2 = \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}, V_3 = \big\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\big\}, \ldots$$

Of course, we have only described a very small fragment of the universe V so far:

$$V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n \subseteq \cdots\cdots\cdots \subseteq V.$$

Every day mathematical objects such as $\omega = \{0, 1, 2, \ldots\}$ "ought" to be in V. Ans similarly collection such as $\{V_0, V_1, V_2, \ldots\}$ should also be V. For one, if ω is in V and each V_n is in V, then one should be able to form the set $\{V_0, V_1, V_2, \ldots\}$ by induction. Setting $V_\omega := \bigcup\{V_0, V_1, V_2, \ldots\}$ to be the union of all the sets we defined so far and letting $V_{\omega+1} = \mathcal{P}(V_\omega)$, $V_{\omega+2} = \mathcal{P}(V_{\omega+1})$,... we can keep going:

$$V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n \subseteq \cdots \subseteq V_\omega \subseteq \cdots \subseteq V_{\omega+n} \subseteq \cdots \subseteq V_{\omega+\omega} \subseteq \cdots\cdots V_\alpha \cdots\cdots \subseteq V.$$

The length of this procedure depends on what constructions we believe we are allowed to perform in order to get new sets from old: if A is a collection of stages which happens to be a set, i.e. $A \in V$, then we can "intuitively" form a new stage $\sup(A)$, that is the supremum of A, and $V_{\sup(A)} := \bigcup_{\alpha \in A} V_\alpha$, so that both are in V. Hence, the length of the hierarchy above depends on how large sets A of stages we can form. This will become more clear soon. Either way, we let $V := \bigcup_\alpha V_\alpha$ to be the universe of all sets.

In order to study the universe V, we will work within first order logic with our language $\mathcal{L}_\in = \{\in\}$ consisting of a binary relation symbol $\in$. In V, the interpretation of $\in$ is the standard belonging relation. Using $\in$ we will be able to define all other necessary set theoretic relations such as: $\bigcup, \bigcap, \subseteq, \ldots$

Besides $(V, \in)$ we will also be considering other $\mathcal{L}_\in$-structures $\mathcal{U} = (U, \in^U)$ which will often also be models of our set-theoretic axioms. In these "non-standard" models, $\in^U$ is just some subset of U^2, a directed graph relation if you want, which satisfies additional axioms (which are perhaps artificial for the theory of directed graphs). Even when U is taken to be a subset of V and $\in^U$ is taken to be the restriction of $\in^V$ on U it is not necessary that the powerset $\mathcal{P}(x)$ of some $x \in U$, when computed in U, is equal to the actual powerset (the one computed in V). Indeed, by the downward Lowenhein-Skolem, if there are models of set theory, then there are countable such models. As a consequence, in such a countable model $(U, \in^U)$, the set $\mathcal{P}(\omega)$ is countable, although within $(U, \in^U)$ we cannot find a bijection between $\mathcal{P}(\omega)$ and ω.

CHAPTER 1

A first course in set theory

1. The Zermelo axioms and the Zermelo universe

The first axiomatization of set theory is due to Zermelo (1908). As we will see, although this axiomatization is strong enough to contain (with the addition of axiom of choice) almost all mathematics, it still does not satisfy (B) above. Hence, in later sections we will need to introduce further axioms.

The Zermelo Axioms.

A.0. Non-triviality. There exists a set: $\exists x(x = x)$.

A.1. Extensionality. Two sets are equal if and only if the contain the same elements: $\forall x \forall y(x = y \iff \forall z(z \in x \iff z \in y))$.

A.2. Pairing. For any two sets there is another set containing exactly these two sets: $\forall x \forall y \exists z \forall w(w \in z \iff (w = x \vee w = y))$.

We introduce here the notation $\{x, y\}$ replacing z. We write $\{x\}$ for $\{x, x\}$.

A.3. Union. Given any set (collection) of sets, there is a set which contains precisely those sets which are members of some set in the collection:

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

The set y above is called the **union** of x and it is denoted by $\cup x$. The set $\cup\{x_1, x_2\}$ will often be denoted by $x_1 \cup x_2$.

A.4. Powerset. Given any sets, there is a set whose elements are precisely all subsets of the original set. If we introduce the notation

$$x \subseteq y \ := \ \forall z(z \in x \implies z \in y),$$

then the powerset axiom is: $\forall x \exists y \forall z(z \in y \iff z \subseteq x)$.

The set y above is called the **powerset** of x and it is denoted by $\mathcal{P}(x)$.

A.5. Subset (scheme). Given any set and any definable with parameters first order property P, there is a set that contains precisely those elements of the original set which satisfy P: for every formula $\varphi(z, x_1, \dots, x_n)$ we have

$$\mathrm{Subset}_\varphi := \quad \forall x, x_1, \dots x_n \exists y \forall z(z \in y \iff (z \in x \wedge \varphi(z, x_1, \dots, x_n))).$$

The set y above will often be denoted by $\{z \in x \mid \varphi(x, x_1, \ldots, x_n)\}$. This axiom is a local version of (1).

A.6. Infinity. There is a set which contains the empty set and it is closed under the powerset operation: $\exists x(\emptyset \in x \wedge \forall y(y \in x \implies \mathrm{P}(y) \in x))$.

The notation $\emptyset$ stands for the **empty set**, which is the set that contains no element. By Lemma 3 such a set exists and it is unique. We will denote the collection A.0.-A.6. of the Zermelo axioms by Z.

REMARK 2. When we informally "define" some collection A and we say that it **exists** (provably from Z) or, equivalently, that it is a **set** (provably from Z), what we mean is that there is a formula $\varphi(x)$ of set theory so that $\mathrm{Z} \models \exists! x\, \varphi(x)$ and when we apply this formula $\varphi(x)$ to the universe V we curve out the collection A. For example, in the introduction, we "defined" V_ω in terms of the "set" $\{V_0, V_1, V_2, \ldots\}$. While we all understand what set the informal expression $\{V_0, V_1, V_2, \ldots\}$ describes, we still need to show that this set can be curved out of V via some formula (which holds for a unique set) in order for it to "exist", i.e., to "be a set". See next lemma for some examples.

LEMMA 3. *The following sets exist (provably from* Z*):*

(1) the set $\emptyset$ which contains no element;
(2) for every natural number n the set V_n described in the intro;
(3) the set $T_\omega := \{V_0, V_1, \ldots, V_n, \ldots\}$;
(4) the set V_ω described in the intro.
(5) for every natural number n, the set $V_{\omega+n}$ described in the intro.

PROOF. For (1) let x be the set given by A.0. and set $\emptyset := \{z \in x \mid z \neq z\}$. Then $\emptyset$ exists by A.5., it contains no element since $\forall z(z = z)$ is an axiom of first order logic, and it is unique by A.1.

(2)Given that $V_0 = \emptyset$ is a set by (1), so is V_n by applying n-times the axiom A.5. Uniqueness is a consequence of A.0 (proved inductively).

For (3), call a set inductive if it contains $\emptyset$ and it is closed under $x \mapsto \mathcal{P}(x)$. By A.6. an inductive sets exists; call it u. Consider the property $\varphi(x)$ defined by

$$\forall z\big((\emptyset \in z \wedge \forall y(y \in z \implies \mathrm{P}(z) \in x)) \implies x \in z\big)$$

Notice now that $T := \{x \in u \mid \varphi(x)\}$ is a set by A.5. and it is the "smallest" inductive set. We leave to the reader to check $T = T_\omega$.

For (4), $V_\omega = \bigcup T_\omega$ exists and is unique by A.3, A.1. (5) is similar to (2). $\square$

Hence the Zermelo axioms are strong enough to produce the beginning of what intuitively should be the V-hierarchy. However, these axioms cannot take us much further. The next exercise is left to the reader.

EXERCISE 4. *The $\mathcal{L}_\in$-structure $(V_\omega, \in)$ is a model of* $\mathrm{Z} \setminus \{A.6.\}$ *(i.e., of the axioms A.0.-A.5). The $\mathcal{L}_\in$-structure $(V_{\omega+\omega}, \in)$ is a model of the Zermelo set theory* Z*.*

COROLLARY 5. *The axioms* $\mathrm{Z} \setminus \{A.6.\}$ *(i.e., A.0.-A.5) cannot prove that* V_ω *is a set. The Zermelo axioms cannot prove that* $V_{\omega+\omega}$ *is a set.*

PROOF. Consider the sentence σ which states that "there exists a set which contains $\emptyset$ and which is closed under the operation $x \mapsto \mathcal{P}(x)$"; see A.6. Assume towards contradiction that $\mathrm{Z}\setminus\{\mathrm{A.6.}\}$ proves that $V_{\omega+\omega}$ exists. Then it can also proves σ. But then, by the previous exercise $(V_\omega, \in) \models \sigma$. This leads to a contradiction since every $x \in (V_\omega, \in)$ is finite and any set that is a witness to σ is infinite.

A similar argument can be used for showing that Z does not prove $V_{\omega+\omega}$ is a set. But here is an alternative argument that is more "universal" and can be used in other situations as well: if Z implies that $V_{\omega+\omega}$ exists, then from the previous exercise it follows that $V_{\omega+\omega} \in V_{\omega+\omega}$. But then we can repeat Russell's Paradox to get a contradiction. More precisely, $S = \{x \in V_{\omega+\omega} \mid x \notin x\}$ is a set by A.5., and therefore it is in $V_{\omega+\omega}$. Following the definitions we have that $S \in S \iff S \notin S$. $\square$

DEFINITION 6. The **Zermelo universe** V_{Z} is simply the structure $(V_{\omega+\omega}, \in)$

From the standpoint of set theory, i.e., from the standpoint of trying to capture V, the Zermelo universe is small since it does not even contain collections such as $\{V_\omega, V_{\omega+1}, \ldots, V_{\omega+n}, \ldots\}$ which intuitively ought to be sets. From the standpoint of classical mathematics however, every commonly used mathematical object can be identified with some set of $V_{\omega+\omega}$. For example, the collection ω of all natural numbers is usual identified with the smallest set that contains $\emptyset$ and is closed under the operation $x \mapsto x \cup \{x\}$. We will prove later that ω exists and it is a subset of V_ω. Hence $\omega \in V_{\omega+1}$. Since an integer k can be identified with an equivalence class $[(a, b)]$ of pairs $(a, b) := \{\{a\}, \{a, b\}\}$ of elements of ω. We can see that $[(a, b)] \in V_{\omega+1}$ and therefore $\mathbb{Z} \in V_{\omega+2}$. Similarly $\mathbb{Q} \subseteq V_{\omega+3}$ and similarly reals r which are identified with the set $\{q \in \mathbb{Q} \mid q < r\}$ are elements of $V_{\omega+3}$. So, $\mathbb{R} \in V_{\omega+4}$. Similarly one can see that $\mathbb{R}^2 \in V_{\omega+4}$ and therefore, since any function $f \colon \mathbb{R} \to \mathbb{R}$ can be identified each graph, $f \in V_{\omega+5}$.

2. Notation and standard constructions

We have already defined $\subseteq, \mathcal{P}(\cdot), \cup$ in terms of $\in$. Given a set X, the **intersection** $\cap X$ of X, also denoted by $\cap\{x \in\mid x \in X\}$ is the set whose elements are contained in every set x which is a member of X. By A.3. and A.5. we have that $\cap X$ exists:

$$\cap X := \{z \in \cup X \mid \forall x(x \in X \implies z \in x)\}.$$

We denote $\cap\{x, y\}$ by $x \cap y$. Similarly we define

$$x \setminus y := \{z \in x \mid z \notin y\} \quad \text{and} \quad x \triangle y := (x \setminus y) \cup (y \setminus x).$$

Given sets x, y, the **pair** (x, y) **of** x, y is the set $\{\{x\}, \{x, y\}\}$. We also let $(x, y, z) := (x, (y, z))$, etc. The fundamental property of the pair is the following.

PROPOSITION 7. $(a, b) = (c, d)$ *if and only if* $a = c$ *and* $b = d$.

PROOF. Easy and left as an exercise. □

The **Cartesian product** $X \times Y$ of the sets X, Y is the set $\{(x, y) \mid x \in X, y \in Y\}$, which exists by applying the A.5., with $\varphi(z) \equiv \exists x \exists y \big(z = (x, y) \wedge x \in X \wedge y \in Y\big)$, to the set $\mathcal{P}(\mathcal{P}(X \cup Y))$. We set $X^2 := X \times X$, $X^3 := X \times (X \times X)$, etc.

By a **relation** is any set R which consists entirely of ordered pairs. We write xRy, whenever $(x, y) \in R$. The **domain** and the **range**, respectively, the sets

$$\operatorname{dom}(R) = \{x \mid \exists y (x, y) \in R\} \quad \text{and} \quad \operatorname{rng}(R) = \{y \mid \exists x (x, y) \in R\}.$$

EXERCISE 8. *Show that* $\operatorname{dom}(R)$ *and* $\operatorname{rng}(R)$ *are sets. Similarly for any other collection in this section which we have left without any justification.*

If $R \subseteq X \times Y$ then we say that R is a **relation from** X **to** Y.

An **equivalence relation** is any $R \subseteq X \times X$ which is **reflexive** (xRx), **symmetric** $(xRy \implies yRx)$, and **transitive** $(xRy \wedge yRz \implies xRz)$. We denote by $[x]_R := \{y \in X \mid xRy\}$ the R-equivalence class of x and set $X/R := \{[x]_R \mid x \in X\}$.

A **function** is any relation which satisfies $xfy \wedge xfz \implies y = z$. We denote xfy by $f(x) = y$. If $\operatorname{dom}(f) = X$ and $\operatorname{rng}(f) = Y$ then we write $f\colon X \to Y$, The function f is **onto** if $\operatorname{rng}(f) = Y$ and it is **injective** if $f(x) = f(x') \implies x = x'$. It is **bijective** if it is both injective and onto. If $Z \subseteq X$, we denote by $f \restriction Z$ the restriction $f \cap (Z \times Y)$ of f on Z. We let $f''A$ and $f^{-1}(B)$ be the front and inverse image of the sets $A \subseteq X$ and $B \subseteq Y$ under f. The **set of all functions** $f\colon X \to Y$ from X to Y is denoted by Y^X.

DEFINITION 9. An **indexed family of sets** $\{X_i\}_{i \in I}$ is any function f whose domain is the set I and $f(i) = X_i$. The **Cartesian product of** $\{X_i\}_{i \in I}$, denoted by

$$\prod_{i \in I} X_i,$$

is the set of all functions g with domain I so that $g(i) \in X_i$. Notice that if $X_i = X$ for all i then this is simply the set X^I.

Notice that if $X_i = \emptyset$ for some $i \in I$ then $\prod_{i \in I} X_i = \emptyset$. If $X_i \neq \emptyset$ for all $i \in I$ then, intuitively, $\prod_{i \in I} X_i \neq \emptyset$. That is, we "should" always be able to produce a function that chooses an element x_i from each X_i. However, this is not provable from the Zermelo axioms and, as we will later see, it is one of the equivalent forms of the axiom of choice.

3. ω and definition by induction

By a **number system** we mean any triple (A, a_0, S) which satisfies the Peano axioms (second order). In other words:

(1) $a_0 \in A$;
(2) $S\colon A \to A$;
(3) $S(a) \neq a_0$ for all $a \in A$;

(4) S is injective;

(5) if $B \subseteq A$ with $a_0 \in B$ and with ($b \in B \implies S(b) \in B$), then $B = A$.

Intuitively, and therefore in V, a number system looks like $\{a_0, S(s_0), S(S(a_0)), \ldots\}$. Notice however, that if $\mathcal{U} = (U, \in^{\mathcal{U}})$ is a model of Z, and $A, a_0, S \in U$, then the above axioms can be phrased in **first order logic** of the language $\{\in\}$ and therefore, $\mathcal{U} = (U, \in^{\mathcal{U}})$ has its "own idea" of what a number system is. By that, we mean that if, for example, U does not contain the "true" powerset $\mathcal{P}(A) := \mathcal{P}^V(A)$ of A, but rather $\mathcal{P}^{\mathcal{U}}(A) \subsetneq \mathcal{P}^V(A)$, then the induction axiom σ_5 above may not be true in V, yet it is possible that $\mathcal{U} \models \sigma_5$. See Exercise 13 for more information.

The following theorem is the first in a series of theorems which will allow to argue that several collections exist as sets and consequentially better approximate the hierarchy $(V_\alpha)_\alpha$ discussed in the introduction.

THEOREM 10 (Definition by induction I). *Let (A, a_0, S) be a number system, let $g\colon C \times A \to C$ a function, and $c_0 \in C$. There exists a unique function $f\colon A \to C$ so that:*

$$f(a_0) = c_0 \quad \textit{and} \quad f(S(a)) = g(f(a), a). \tag{2}$$

PROOF. The uniqueness is immediate from the induction property of (A, a_0, S): If f_1, f_2 are two functions satisfying (2) then let $B = \{a \in A \mid f_1(a) = f_2(a)\}$ and notice that (2) implies the assumptions of the induction axiom. Hence $A = B$.

To prove existence we need the following notion of approximation of f. A set $B \subseteq A$ is an **initial segment** of A if $a_0 \in B$ and $S(a) \in B \implies a \in B$. An **approximation to** f is a map $u\colon B \to C$, where B is an initial segment of f, $u(a_0) = c_0$, and for every $a \in A$ with $S(a) \in B = \mathrm{dom}(u)$, we have that $u(S(a)) = g(u(a), a)$. Notice that the collection

$$F := \{u \mid u \text{ is an approximation to } f\},$$

is a set by applying the appropriate formula to the set $\mathcal{P}(A \times C)$. It is also non-empty, since $\{(a_0, c_0)\} \in F$. Let $f := \bigcup F$. We claim that f is a function and it satisfies (2) above.

Clearly f is a relation since it is a union the relations $u \in F$. By the induction axiom we also see that $\mathrm{dom}(f) = A$, since $\{(a_0, c_0)\} \in F \implies = a_0 \in \mathrm{dom}(f)$, and whenever $a \in \mathrm{dom}(u)$, with $u \in F$ we have that $u' = u \cup \{(S(a), g(u(a), a))\} \in F$.

To see that f is a function, let $B = \{a \in A \mid f \text{ is uniquely defined}\}$. That is,

$$B = \{a \in A \mid \exists c \in C \forall u \in F (u(a) = c)\}.$$

Now a_0 is clearly in B and if for some fixed a there is a fixed c so that all $u \in F$ agree that $u(a) = c$, the definition of approximation implies that $S(a) \in B$. By induction axiom we have that $B = A$ and therefore f is a function.

We leave to the reader to similarly check that (2) holds for f. □

COROLLARY 11. *Any two numbers systems (A, a_0, S) and (B, b_0, T) are isomorphic, i.e., there is a bijection $\pi\colon A \to B$, s.t. $\pi(a_0) = b_0$, $\pi(S(a)) = T(\pi(a))$ for all $a \in A$.*

Proof. Apply Theorem 2 with $c_0 := b_0$ and $g(a,c) := T(c)$ to get π. Use induction axiom (axiom 5 above) to show that π is an injection, a surjection, and that it preserves the number system structure. $\square$

Remark 12. Notice that the above corollary shows that second order logic behaves very differently than first order logic. Indeed, by the upward Lowenheim-Skolem the usual first order formulation of Peano arithmetic has many pairwise non-isomorphic models, while the above corollary says that all models of second order PA are isomorphic. In other words, the upward Lowenheim-Skolem is not true in second order logic. How can we reconcile this observation with the discussion following the definition of a number system? Well, that is the content of the next exercise.

Exercise 13. *Let $\mathcal{U} = (U, \in^{\mathcal{U}})$ and $\mathcal{W} = (W, \in^{\mathcal{W}})$ be models of* Z. *Prove that:*

1. *If $(A, a_0, S) \in U$ and $\mathcal{U} \models$ "(A, a_0, S) is a number system″ then (A, a_0, S) is a model of* **first order Peano arithmetic** *in $\mathcal{L} = \{0, s\}$;*
2. *If $(A, a_0, S), (B, b_0, T) \in U$, $\mathcal{U} \models$ "(A, a_0, S) is a number system″, and $\mathcal{U} \models$ "(B, b_0, T) is a number system″ then (A, a_0, S) and (B, b_0, T) are isomorphic $\{0, s\}$-structures;*
3. *Assume that $(A, a_0, S) \in U$, $(B, b_0, T) \in W$, $\mathcal{U} \models$ "(A, a_0, S) is a number system″, and $\mathcal{W} \models$ "(B, b_0, T) is a number system″. Does the above theorem imply that (A, a_0, S) and (B, b_0, T) are isomorphic $\{0, s\}$-structures?*

Next we describe a specific number system of V, which we will identify with the set ω of all "natural numbers", and we show that the existence of this number system is guaranteed by the axioms we have so far. By the corollary which follows Theorem 2 any other number system in V is isomorphic to ω.

Let x be a set. The **successor** x^+ **of** x is the set $x \cup \{x\}$. In V we can intuitively form the set $\{\emptyset, \emptyset^+, \emptyset^{++}, \ldots\}$. We take this to be the **set** ω **of all natural numbers**. The formal definition is as follows: a **successor set** z is any set which contains $\emptyset$ and which is closed under the operations $x \mapsto x^+$, that is, any set z which satisfies:

$$\varphi_{\text{succ}}(z) := (\emptyset \in z) \wedge (\forall x\, x \in z \implies x^+ \in z)$$

We define ω to be the "smallest successor set", that is the unique set satisfying:

$$\varphi_\omega(w) := \forall x\big((x \in w) \iff (\forall z\, \varphi_{\text{succ}}(z) \implies x \in z)\big)$$

Theorem 14. Z *proves that there exists a unique set satisfying φ_ω*

Proof. Uniqueness of ω follows immediately from the extensionality axiom A.5. since any two sets w_1, w_2 which satisfy the formula $\varphi_\omega(w)$ are forced by this formula to the exact same x, that is, all x with $(\forall z\, \varphi_{\text{succ}}(z) \implies x \in z)$.

For existence, it suffice to show that there exists some successor set S because then we can use the subset axiom A.5., to attain ω as $\{w \in \mathcal{P}(S) \mid \varphi_\omega(w)\}$. This follows from the next lemma. $\square$

Lemma 15. *V_ω, which exists by Lemma 3, is a successor set.*

Before we proceed to the proof of this lemma we introduce the following notion that is going to play important role, not only in the proof of this lemma, but also for the development of the general theory of the class ORD of all "ordinals" which we will define later on.

Definition 16. A set X is called **transitive** if whenever $v \in x$ and $x \in X$, we have that $v \in X$. Equivalently, whenever $x \in X \implies x \subseteq X$.

Proof of 15. By definition, $V_\omega := \bigcup T_\omega$ where T_ω is the "smallest set" that contains $\emptyset$ and it is closed under the operation $x \mapsto \mathcal{P}(x)$. Since $\emptyset \in T_\omega$, we have that $\{\emptyset\} = \mathcal{P}(\emptyset) \in T_\omega$ and therefore $\emptyset \in \bigcup T_\omega$.

To see that $\bigcup T_\omega$ is closed under $x \mapsto x^+$, notice first that every element $X \in T_\omega$ is transitive. This because

$$T := \{X \in T_\omega \mid X \text{ is transitive}\}$$

contains $\emptyset$ and it is closed under $X \mapsto \mathcal{P}(X)$ (if $x \in \mathcal{P}(X)$ then $x \subseteq X$; so if $v \in x$ then $v \in X$, and by transitivity of X, $v \subseteq X$; hence $v \in \mathcal{P}(X)$) and by minimality of T_ω we have $T_\omega = T$.

Let now $x \in \bigcup T_\omega$. Then, there is $X \in T_\omega$ with $x \in X$. By transitivity of X we have that $x \subseteq X$ and therefore $x \in \mathcal{P}(X)$. So $x \in X$ and $x \subseteq X$. In other words, $x^+ := x \cup \{x\} \subseteq X$. That is $x^+ \in \mathcal{P}(X)$, and since $\mathcal{P}(X) \in T_\omega$ we have that $x^+ \in \bigcup T_\omega$. □

Theorem 17. *The triple $(\omega, \emptyset, s)$, where $s(x) = x^+$, is a number system.*

Proof. (1), (2) are immediate. For (3) notice that $x^+ \neq \emptyset$ since $x \in x^+$. For (4), notice first that each $x \in \omega$ is transitive. Indeed, let $\Omega := \{x \in \omega \mid x \text{ is transitive}\}$ and notice that $\emptyset \in \Omega$ and Ω is closed under $x \mapsto x^+$: if $v \in y \in x^+$, then either $v = x$ or $v \in y \in x$; the first immediately implies $v \in x^+$ and so does the second after invoking transitivity of x. By minimality of omega we have that $\omega = \Omega$.

If now $x^+ = y^+$ with $x, y \in \omega$ the $x \cup \{x\} = y \cup \{y\}$. It follows that

$$(x \in y \text{ or } x = y) \text{ and } (y \in x \text{ or } x = y)$$

In the worst case scenario we have that $x \in y$ and $y \in x$. But then, by transitivity we have that $x \subseteq y$ and $y \subseteq x$. Hence, by extensionality A.2. we have $x = y$.

Property (5) is immediate by minimality of ω. □

We will usually denote elements of ω by $k, l, m, n, \ldots$ and we set

$$0 := \emptyset, \quad 1 := \emptyset^+, \quad 2 := \emptyset^{++}, \quad 3 := \emptyset^{+++}, \ldots$$

Notice that for every $n \in \omega$ we have that $n^+ = \{0, 1, \ldots, n\}$. Theorem 10 easily generalizes (with the same proof) to the next theorem:

Theorem 18 (Definition by induction II). *Let $h \colon X \to Y$ and $g : Y \times \omega \times X \to Y$ be to functions. Then there exists a unique $f \colon X \times \omega \to Y$ so that*

$$f(0, x) = h(x) \quad \text{and} \quad f\big(n^+, x\big) = g(f(n, x), n, x\big).$$

Using this we can define **addition**, **multiplication**, and **exponentiation** by:

$$f_+(0,m) = m \quad \text{and} \quad f_+\big(n^+,m\big) = f(n,x)^+;$$
$$f_*(0,m) = 0 \quad \text{and} \quad f_*\big(n^+,m\big) = f_+\big(f_*(n,m),m\big);$$
$$f_{\exp}(0,m) = 1 \quad \text{and} \quad f_{\exp}\big(n^+,m\big) = f_*\big(f_{\exp}(n,m),m\big);$$

We will denote $f_+(n,m)$, $f_*(n,m)$, $f_{\exp}(n,m)$ simply by $n+m$, $n*m$, m^n. Finally

$$\text{for} \ \ n,m \in \omega \ \ \text{we let} \ \ (n < m \iff n \in m)$$

Theorem 19. *The relation $<$ is a linear ordering on ω*

Proof. We need to show that $<$ is a transitive, non-reflexive, total relation. We already showed that every $l \in \omega$ is transitive: if $n \in m \in l$ then $n \in l$.

We show that $n \not< n$ by induction on n. It is clear that $\emptyset \notin \emptyset$. Moreover notice that if $n \cup \{n\} \in n \cup \{n\}$ then either $n \cup \{n\} \in n$ or $n \cup \{n\} = n$. In both cases it would follow $n \in n$ which contradicts the inductive assumption that $n \not< n$.

We finally need to show that $<$ is total (since $n < m \implies n$ follows from that and non-reflexivity). We will show by induction on n that the following holds:

$$T(n) \iff \forall m\big((n \in m) \vee (n = m) \vee (m \in n)\big)$$

For $n = 0$ an easy induction shows that $\forall m\big((0 \in m) \vee (0 = m)\big)$. We leave this to the reader. Assume now that $T(n)$ holds and we prove $T(n^+)$. We do this by a second induction on m. For $m = 0$ it is easy to see that $m \in n^+$ since $0 \in \{0\} = 0^+$ and $0^+ \subseteq 1^+ \subseteq \cdots \subseteq n^+$. Assume now that

$$\big((n^+ \in m) \vee (n^+ = m) \vee (m \in n^+)\big),$$

we need to show that

$$\big((n^+ \in m^+) \vee (n^+ = m^+) \vee (m^+ \in n^+)\big).$$

If $n^+ \in m$ then $n^+ \in m^+$ by transitivity and the fact that $m \in m^+$. If $n^+ = m$ then $n^+ \in m^+$. Finally, if $m \in n^+$, we have two cases to consider. First case is $m = n$, in which case $m^+ = n^+$ and we are done. Second case is $m \in n$. We also have:

$$\big((n \in m^+) \vee (n = m^+) \vee (m^+ \in n)\big),$$

since, by inductive assumption, $T(n)$ holds. But $m \in n$ contradicts the first alternative since $(n \in m^+) \implies (n \in m) \vee n = m$, and $T(n)$ holds by assumption. The other two cases both imply that $m^+ \in n^+$ and we are done. □

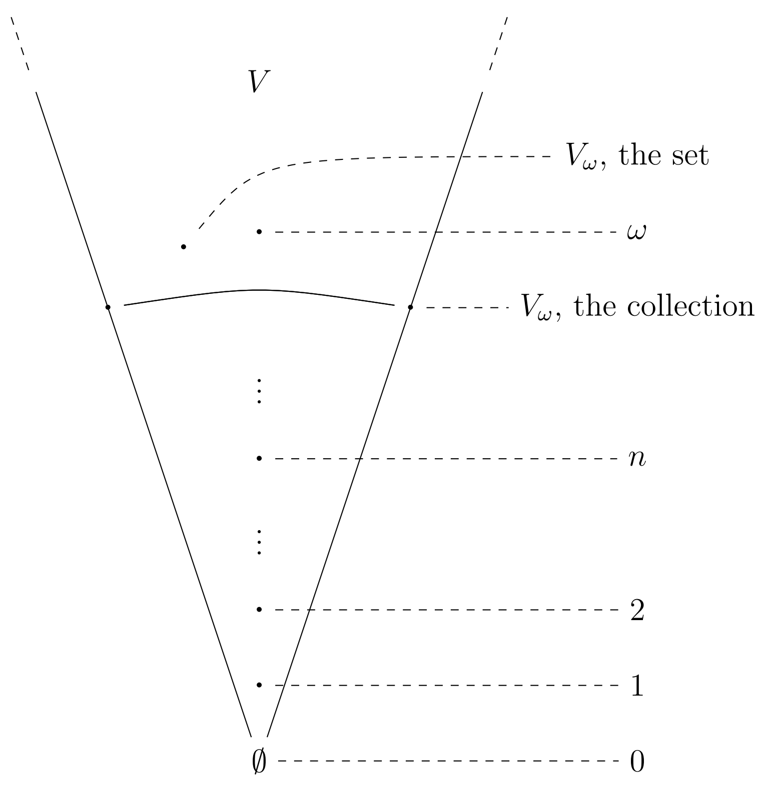
Theorem 20. *The relation $<$ is a well ordering on ω. That is, for every $A \subseteq \omega$, if $\emptyset \neq A$, there is $a \in A$ so that for all $b \in A$ we have $a \leq b$. Moreover, for all $n \in \omega$ we have that there is no $k \in \omega$ with $n < k < n+1$.*

Proof. For the second assertion, if $n < k < n+1$, then $n \in k \in n \cup n \cup \{n\}$. So either $n < n$ by transitivity $n \in k \in n$, a contradiction; or $n < n$ by since $n \in k = n$, again a contradiction.

For the first statement, let $B = \{n \mid \forall k \leq nk \notin A\}$. Assume towards contradiction that A has no least element then we will show $B = \emptyset$. This would imply that $A = \emptyset$. Clearly $0 \in B$ since otherwise, $0 \in A$ and it would clearly be the least element of A. If $n \in B$, i.e., $\forall k \leq nk \notin A$ we want to show $n + 1 \in B$, that is, $\forall k \leq n^+ k \notin A$. But if that fails than $n + 1 \in A$ is the least element of A. $\square$

To summarize, ω was chosen to be a "canonical" number system inside V: the well founded linear ordering $<$ that is associated with the system $(\omega, 0, S)$ is no other than $\in$ itself. The notion of a transitive set (Definition 16) was used extensively in the process of establishing the relationship between the various elements of ω. It is not difficult to see that not only the elements n of ω are transitive sets but also ω itself. Insisting on viewing $\in$ are an ordering $<$, we have

$$0 < 1 < 2 < \ldots < n < \ldots < \omega.$$



Of course, we can now start applying the operation $x \mapsto x^+$ to continue this ordering to $\omega + 1 := \omega^+$, $\omega + 2 := \omega^{++}$. Every set of the form $\omega + n$ is easily seen to be transitive and well-ordered by $\in$. These sets form the initial segment of the class ORD of all ordinals. Ordinals is the formal name for what we referred to as "stages" in the introduction. The class ORD will form for us the "spine" of the universe V

which we will use, among others, to formally define the hierarchy $(V_\alpha)_{\alpha\in\mathrm{ORD}}$ that exhausts the universe V.

Definition 21. An **ordinal** (in the sense of Von Neumann) is any set α which is transitive and well-ordered by $\in$, that is, $\{(x,y) \mid x,y\in\alpha,\ x\in y\}$ is a well-ordering.

4. Cardinality: finite vs infinite

Two sets x,y are **equinumerable**, or **have the same cardinality**, if there is a map $f\colon x\to y$ which is bijective. We denote this by

$$x\approx y.$$

It is clear that for all x,y,z we have that $x\approx x$, $x\approx y\implies y\approx x$ and $x\approx y\wedge y\approx z\implies x\approx z$. In other words, if it was not for the fact that $\approx$ is not a set, then $\approx$ would be an equivalence relation. A set x is **finite** if there is $n\in\omega$ with $x\approx n$. It is **infinite** if it is not finite. Being finite is a "smallness" property:

Proposition 22. *The collection of all finite sets forms an ideal. That is:*
(1) $\emptyset$ is finite;
(2) if x is finite and $y\subseteq x$, then y is finite;
(3) if x,y are finite then $x\cup y$ is finite.

Proof. For (1), notice that $\emptyset\colon\emptyset\to\emptyset$ is a bijection. For (3), notice that if $f\colon n\to x$ and $g\colon m\to y$ are bijections, then $h\colon n+m\to x\cup y$ with $h(k)=f(k)$, if $k<n$ and $h(l+n)=g(l)$ if $l<m$; is a bijection. Property (2) follows from the lemma:

Lemma 23. *If $x\subsetneq n$, then $x\approx m$ for some $m<n$.*

Proof of Lemma. By induction on n. For $n=0$ it is clear. Assume now that it holds for n and assume that $x\subsetneq n^+=n\cup\{n\}$. If $n\notin x$ then either $x=n$, in which case we are done since $x\approx m:=n<n^+$; or $x\subseteq n$, in which case we are done by inductive hypothesis and transitivity of $<$. So assume that $n\in x$. But then let $i\in n$ with $i\not\in x$ and consider the function $\{(k,k)\mid k\in x\setminus n\}\cup\{(n,i)\}$. This function shows that $x\approx x'$ where $x'\subseteq n$. Then either $x'=n$ and we are done since $m:=n<n^+$; or $x'\subsetneq$ and we are done by inductive hypothesis. □

□

Taking the contrapositive we have that "being infinite" is a "largeness" property.

Corollary 24. *The collection of all infinite sets forms a filter. In particular, if $y\subseteq x$ and y is infinite, then so is x.*

Theorem 25 (Non-compressibility). *If x is finite and $y\subsetneq x$ then $x\not\approx y$.*

Proof. By next lemma it suffices to show that if $m<n\in\omega$ then $m\not\approx n$. We need to show that the set $A:=\{n\in\omega\mid\ \forall m<n\ m\not\approx n\}$ is equal to ω. But 0 is clearly in ω and assuming $n\in A$ we can easily show that $n^+\in A$: if $n^+\approx m<n^+$

then $n \approx x \subsetneq m$. By the above lemma we have $n \approx x \approx k < m$. Since $m < n^+$, by the second part of Theorem 20 we have that $k < n$. This, together with $n \approx k$, contradicts $n \in A$. $\square$

Corollary 26. *ω is infinite.*

Proof. Notice that $n \mapsto n + 1$ is a "compression" of ω. $\square$

The fact that ω is "compressible" while no finite set is "compressible" suggest that the following alternative definition of "infinite." A set x is **Dedekind infinite**, or **compressible**, if for some set y we have $x \approx y \subsetneq x$. By Theorem 25, we have that

$$\text{Dedekind infinite} \implies \text{Infinite}.$$

However, the converse is independent of the Zermelo axioms. That is a good excuse to introduce a new axiom. Let x be a set. A **choice function** for $\mathcal{P}(x)$ is any function $\varphi\colon \mathcal{P}(x) \setminus \{\emptyset\} \to x$ so that $\varphi(y) \in y$ for all non-empty $y \subseteq x$.

The Axiom of Choice.

AC. If x is non-empty then there exists a choice function for $\mathcal{P}(x)$.

We should emphasize that for many sets x which come together with some reasonable structure, such as ω, one can directly define a choice function for $\mathcal{P}(x)$. For example, Theorem 20 proves that the assignment $A \mapsto \min_<(A)$ is a choice function for $\mathcal{P}(\omega)$

Theorem 27. *Assuming the axiom of choice: Infinite $\implies$ Dedekind infinite.*

Proof. Let x be infinite. It suffices to find an injective map $f\colon \omega \to x$ because then we can define a compression $g\colon x \to x$ by setting $g(y) = y$ if $y \notin \mathrm{rng}(f)$ and $g(y) = f(n+1)$ if $f(n) = y$.

To define f, fix any choice function $\varphi\colon \mathcal{P}(x) \setminus \{\emptyset\} \to x$ for $\mathcal{P}(x)$ and set:

$$f(0) = \varphi(x) \quad \text{ and } \quad f(n^+) = \varphi(x \setminus \{f(0), \dots, f(n)\}).$$

Formally, this is defined using Theorem 10: first define by induction a map $F\colon \omega \to \mathcal{P}(x)$ using, in the notation of Theorem 10, $c_0 = x \in \mathcal{P}(x)$ and $g\colon \mathcal{P}(x) \times \omega \to \mathcal{P}(x)$ with $g(c, n) = c \setminus \varphi(c)$. Then let $f(n) = \varphi(F(n))$. $\square$

For every finite set x we say that the **cardinality of** x **is** n if $x \approx n$. We denote this by $|x| = n$. Notice by the results of this section such a number is unique for each x. Here are some basic properties of cardinalities of any two finite sets x, y:

(1) for all $n \in \omega$ we have that $|n| = n$;
(2) $|x \times y| = |x| * |y|$;
(3) $|x^y| = |x|^{|y|}$;
(4) $|\mathcal{P}(x)| = 2^{|x|}$;
(5) $x \cap y = \emptyset \implies |x \cup y| = |x| + |y|$.

We say that the **cardinality of a set x is less than or equal to the cardinality of** y, or $x \lessapprox y$, if there is an injection $f: x \to y$, that is, if $x \approx z \subseteq y$. We have that:

(1) under (AC), x is infinite if and only if $\omega \lessapprox x$;
(2) if x, y are finite then $x \lessapprox y$ if and only if $|x| \leq |y|$;
(3) if x is finite and $f: x \to y$ then $f(x)$ is finite and $|f(x)| \leq |x|$.

5. Compactness and combinatorics on the boundaries of Peano Arithmetic

When it comes to infinite sets, the following trivial fact is the starting point for many, often non-trivial, combinatorics.

Pigeonhole Principle. Assume that $X_0, \ldots, X_n$ is a covering of an infinite set X, i.e., $X_0 \bigcup \cdots \bigcup \cdots X_n = X$ and $X_i \bigcap X_j = \emptyset$ for all distinct $i, j \leq n$. Then there is $i \leq n$ with X_i infinite.

To see this, notice that it suffices to prove that if all X_i are finite then so is the union of them. This is proved by induction of n using Proposition 22. The first application of the Pigeonhole Principle will be a combinatorial compactness principle, known as König's Lemma, which often allows us to bridge finite combinatorics with infinite combinatorics.

Let A be a set. A **tree on** (alphabet) A is any set T of finite sequences $(a_0, \ldots, a_{n-1})$ from A closed under initial segments, that is:

$$(a_0, \ldots, a_{n-1}) \implies \forall m \leq n (a_0, \ldots, a_{m-1}) \in T$$

Formally, a finite sequence is any function $s: n \to A$ with $n \in \omega$. For $n = 0$ we get the empty sequence $\emptyset$, which is obviously contained in any non-empty tree. A **node** of T is simply an element of T.

Examples. $\omega^{<\omega}$, $2^{<\omega}$; draw pictures.

By an **infinite branch** of T we mean any function $f: \omega \to A$ so that $f \restriction n \in T$ for all $n \in \omega$. Notice that an infinite branch of T is not an element of T but it is "approximable" by elements of T. A tree is **finite splitting** if for every $s = (a_0, \ldots, a_{n-1}) \in T$, the collection of all $a \in A$ so that $s^\frown a := (a_0, \ldots, a_{n-1}, a) \in T$ is finite. If $s^\frown a \in T$ we call $s^\frown a$ and **immediate extension of** s.

Theorem 28 (König's Lemma (AC)). *Every finite splitting infinite tree has an infinite branch.*

Proof. Let T be a finite splitting tree on some alphabet A. Fix a choice function $\varphi: \mathcal{P}(A) \to A$. We define inductively, using Theorem 10, the element $a_n \in A$ by

$$a_0 := \varphi(\{a \in A \mid (a) \in T \text{ and } T_{(a)} \text{ is infinite}\})$$

$$a_{n+1} := \varphi(\{a \in A \mid (a_0, \ldots, a_n, a) \in T \text{ and } T_{(a_0, \ldots, a_n, a)} \text{ is infinite}\}),$$

where for any sequence $s = (b_0, \ldots, b_{k-1})$ we let

$$T_s := \{t \in T \mid t = (b_0, \ldots, b_{k-1}, b_k, \ldots, b_{l-1}), \text{ for some } b_k, \ldots, b_{l-1} \in A\}$$

The fact that each a_n exists follows from the pigeonhole principle applied on the assumption that T is finitely branching. It is immediate to prove by induction that $(a_0, \ldots, a_{n-1}) \in T$ for all n, and therefore $(a_n)_{n\in\omega}$ is an infinite branch. □

Example. For an example of an infinite tree without infinite branch consider the subtree of $\omega^{<\omega}$ consisting of all sequences $(n_0, \ldots, n_{k-1})$ with $i \leq n_i$.

König's Lemma is often used to connect the infinitary with finitary results. We will will illustrate this by reflecting the infinite Ramsey theorem to its finitary counterpart. Lets start with the infinitary Ramsey theorem which we will simply call the "Ramsey theorem." The Ramsey theorem is a higher-dimensional analogue of the Pigeonhole Principle. The *2-dimensional Pigeonhole Principle* known as the *Ramsey theorem for edges* can be stated as follows. Given a set X, let

$$[X]^2 = \big\{\{a, b\} \mid a, b \in X, a \neq b\big\}.$$

One may think of the pair $(X, [X]^2)$ as the complete graph on domain X. By a **finite coloring** of $[X]^2$ we mean a partition $[X]^2 = C_1 \sqcup \cdots \sqcup C_m$ of $[X]^2$ into finitely many sets $C_1, \ldots, C_m$ which we call **colors**. The Ramsey theorem for edges states that if we finitely color the edges of any infinite complete graph then one of the colors is "large", in that, it contains the edges of an infinite complete subgraph.

THEOREM 29 (Infinite Ramsey for edges (AC)). *Let X be an infinite set. If $[X]^2 = C_1 \sqcup \cdots \sqcup C_m$ is a finite coloring of $[X]^2$ then there is some infinite $Y \subseteq X$ and some $i \leq m$, with $[Y]^2 \subseteq C_i$.*

PROOF. Notice that it suffices to prove this for colorings $[X]^2 = B \sqcup R$ consisting of 2 colors. Indeed by a simple induction we can reduce the general case which uses m colors to the one which uses $m - 1$ many colors by setting $B = C_1 \sqcup \cdots \sqcup C_{m-1}$ and $R := C_m$. Moreover using AC we can assume without loss of generality that $X = \omega$. We will use the notation $\{n < n'\}$ to denote the set $\{n, n'\} \in [\omega]^2$ where $n < n'$.

We will first construct an infinite subset $X_\infty \subseteq \omega$ which is not necessarily "monochromatic" but the colors of the edges are nicely organized: the color of the edge $\{n, n'\}$ with $n, n' \in X_\infty$ depends only on the minimum element n, i.e. if $\{n, n''\}$ is also an edge in X_∞ then then the color of $\{n, n'\}$ is the same as the color of $\{n, n''\}$.

Let $n_0 = 0$. By Pigeonhole Principle there is an infinite subset $X_0 \subseteq \omega$ and a color $C \in \{B, R\}$ so that for all $n \in X_0$ we have $\{n_0 < n\} \in C$.

Assume that n_k and X_k have been defined so that X_k is infinite. Let $n_{k+1} := \min X_k$. By the Pigeonhole Principle we have that there is an infinite subset $X_{k+1} \subseteq X_k$ and a color $C \in \{B, R\}$ so that for all $n \in X_{k+1}$ we have $\{n_{k+1} < n\} \in C$.

Let $X_\infty := \{n_0 < n_1 < n_2 < \ldots < n_k < \ldots\}$ and notice that indeed, if $\{n_i < n_j\} \in [X_\infty]^2$ then n_j is an element of X_i and therefore the color of $\{n_i, n_j\}$

depends only on n_i and it is specified at the i-th stage of the above induction. This way we define a coloring $X_\infty = B' \sqcup R'$ of the **vertexes** of X_∞:

$$n \in B' \iff \forall n' > n \; \{n < n'\} \in B \iff \exists n' > n \; \{n < n'\} \in B$$

By a final use of the Pigeonhole Principle we find an infinite $Y \subseteq X_\infty$ so that $Y \subseteq B'$ or $Y \subseteq R'$. This implies that wither $[Y]^2 \subseteq B$ or $[Y]^2 \subseteq R$ as desired. $\square$

In the Pigeonhole principle we colored vertexes and in the Theorem above we colored edges. More generally, for any dimension $d \in \{1, 2, 3, \ldots\}$ let

$$[X]^d := \big\{\{x_1, \ldots, x_d\} \mid x_i \in X, x_i \neq x_j\big\},$$

be the set of all d-dimensional faces with vertexes from X. We have:

THEOREM 30 (Infinite Ramsey (AC)). *Let X be an infinite set. If $[X]^l = C_1 \sqcup \cdots \sqcup C_m$ is a finite coloring of $[X]^l$ then there is some infinite $Y \subseteq X$ and some $i \leq m$, with $[Y]^l \subseteq C_i$.*

PROOF. The proof is exactly the same as in Theorem 29. The only difference is that the green uses of the Pigeonhole Principle are replaced with the Ramsey theorem of dimension $(d-1)$, so that the color of $\{n_k\} \cup \sigma$, for any $\sigma \in [X_k]$, is entirely determined by n_k. $\square$

The finite version of the general Ramsey theorem can be stated as follows. First let $a, b, c \in \omega$ which, intuitively, will satisfy $a \leq b \leq c$ and let $m\omega$ be the number of colors. We will use the notation

$$c \to (b)^a_k$$

whenever these numbers satisfy: "if the $[c]^a$ is colored with k-many colors then there is a subset $Y \subseteq a$ of cardinality b so that $[Y]^a$ is monochromatic."

THEOREM 31 (Ramsey Theorem (finite)). *For all $a, b, c, m \in \omega$ there is $c \in \omega$ with*

$$c \to (b)^a_k.$$

PROOF. Assume that the statement is false. That is, for every c there is a "bad" m-coloring of $[c]^a$ which has no monochromatic $[Y]^a$ with $|Y| = b$. We will identify an m-coloring with a function $t\colon [c]^a \to k$. Let T be the collection of all such colorings $t\colon [c]^a \to k$, for all $c \in \omega$ and notice that T can be given the structure of a tree, where $t_2\colon [c_2]^a \to k$ extends $t_1\colon [c_1]^a \to k$ if $t_2 \restriction c_1 = t_1$. It is easy to see that T is finitely branching.

Consider the subcollection $T_{\text{bad}} \subseteq T$ which contains all bad colorings. Clearly T_{bad} is a subtree of T and therefore it is a finitely branching tree. Moreover the assumption: "for every c there is a bad m-coloring of $[c]^a$" implies that T_{bad} is also infinite. By König's Lemma there is and infinite branch $(t_n)_n$ in T_{bad}. Let $f = \bigcup_n t_n$. It is easy to see that f is a bad coloring of $[\omega]^a$, in that, it has no monochromatic $[Y]^a$ with $|Y| = b$. This contradicts, of course, Theorem 30. $\square$

Using Gödel's coding tricks one can actually write a sentence σ_{finR} in the (first order) language of Peano arithmetic, so that $(\omega, 0, S, +*) \models \sigma_{\text{finR}}$ if and only if Theorem 31 is true. Moreover, PA proves σ_{finR} (which is equivalent to PA $\models \sigma_{\text{finR}}$, by Gödel's completeness theorem). Of course the statement in the Theorem 30 is intrinsically infinitary and it cannot be "phrased" in Peano arithmetic. More interesting is the situation with the following statement known as the **Paris-Harrington principle**.

DEFINITION 32. Let $a, b, c, m \in \omega$. We write $c \to_{\text{PH}} (b)^a_m$ if whenever $[c]^a$ is m-colored, then there is a set $Y \subseteq c$ so that $[Y]^a$ is monochromatic and moreover:

(1) $|Y| \geq b$;
(2) $|Y| \geq \min Y$.

Using the same argument as in the proof of Theorem 31 we have that

THEOREM 33 (Paris-Harrington Principle). *For all $a, b, m \in \omega$ there is $c \in \omega$ with*

$$c \to_{\text{PH}} (b)^a_m$$

As in the case of Theorem 31, the above statement can be viewed as a sentence σ_{PH} in the language of first order Peano Arithmetic. However, it is a theorem of Paris and Harrington that PA cannot prove σ_{PH}! The interested reader may consult Marker's model theory book or my notes (Topics in Computability) for a sketch of the proof of this "negative result." This additional strength of the Zermelo set theory which allows us to prove Theorem 33 comes from the assumption that there exists an infinite set (axiom A.6.). This theme repeats in set-theory we will see for example later on that if we assume the existence of certain "large cardinals" which cannot be proved to exist in Z, then we can prove various results in analysis which cannot be proved in Z.

6. Countable sets

A set x is **countable** if $x \lesssim \omega$. It is **countably infinite** if $x \approx \omega$. We write $|x| \leq \aleph_0$ and $|x| = \aleph_0$ to denote that x is countable and countably infinite, respectively, and we say that the **cardinality of x is less than** $\aleph_0$ and **equal to** $\aleph_0$, respectively.

Here are some properties which follow directly by the definitions:

(1) $|x| = \aleph_0$ if and only if there is a surjective function $f \colon \omega \to x$;
(2) if $|x| \leq \aleph_0$ and $|x| \neq \aleph_0$, then $|x| = n$, for some $n \in \omega$;
(3) if $|x| \leq \aleph_0$ and $y \subseteq x$ then $|y| \leq \aleph_0$;
(4) if $|x| \leq \aleph_0$ and $f \colon x \to y$ then $|\text{rng}(x)| \leq \aleph_0$.

Here are some additional closure properties for countable sets

LEMMA 34. *Let $x, x_1, \ldots, x_n$ be countable sets. Then*

(1) $x_1 \times \cdots \times x_n$ is countable;
(2) $x_1 \cup \cdots \cup x_n$ is countable;

(3) $x^{<\omega} := \bigcup_{n\in\omega} x^n$ *is countable.*

Proof. (1) It suffices to show that $\omega \times \omega \approx \omega$. Consider the map $f\colon \omega\times\omega \to \omega$ with

$$f((m,n)) = \frac{(m+n+1)(m+n)}{2} + m.$$

It is easy "geometric" argument based on the lattice shows that this is a bijection:

$$\begin{array}{cccc} (0,0) & (0,1) & (0,2) & \dots \\ (1,0) & (1,1) & (1,2) & \dots \\ (2,0) & (2,1) & (2,2) & \dots \\ \vdots & \vdots & \vdots & \end{array}$$

(2) This is left as an exercise.

(3) It suffice to show that $\omega^{<\omega} \approx \omega$. For every $(k_0, \dots, k_{n-1}) \in \omega^n$ let

$$\langle k_0, \dots, k_{n-1}\rangle := p_0^{k_0+1} * \cdots * p_{n-1}^{k_{n-1}+1} - 1,$$

where $p_0, p_1, \dots$ is the unique increasing enumeration of primes. Let also $\langle\emptyset\rangle := 0$. $\square$

Theorem 35 ((AC)). *A countable union of countable sets is countable.*

Proof. Let $\{x_i \mid i \in \omega\}$ be a countable family of countable sets. Recall that, formally, this is just a function h with $\mathrm{dom}(h) = \omega$ and $h(i) = x_i$. For every $i \in \omega$ we pick a bijection $f_i\colon \omega \to x_i$. This is where (AC) is used. Let $g\colon \omega\times\omega \to \bigcup_{i\in\omega} x_i$ with $g(i,j) = f_i(j)$. Clearly g is onto $\bigcup_{i\in\omega} x_i$.

Formally, we let $\varphi\colon \mathcal{P}\big(\mathcal{P}(\omega\times\bigcup_{i\in\omega} x_i)\big) \to \mathcal{P}(\omega\times\bigcup_{i\in\omega} x_i)$ be a selection map by (AC). Let also

$$G = \{(n,F) \in \omega \times \mathcal{P}\big(\mathcal{P}(\omega \times \bigcup_{i\in\omega} x_i)\big) \mid F = \{\text{ all bijections } f\colon \omega \to h(n)\}\}$$

Then G is a function and $\varphi \circ G$ is a function assigning to each $i \in \omega$ the desired function f_i above. $\square$

Using these results we can show that many other sets such as $\mathbb{Z}$, $\mathbb{Q}$, and the set of all algebraic numbers, are countable. That being said, by the following theorem we also have that $\mathcal{P}(\omega)$ is not countable.

Theorem 36 (Cantor). *For any set X we have that $\mathcal{P}(X) \not\approx X$.*

Proof. Let $f\colon X \to \mathcal{P}(X)$ be any function. Then f cannot be onto: let

$$Z := \{x \in X \mid x \notin f(x)\}$$

Then $Z \in \mathcal{P}(X)$ and $Z \notin \mathrm{rng}(f)$ because if $f(z) = Z$ then we have a contradiction:

$$z \notin Z \iff z \in f(z) \iff z \in Z.$$

$\square$

We clearly have that $X \lnapprox \mathcal{P}(X)$ via $x \mapsto \{x\}$. Hence the "cardinality" of $\mathcal{P}(X)$ is strictly bigger than the cardinality of X. Notice that, by Cantor's theorem, even in the Zermelo universe $V_{\omega+\omega}$ we have infinitely many $\approx$-different infinities.

7. The Cantor-Schröder-Bernstein Theorem

Even if two sets X, Y can be brought into bijective correspondence, it is often difficult to find an explicit map $h\colon X \to Y$ which is both 1-1 and onto. It is often much easier to produce an injection $f\colon X \to Y$ and an injection $g\colon Y \to X$. The next theorem shows that these two injections can always be combined into the desired bijection. We observe that while a big part of the theory of cardinality depends on the axiom of choice, this is not the case for the next theorem.

Theorem 37 (Cantor-Schröder-Bernstein). $(X \lessapprox Y) \wedge (Y \lessapprox X) \implies X \approx Y.$

Proof. Going back to Hilbert's hotel thought experiment: think of X as being a collection of guests and Y as being a collection of rooms in the hotel. The given injection $g\colon Y \to X$ has assigned to each room a guest but there are possibly some guest left without a room: the guests in $X_0 := X \setminus g''Y$. We can use the injection $f\colon X \to Y$ to move the guests in X_0 to the rooms $f''X_0$ but these are already occupied by the guests in $X_1 := g''f''X_0$. So we have to displace X_1 using f to the rooms currently used by $X_2 := g''f''X_1$ and so on...

So define inductively $X_0 := X \setminus g''Y$ and $X_{n+1} := g''f''X_n$. Let also $X_\infty = \bigcup_n X_n$. Define $h\colon X \to Y$ with $h(x) = f(x)$, if $x \in X_\infty$; and $h(x) = g^{-1}(x)$, if $x \in X \setminus X_\infty$.

h is a function, since g is injective, whose domain is clearly X. It also clearly injective on X_∞ and injective on $X \setminus X_\infty$. So let $x \in X_\infty$ and $x' \in X \setminus X_\infty$. Assume that $h(x) = h(x')$. That is $f(x) = g^{-1}(x')$ and therefore $x' = g \circ f(x)$. Since $x \in X_n$ for some n, this implies that $x' \in X_{n+1}$, a contradiction.

Surjectivity is easy to see since at its stage of the "displacement procedure" above every room is always occupied (left as an exercise). ☐

8. The cardinality of the continuum

Classically, the continuum is the space $\mathbb{R}$ of all reals. Recall that a real r is by definition a Dedekind cut, i.e., any subset $r \subseteq \mathbb{Q}$ so that

(1) (bounded) there is $q \in \mathbb{Q}$ so that for all $p \in r$ we have $p < q$.
(2) (downward closed) if $p < q$ and $q \in r$, then $p \in r$.

We say that X has the **cardinality of the continuum** and we denote this by $|X| = 2^{\aleph_0}$ if $X \approx \mathbb{R}$. In modern set theory, "the continuum" refers to either of the following sets, or more generally to any *Polish space* (to be defined):

$$2^\omega \approx n^\omega \approx \omega^\omega \approx \mathcal{P}(\omega) \approx \mathbb{R}.$$

Lemma 38. *For all $n \in \omega$, with $n > 1$, we have $2^\omega \approx n^\omega \approx \omega^\omega$.*

PROOF. It is clear that $2^\omega \lessapprox n^\omega \lessapprox \omega^\omega$, since $2^\omega \subseteq n^\omega \subseteq \omega^\omega$. By Cantor-Schröder-Bernstein it suffices to show that $\omega^\omega \lessapprox 2^\omega$. The desired injection is given by:

$$(n_0, n_1, n_2, \ldots) \mapsto (\overbrace{1, \ldots, 1}^{n_0+1}, 0, \overbrace{1, \ldots, 1}^{n_1+1}, 0, \overbrace{1, \ldots, 1}^{n_2+1}, 0, \ldots).$$

□

LEMMA 39. *$2^\omega \approx \mathcal{P}(\omega) \approx \mathbb{R}$.*

PROOF. $2^\omega \approx \mathcal{P}(\omega)$ as witnessed by the bijection which sets any $A \in \mathcal{P}(\omega)$ to its characteristic map $\chi_A \in 2^\omega$. Since $\mathbb{R} \subseteq \mathcal{P}(\mathbb{Q})$ and $\mathbb{Q} \approx \omega$, we have that $\mathbb{R} \lessapprox \mathcal{P}(\omega)$. By Cantor-Schröder-Bernstein it suffices to find an injection from 2^ω to $[0,1] \subseteq \mathbb{R}$.

We define inductively a closed interval I_s for every $s \in 2^{<\omega}$ as follows. Let $I_\emptyset = [0,1]$. Given $I_s = [a,b]$ let $I_{s^\frown 0}$ be the left and $I_{s^\frown 1}$ be the right third of $[a,b]$:

$$I_{s^\frown 0} := [a, a + \frac{b-a}{3}], \quad I_{s^\frown 1} := [a + 2\frac{b-a}{3}, b]$$

By compactness of $[0,1]$ (see HW2), for every $\alpha \in 2^\omega$ the intersection $\bigcap_{n\in\omega} I_{\alpha|n}$ is non-empty. Moreover, since the diameter of $I_{\alpha|n}$ goes to 0 as n increases we have that $\bigcap_{n\in\omega} I_{\alpha|n}$ is a singleton. This defines a function $f\colon 2^\omega \to [0,1]$ with

$$\{f(\alpha)\} = \bigcap_{n\in\omega} I_{\alpha|n}.$$

We leave to the reader to check that f is injective. □

LEMMA 40. *If $A \subseteq \mathbb{R}$ is countable then $|\mathbb{R} \setminus A| = 2^{\aleph_0}$.*

PROOF. It suffices to find $B \subseteq \mathbb{R} \setminus A$ with $|B| = \aleph_0$ because then we can break $B = B_0 \sqcup B_1$ to the even B_0 and odd B_1 elements of the enumeration witnessing $|B| = \aleph_0$ and define a bijection $h\colon \mathbb{R} \setminus A \to \mathbb{R}$ where $h''B_0 = B$, $h''B_1 = A$ and $h(x) = x$ for all $x \in \mathbb{R} \setminus (A \bigcup B)$.

By (AC) we immediately get such B by we leave to the reader to use a diagonal argument to get such B without invoking (AC). □

THEOREM 41. *Let $|X| = 2^{\aleph_0}$, $|Y| = 2^{\aleph_0}$, $|Z| = \aleph_0$. Then*
(1) $|X \bigcup Y| = 2^{\aleph_0}$;
(2) $|X \times Y| = 2^{\aleph_0}$;
(3) $|X^Z| = 2^{\aleph_0}$.

PROOF. For (1) it is clear that $2^\omega \approx X \lessapprox X \bigcup Y$. By CSB it suffices to show that $X \bigcup Y \lessapprox 2^\omega$. Let $f\colon X \to 2^\omega$ and $g\colon Y \to 2^\omega$ be injections (which exist by assumption) and let $h\colon X \bigcup Y \to 2^\omega$ by $h(x) = (0)^\frown f(x)$, if $x \in X$; and $g(x) = (1)^\frown g(x)$ otherwise. Then h is clearly an injection.

For (2), notice that $((a_0, a_1, \ldots), (b_0, b_1, \ldots)) \mapsto (a_0, b_0, a_1, b_1, \ldots)$ is a bijection between $2^\omega \times 2^\omega$ and 2^ω.

For (3), notice that if $f\colon \omega\times\omega \to \omega$ is a bijection then the map

$$\big((a_0^0, a_1^0, a_2^0, \ldots), (a_0^1, a_1^1, a_2^1, \ldots), (a_0^2, a_1^2, a_2^2, \ldots), \cdots\cdots\big) \mapsto (b_0, b_1, b_2, \ldots)$$

with $b_{f(n,m)} = a_m^n$, is a bijection from $(2^\omega)^\omega$ to 2^ω. $\square$

Corollary 42. *The set $C(\mathbb{R},\mathbb{R})$ of all continuous real functions has the cardinality of the continuum. The sets* $\mathrm{Open}(\mathbb{R})$, $\mathrm{Closed}(\mathbb{R})$ *of all open subsets and all closed subsets of $\mathbb{R}$ has the cardinality of the continuum.*

Proof. Notice that $\mathbb{R} \lessapprox C(\mathbb{R},\mathbb{R}) \lessapprox \mathbb{R}^{\mathbb{Q}}$ since the constant functions are continuous and since every continuous function is entirely determined its values on the rationals. The rest follows from CSB and Theorem 41.

It is clear that $\mathrm{Open}(\mathbb{R}) \approx \mathrm{Closed}(\mathbb{R})$ and that $\mathbb{R} \lessapprox \mathrm{Open}(\mathbb{R})$. To see that $\mathrm{Open}(\mathbb{R}) \lessapprox \mathbb{R}$ notice that $\mathrm{Open}(\mathbb{R}) \lessapprox \mathcal{P}(\mathbb{Q}\times\mathbb{Q})$, since every open set U is determined by all open intervals $(p,q) \subseteq U$ with rational coefficients. $\square$

We write $|X| \leq 2^{\aleph_0}$ there is an injection from X to $\mathbb{R}$

Theorem 43 ((AC)). *(1) If $f\colon X \to Y$ and $|X| = 2^{\aleph}$ then $|\mathrm{rng}(f)| \leq 2^{\aleph_0}$.*

(2) The union of continuum many sets of the cardinality of the continuum has the cardinality of the continuum.

Proof. The ideas used in this proof have been already laid out in previous sections and are left to the reader. We are going to repeat one final time in later section when we develop the general theory of cardinality under (AC). $\square$

9. The continuum hypothesis: part I

We have established (under AC) that $\aleph_0$ is the least infinite cardinality. That is, if X is an infinite set then $\omega \lessapprox X$. We have also seen that applications of the powerset operation strictly increase the cardinality. That is,

$$\aleph_0 \lneq 2^{\aleph_0}.$$

The obvious question is whether there is a set whose cardinality is strictly between $\aleph_0$ and $2^{\aleph_0}$. To phrase it in more classical terms:

Problem 44. Is there a set $A \subseteq \mathbb{R}$ which is neither countable, nor of the size of the continuum?

Cantor believed that this is not the case. In fact, he put a lot of energy in proving what we now call the "continuum hypothesis."

The Continuum Hypothesis.

CH. Every uncountable set of reals has the cardinality of the continuum.

In 1939 Gödel proved that if set-theory (ZFC) is consistent there is a model of set theory (ZFC) in which (CH) holds. In 1963 Cohen developed the forcing technique which he used to show that if set-theory (ZFC) is consistent then there is a model of set theory (ZFC) in which $\neg$(CH) holds. In other words Cantor's efforts to prove (CH) using the standard axioms of set theory (ZFC) were doomed to fail.

That being said, we now know that any set $A \subseteq \mathbb{R}$ which witnesses the negation of (CH) has to be too complicated to be accessible through countable limiting constructions which naturally occur in topology and analysis. In other words, every "sufficiently simple" set $A \subseteq \mathbb{R}$ is either countable or of the size of the continuum.

In this section we will establish that if $A \subseteq \mathbb{R}$ is closed (in the standard topology of $\mathbb{R}$) then A satisfies (CH). Next quarter we will extend the analysis of this section to show the same is true when A is Borel or even analytic (which makes precise what I above called "*accessible through countable limiting constructions which naturally occur in topology and analysis*"). In a later section we will also construct Gödel's universe and next quarter we will develop the forcing method and construct Cohen's universe.

DEFINITION 45. A subset U of $\mathbb{R}$ is closed if it is the union of some family $\mathcal{U}$ consisting of **open** intervals (a, b). A set F is **closed** if F^c is open. A set P is **perfect** if it is closed and it has no isolated points, i.e., whenever $(a, b) \cap P \neq \emptyset$ then $|(a, b) \cap P| > 1$.

It is immediate from the definition that if P is perfect and $(a, b) \cap P \neq \emptyset$ then $|(a, b) \cap P| \geq \aleph_0$. In fact we can work a little harder to prove the following theorem

THEOREM 46. *If $P \subseteq \mathbb{R}$ is perfect and non-empty then $|P| = 2^{\aleph_0}$.*

PROOF. Since $P \neq \emptyset$ we can find an interval $I = [a, b]$ so that $(a, b) \cap P \neq \emptyset$. For each $s \in 2^{<\omega}$ we define a closed interval $I_s = [a_s, b_s]$ so that:

(1) $I_\emptyset = I$;
(2) $(a_s, b_s) \cap P \neq \emptyset$;
(3) $I_{s^\frown 0} \cap I_{s^\frown 1} = \emptyset$.
(4) $\mathrm{diam}(I_s) \leq [a, b]/2^{\mathrm{length}(s)}$

For the construction, assume that $I_s = [a_s, b_s]$ is given as above. Since $(a_s, b_s) \cap P \neq \emptyset$ and P is perfect there are at least two elements $x_0 < x_1$ of P inside $[a_s, b_s]$. Let $I_{s^\frown 0}$ be the interval with center x_0 and radius less than $(x_1 - x_0)/2$. Similarly $I_{s^\frown 1}$ be the interval with center x_1 and radius less than $(x_1 - x_0)/2$. It is clear that $I_{s^\frown 0}, I_{s^\frown 1}$ satisfy (3) and (4) above. They also clearly satisfy (2), perpetuating this way the induction. As in the case of Lemma 39 we define a map $f \colon 2^\omega \to I$ with

$$\{f(\alpha)\} = \bigcap_{n \in \omega} I_{\alpha|n}$$

Compactness of I guarantees that the intersection is indeed non empety and property (4) above guarantees that the intersection is a singleton. Hence f is well defined. Property (3) above guarantees that f is an injection. Finally property (2) above and the fact that P is closed shows that $\mathrm{rng}(f) \subseteq P$. Hence we have that $2^\omega \lessapprox P$, which finishes the proof. $\square$

THEOREM 47 (Cantor-Bendixson). *If $F \subseteq \mathbb{R}$ is a closed then $F = P \cup N$ where P is perfect, N is countable, and $P \cap N = \emptyset$. Moreover, this decomposition is unique, i.e., if $F = P' \cup N'$ satisfies the above conditions then $P = P'$ and $N = N'$.*

Before we prove this theorem we record the following immediate corollary.

COROLLARY 48. *If F is closed and uncountable then $|F| = 2^{\aleph_0}$.*

PROOF OF THEOREM 47. For every $A \subseteq \mathbb{R}$ and every $x \in \mathbb{R}$ we say that x is a **condensation point** of A if for every open interval (a, b) with $x \in (a, b)$ we have that $(a, b) \cap A$ is uncountable. Let $\widetilde{A} = \{x \in \mathbb{R} \mid a \text{ is a condensation point of} A\}$. Notice, for example, that $\widetilde{A} = \emptyset$ if A is countable.

Notice that is F is closed then $\widetilde{F} \subseteq F$, since every condensation point is, in particular, a limit point. We therefore have that

$$F := \widetilde{F} \bigcup (F \setminus \widetilde{F}),$$

and we claim that setting $P := \widetilde{F}$ and $N = F \setminus \widetilde{F}$ works.

$F \setminus \widetilde{F}$ **is countable.** Notice that if $x \in F \setminus \widetilde{F}$ then there is an interval (p, q) with with rational coefficients so that $x \in (p, q)$ and $(p, q) \cap A$ is countable. We have

$$F \setminus \widetilde{F} = \bigcup \{(p, q) \cap A \mid (p, q) \text{ as above}\}$$

which is countable, as a union of countably many countable sets.

$\widetilde{F}$ **is perfect.** Notice first that is closed since if $x \in \widetilde{F}^c$ then there is (a, b) with $x \in (a, b)$ and $(a, b) \cap F$ countable and therefore $(a, b) \subseteq \widetilde{F}^c$. It also has no isolated points since if $x \in \widetilde{F}$ was isolated then we could find an interval (a, b) so that $(a, b) \cap \widetilde{F} = \{x\}$. But then $(a, b) \cap F$ would be countable since we already showed that $F \setminus \widetilde{F}$ is countable, contradicting that x was a condensation point.

For the second part of the statement, we first point out that the "local" version of Theorem 46 also holds, i.e., if $P \subseteq \mathbb{R}$ is perfect and $P \cap (a, b) \neq \emptyset$ then $P \cap (a, b) \neq \emptyset$ is uncountable. We leave to the reader to confirm this. Assume now that $F = P' \sqcup C'$ is another decomposition of F into a perfect set and a countable set. By the local version Theorem 46 we have that every x in P' is a condensation point of F. Hence $P' \subseteq P$. Conversely, we also have that $P \subseteq P'$ since if $x \in P \setminus P'$, then we can find an interval (a, b) with $x \in (a, b)$ and $(a, b) \cap P' = \emptyset$. This because P' is closed. But then $x \in (a, b) \cap P \subseteq N'$ which contradicts that N' is countable by the local version of Theorem 46. $\square$

10. Wellorderings and transfinite induction

Looking back at the notes, ω was the first infinite set for which we had a systematic way of defining functions out of. This "systematic method" was definition by induction–see Theorem 10 and Theorem 18–and relied on the fact that ω came equipped with a linear order that satisfied the wellordering principle. We will extend the nice properties of ω to more general wellorderings.

Let $(W, <)$ be a linear order. We say that it is a **wellordering** if every non-empty $X \subseteq W$ has a minimum element, that is:

$$\forall X \subseteq W\big(X \neq \emptyset \implies \exists x \in X\ \forall y\ (y \in X \implies x \leq y)\big)$$

Notice that $(\omega, <)$ and $(n, <)$ are wellorderings while $(\mathbb{Z}, <)$ and $(\mathbb{Q}, <)$ are not.

If $\mathcal{A} = (A, <)$ and $\mathcal{B} = (B, <)$ are two linear orderings with $A \cap B = \emptyset$ then we can define the linear ordering $\mathcal{A} + \mathcal{B} = (A \cup B, <^+)$ so that $<^+$ on the A part looks like $\mathcal{A}$, $<^+$ on the B part looks like $\mathcal{B}$ and for every $a \in A$, $b \in B$ we have $a <^+ b$. Similarly we can define $\mathcal{A} \cdot \mathcal{B} = (A \times B, <^\times)$ as the antilexicographic ordering:

$$(a, b) <^\times (a', b') \iff \big((b < b') \vee (b = b' \wedge a < a')\big).$$

Notice that $\omega \cdot 2$ looks like $\omega + \omega$ (assume the copies are disjoint),

$$\underbrace{\bullet\bullet\bullet\dots}_{\omega}\quad\underbrace{\bullet\bullet\bullet\dots}_{\omega}$$

while $2 \cdot \omega$ looks like ω.

EXERCISE 49. *If $\mathcal{A} = (A, <)$ and $\mathcal{B} = (B, <)$ are wellorderings then so are*

$$\mathcal{A} + \mathcal{B} \quad \text{and} \quad \mathcal{A} \cdot \mathcal{B}.$$

For every wellordering $\mathcal{W} = (W, <)$ there are three types of points: a unique "least"; the "successor" points; and the "limit" points. We first need a lemma.

LEMMA 50. *Let $\mathcal{W} = (W, <)$ be a wellordering. Then*
1. *W has a least element;*
2. *if $x \in W$ and there is $y \in W$ with $x < y$, then x has a successor;*
3. *if $X \subseteq W$ is bounded from above (there is $y \in W$ with $x \leq y$ for all $x \in X$), then X has a least upper bound (a least $z \in W$ with $x \leq z$ for all $x \in X$).*

PROOF. For (3), if X has a maximum element then that element is clearly a least upper bound. Otherwise, the least upper bound is the least element of

$$\{z \in W \mid \forall x \in X\ x < z\}.$$

□

Let x be a point in the wellordering $\mathcal{W} = (W, <)$. We say that x is a **successor** if $\{y \in W \mid y < x\}$ has a largest element. If x is neither the **least** element nor a **successor** then we say that x is a **limit**.

THEOREM 51 ((AC)). *Let $\mathcal{W} = (W, <)$ be a linear ordering. It is a wellordering if and only if it has no infinite decreasing chain $x_0 > x_1 > x_2 > \dots$.*

PROOF. If $(x_n)_{n\in\omega}$ is an infinite decreasing sequence then $X := \{x_n \mid n \in \omega\}$ has no least element. Conversely assume that $Y \subseteq W$ has no least element Fix a choice function $\varphi \colon \mathcal{P}(Y) \to Y$. Define inductively:

$$x_0 := \varphi(Y) \quad \text{and} \quad x_{n+1} = \varphi(\{x \in Y \mid x < x_n\}).$$

□

An **initial segment** of the wellordering $(W, <)$ is any set $S \subseteq W$ which is $<$-downward closed, i.e., if $y \in S$ and $x < y$ then $x \in S$. It is **proper** if $S \neq W$. If $x \in W$ then set $W_x := \{y \in W \mid y < x\}$. Notice that $W_x = \emptyset$ if and only if x is the least element.

LEMMA 52. *S is a proper initial segment of $(W, <)$ iff $S = W_x$ for some $x \in X$.*

PROOF. Let x be the least element of $W \setminus S$. □

THEOREM 53 (Transfinite induction principle). *Let $\mathcal{W} = (W, <)$ be a wellordering and $X \subseteq W$. If for all $x \in W$ we have $(W_x \subseteq X \implies x \in X)$, then $X = W$.*

PROOF. If not, then let x be the least element of $W \setminus X$ and notice that $W_x \subseteq X$. □

The next theorem gives us a systematic way of defining maps on domain W, for every wellordering $\mathcal{W} = (W, <)$.

THEOREM 54. *Let $\mathcal{W} = (W, <)$ be a well ordering and let A be any set. Let also $g \colon S \to A$ be a function from the set $S := \bigcup_{x \in W} A^{W_x}$ of all maps from any proper initial segment of W to A. Then there is a unique function $f \colon W \to A$ with*

$$f(x) = g(f{\restriction}W_x, x)$$

PROOF. The proof follows the lines of the proof of Theorem 10.

For uniqueness. If $f_1 = g(f_1{\restriction}W_x, x)$ and $f_2(x) = g(f_2{\restriction}W_x, x)$ for all $x \in W$, then by Theorem 53 the set

$$X = \{x \in W \mid f_1(x) = f_2(x)\}$$

is easily shown to be equal to W.

For existence. Consider the set $\mathcal{F}$ of all "approximations" to f:

$$\mathcal{F} = \{u \in S \mid \text{ if } u \colon W_x \to A \text{ then } \forall y < x \ \big(u(y) = g(u{\restriction}W_y, y)\big)\},$$

and set $f := \bigcup \mathcal{F}$. We need to show that f is the desired object:

(1) f is a relation;
(2) $\mathrm{dom}(f) \subseteq W$ and $\mathrm{rng}(f) \subseteq A$;
(3) f is a function;
(4) $\mathrm{dom}(f)$ is an initial segment of W and if $x \in \mathrm{dom}(f)$, then $f(x) = g(f{\restriction}W_x, x)$;
(5) $\mathrm{dom}(f) = W$.

Claims (1), (2) follow immediately from the fact that $f = \mathcal{F}$, since the elements $u \in \mathcal{F}$ satisfy (1),(2). For (3) notice that if $x \in \mathrm{dom}(u) \cap \mathrm{dom}(v)$ for some $u, v \in \mathcal{F}$ then $\mathrm{dom}(u) \cap \mathrm{dom}(v) = W_{x_0}$ for some x_0 since the intersection of initial segments is an initial segment. As in the proof of uniqueness of f we can use transfinite induction to show that $u{\restriction}W_{x_0} = v{\restriction}W_{x_0}$, and therefore $u(x) = v(x)$. So the relation f is in fact a function. For claim (4) notice that the union of initial segments is an initial segment. Hence $\mathrm{dom}(f) = \mathrm{dom}(\bigcup \mathcal{F})$ is an initial segment. Moreover if $x \in \mathrm{dom}(f)$ then

$x \in \mathrm{dom}(u)$ for some $u \in \mathcal{F}$. Hence $f(x) = u(x) = g(u \restriction W_x, x) = g(f \restriction W_x, x) = f(x)$. For (5), if not, then $\mathrm{dom}(f) = W_{x_0}$ for some $x_0 \in W$. But then the following map is clearly in $\mathcal{F}$:

$$f \bigcup \{(x_0, g(f \restriction W_{x_0}, x_0))\},$$

contradicting that $x_0 \notin \mathrm{dom}(\bigcup \mathcal{F})$. $\square$

11. Comparing wellorderings

One important aspect of wellorderings is that they are pairwise comparable with respect to embeddings. Moreover, whenever two wellorderings are isomorphic, there is a unique such isomorphism.

Theorem 55. *Let $(W, <)$ be a wellordering and let $f \colon W \to W$ be any map with $x < y \implies f(x) < f(y)$. Then $f(x) \geq x$ for all $x \in W$.*

Proof. Assume that for some x we have $f(x) < x$. Since f is order preserving we have $f(f(x)) < f(x)$, $f(f(f(x))) < f(f(x))$, etc. But then, using Theorem 10, we may build an infinite decreasing sequence

$$x > f(x) > f(f(x)) > \cdots$$

which contradicts (via Theorem 51) that $(W, <)$ is a wellordering. $\square$

Corollary 56. *If $(W, <)$ and $(W', <)$ are isomorphic wellorderings, then there is a unique isomorphism between them. In particular, $\mathrm{Aut}((W, <)) = \{\mathrm{id}_W\}$.*

Proof. Assume that $f \colon W \to W'$ and $g \colon W \to W'$ are two different isomorphisms. But then both $g^{-1} \circ f \colon W \to W$ and $f^{-1} \circ g \colon W \to W$ are order preserving. By the previous theorem we respectively have $g^{-1} \circ f(x) \leq x$ and $f^{-1} \circ g(x) \leq x$. But

$$g^{-1} \circ f(x) \leq x \implies f(x) \leq g(x), \quad \text{and} \quad f^{-1} \circ g(x) \leq x \implies g(x) \leq f(x),$$

Since g and f are order preserving. It follows that $f(x) = g(x)$ for all $x \in W$. $\square$

Corollary 57. *Let $(W, <)$ be a wellordering.*

(1) for each $z \in W$ we have that $(W_z, <) \not\simeq (W, <)$.
(2) if $x, y \in W$ with $x \neq y$, then $(W_x, <) \not\simeq (W_y, <)$.

Proof. If $f \colon W \to W_x$ was an isomorphism, then f is, in particular a map from W to W with $x < y \implies f(x) < f(y)$. But then, $f(x) \in W_x \implies f(x) < x$, contradicting Theorem 55. Similarly for (2). $\square$

The following result is fundamental for the theory of wellorderings

Theorem 58. *Let $(W, <)$, $(W', <)$ be two wellorderings. Then precisely one holds:*

(1) $(W, <) \simeq (W', <)$;
(2) there is unique $x' \in W'$ so that $(W, <) \simeq (W'_{x'}, <)$;

(3) there is unique $x \in W$ so that $(W_x, <) \simeq (W, <)$.

Proof. Uniqueness is easily established using (2) of the last corollary. It therefore suffices to prove the above without any mention to uniqueness.

Fix some new element ∞ outside of W' and define using Theorem 54 a map:

$$f \colon W \to W' \cup \{\infty\}$$

by setting $f(x) :=$"the least element of $W' \setminus \{f(z) \mid z \in W_x\}$", if the latter set is non-empty; and $f(x) := \infty$, otherwise. We may now consider two cases:

Case 1. $f(x) \neq \emptyset$ for all $x \in W$. Then $f \colon W \to W'$. Clearly f is order preserving. We are left to prove the following claim which implies that one of the alternatives (1) or (2) of the trichotomy holds:

Claim. $\mathrm{rng}(f)$ is an initial segment of W'.

To prove the claim, notice that if $\mathrm{rng}(f) = W'$ then we are done. Otherwise let x' be the least element of $W' \setminus \mathrm{rng}(f)$. It is clear that $W'_{x'} \subseteq \mathrm{rng}(f)$, so it suffices to show the converse as well. Assume towards contradiction that $\mathrm{rng}(f) \nsubseteq W'_{x'}$ and let $w \in W$ be the least element with $x' \leq f(w)$. But then, since $f(x) < x'$ for all $x < w$ we have by the definition of f that $f(w) \leq x'$. But then $f(w) = x'$ contradicting that x' is not in $\mathrm{rng}(f)$.

Case 2. If $f(x) = \infty$ for some $x \in W$ then pick x to be the least such. This implies that $f \restriction W_x \colon W_x \to W'$. It is clear that $f \restriction W_x$ is order preserving. As in case 1 one argues that the range of $f \restriction W_x$ is an initial segment of W'. This initial segment cannot be proper since this would imply that $f(x) \neq \infty$. It follows that $f \restriction W_x$ is an isomomorphism from W_x to W'. □

Definition 59. Let $\mathcal{W} = (W, <)$ and $\mathcal{W}' = (W', <)$ be two wellorderings. We set:

$$\mathcal{W} <_{\mathrm{seg}} \mathcal{W}' \iff \mathcal{W} \simeq (W'_{x'}, <) \text{ for some } x' \in W'$$

$$\mathcal{W} \leq_{\mathrm{seg}} \mathcal{W}' \iff \big(\mathcal{W} <_{\mathrm{seg}} \mathcal{W}' \vee \mathcal{W} \simeq (W', <)\big),$$

and we say that $\mathcal{W}$ **is isomorphic to a proper initial segment of** $\mathcal{W}'$ and $\mathcal{W}$ **is isomorphic to an initial segment of** $\mathcal{W}'$, respectively.

The collection WELLORD of all wellorderings is not a set (formally it is just the formula $\varphi_{\mathrm{wellord}}(x)$, stating that x is a wellordering) behaves itself as a wellordering under $<_{\mathrm{seg}}$, if we identify isomorphic wellorderings. We make this precise in the following theorem.

Theorem 60. *For any wellorderings $\mathcal{W}, \mathcal{W}', \mathcal{W}''$ we have that:*

(1) $\mathcal{W} <_{\mathrm{seg}} \mathcal{W}'$ of $\mathcal{W}' <_{\mathrm{seg}} \mathcal{W}$ or $\mathcal{W} \simeq \mathcal{W}'$.
(2) $\mathcal{W} <_{\mathrm{seg}} \mathcal{W}' \implies \mathcal{W}' \nless_{\mathrm{seg}} \mathcal{W}$.
(3) $\mathcal{W} <_{\mathrm{seg}} \mathcal{W}'$ and $\mathcal{W}' <_{\mathrm{seg}} \mathcal{W}''$ implies $\mathcal{W} <_{\mathrm{seg}} \mathcal{W}''$.
(4) if COLL is a non-empty collection of wellorderings (given by $\psi(x)$, say), then there is a $<_{\mathrm{seg}}$-least element of COLL.

Moreover, if X is a set of wellorderings then there is a wellordering $\mathcal{W}_X = (W_X, <)$ which is the $\leq_{\text{seg}}$-(least upper bound) of X.

Proof. The first three properties follow immediately from Theorem 58.

For property (4), let $\mathcal{W} = (W, <)$ in COLL. Let x be the least element W for which there is $\mathcal{W}'$ in COLL with $(W_x, <) \simeq \mathcal{W}'$ (if such x does not exist then $\mathcal{W}$ is already least). Then $\mathcal{W}'$ is the least element.

For property (5) notice that, by property (4), it suffices to find a wellordering that is $\leq_{\text{seg}}$-above all elements of X. We may assume without loss of generality that if $(W, <), (W', <) \in X$ then $W \cap W' = \emptyset$. Let $A = \bigcup\{W \mid (W, <) \in X\}$ and consider the equivalence relation

$$a \sim b \iff W_a^{(a)} \sim W_b^{(b)},$$

where $W^(c)$ is the unique element of X containing c, for every $c \in A$. Let $[a]$ be the equivalence class of $[a]$ under $\sim$ and let $W_X := W/\sim$. Define

$$[a] <_X [b] \iff W_a^{(a)} <_{\text{seg}} W_b^{(b)}$$

It is easy to see that $(W_X, <_X)$ is a wellordering and $a \mapsto [a]$ shows that each $(W^{(a)}, <)$ in X is $\leq_{\text{seg}}$-below $(W_X, <_X)$. □

12. Zermelo's wellordering theorem

Recall that when a set, like ω, admits a nice structure, e.g. a wellordering $<$, then (AC) is not required for defining a choice function $\varphi\colon \mathcal{P}(\omega) \setminus \{\emptyset\} \to \omega$. Zermelo's theorem shows that (AC) is, in fact, equivalent to the assumption that we can "nicely structure" every set.

Theorem 61. *The axiom of choice is equivalent to the assumption that every set X admits a wellordering.*

Proof. If $<$ is a wellordering on X, then $Z \mapsto \min_<(Z)$ is clearly a choice function for $\mathcal{P}(X)$.

Conversely, let $\mathcal{I} := \{\mathcal{W} \mid \mathcal{W}$ is a wellordering with domain $W \subseteq X\}$. By Theorem 60 we have a wellordering $\mathcal{W}_\mathcal{I} = (W_\mathcal{I}, <_\mathcal{I})$ so that $\mathcal{W} <_{\text{seg}} \mathcal{W}_\mathcal{I}$, for all $\mathcal{W} \in \mathcal{I}$. Indeed one may find some $\mathcal{W}_0$ so that $\mathcal{W} \leq_{\text{seg}} \mathcal{W}_0$, for all $\mathcal{W} \in \mathcal{I}$, and set $\mathcal{W}_\mathcal{I} := \mathcal{W}_0 + 1$. Using transfinite induction (Theorem 54), for some $\infty \notin W_\mathcal{I}$ we define a map

$$f\colon W_\mathcal{I} \to X \cup \{\infty\},$$

, by $f(w) = \varphi(X \setminus \text{rng}(f \restriction (W_\mathcal{I})_w))$, where φ is a choice function for $\mathcal{P}(X)$.

If $f(w) \neq \infty$ for all w, then $f\colon \mathcal{W}_\mathcal{I} \to X$. Since f is clearly injective, $\text{rng}(f)$ may be endowed with a wellordering $\mathcal{W} := (\text{rng}(f), <_\mathcal{I}^f)$, which is the push-forward of $<_\mathcal{I}$ under f. Clearly $\mathcal{W} \in \mathcal{I}$ and therefore $\mathcal{W} <_{\text{seg}} \mathcal{W}_\mathcal{I}$. This implies that f is also surjective by definition of f and since there is still "room" in the domain. Hence $\mathcal{W} \simeq \mathcal{W}_\mathcal{I}$ by f, contradicting that $\mathcal{W} <_{\text{seg}} \mathcal{W}_\mathcal{I}$.

We therefore have that $f(w) = \infty$ for some w; and pick w to be the least such. As in the previous paragraph we can then see that $f\restriction(\mathcal{W}_{\mathcal{I}})_w$ is a bijection between $W_{\mathcal{I}})_w$ and X, so we can take the push-forward wellordering. $\square$

Corollary 62 (AC). *For any two sets x, y we have that either $x \lessapprox y$ or $y \lessapprox x$.*

Proof. Use Theorem 61 and then Theorem 58. $\square$

Corollary 63 (AC). *For every collection* COLL *of sets there is an element x in* COLL *of least cardinality, i.e. $x \lessapprox y$ for all y in* COLL.

Proof. Use Theorem 61 and then Theorem 60. $\square$

13. Ordinals I

We saw that if we mod out the class WELLORD of all wellorderings by the isomorphism relation $\simeq_{\mathrm{ISO}}$ between wellorderings then we get a class (not a set) which behaves exactly as a well ordering. We will view the collection ORD of all ordinals as a collection of canonical representatives for the equivalence classes in

$$\mathrm{WELLORD}/\simeq_{\mathrm{ISO}}$$

Recall that x is **transitive** if for all $y \in x$ and all $z \in y$ we have $x \in z$. Recall also:

Definition 64. A set α is an **ordinal** if it is transitive and wellordered by $\in$, i.e.,

(a) (transitive) $\forall y, z\, (z \in y \in \alpha \implies z \in \alpha)$;
(b) $\forall x \in \alpha\, (x \notin x)$;
(c) $\forall x, y \in \alpha\, (x \in y \vee x = y \vee y \in x)$;
(d) $\forall y, z\, \forall x \in \alpha\, (z \in y \in x \implies z \in x)$;
(e) $\forall X \subseteq \alpha (X \neq \emptyset \implies \exists x \in X \forall y \in X (x = y \vee x \in y))$.

We will denote ordinals by letters $\alpha, \beta, \gamma, \ldots$. We denote by ORD the collection of all ordinals. We often write $\alpha < \beta$ for $\alpha \in \beta$ and $\alpha \leq \beta$ for $(\alpha < \beta) \vee \alpha = \beta$.

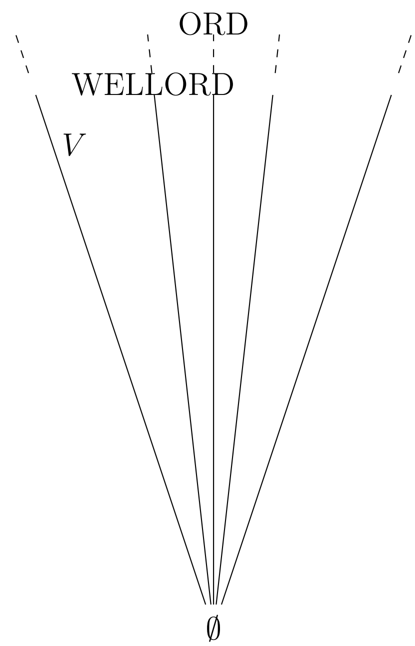
Examples. Every natural number n is an ordinal and ω is an ordinal. Moreover we can produce new ordinals such as $\omega^+ := \omega \cup \{\omega\}$ using the following operations:

(1) If α is an ordinal then so is $\omega^+ := \alpha \cup \{\alpha\}$.
(2) If A is a set of ordinals then $\mathrm{lub}(A) := \bigcup A$ is an ordinal.

It is immediate from the definitions that α^+ is an ordinal. Before we prove that $\mathrm{lub}(A)$ is an ordinal we will establish some basic properties of ordinals.

Lemma 65. *Let α, β be ordinals. We have*

(1) $\alpha \notin \alpha$;
(2) if $x \in \alpha$ then x is an ordinal;
(3) if $x \subseteq \alpha$ and x is transitive then x is an ordinal;
(4) $\alpha \in \beta \iff \alpha \subsetneq \beta$;
(5) $(\alpha < \beta) \vee (\alpha = \beta) \vee (\beta < \alpha)$;



PROOF. (1) if $\alpha \in \alpha$ then $\alpha \in \alpha \in \alpha$, contradicting (a) from the definition of an ordinal. For (2), by transitivity of α we have that $x \subseteq \alpha$ hence $(x, \in)$ is a subordering of $(\alpha, \in)$, and therefore, a wellordering. Finally x is transitive by (d) from the definition of an ordinal. The same argument shows (3).

For (4), by transitivity, $\alpha \in \beta$ implies $\alpha \subseteq \beta$; and by (1) the inclusion is strict. For the converse, assume that $\alpha \subsetneq \beta$. Then $\beta \setminus \alpha$ has a least element, call it γ. We will show that $\gamma = \alpha$. Notice that if $\delta \in \gamma$ then by minimality of γ we have $\delta \notin \beta \setminus \alpha$. Hence $\delta \in \alpha$. So $\delta \subseteq \alpha$. But also if $\delta \in \alpha$ then $(\gamma \in \delta) \vee (\gamma = \delta) \vee (\delta \in \gamma)$. The first two cases imply that $\gamma \in \alpha$, a contradiction since $\gamma \in \beta \setminus \alpha$.

For (5) assume that $\alpha \neq \beta$ and consider the set $\alpha \cap \beta$. It is clear that $\alpha \cap \beta$ is transitive and hence, by (3) it is an ordinal. Clearly $\alpha \cap \beta \subseteq \alpha$ and $\alpha \cap \beta \subseteq \beta$. If $\alpha \cap \beta = \beta$ then $\beta \subseteq \alpha$. By assumption we then have $\beta \subsetneq \alpha$ which by (4) implies $\beta \in \alpha$. Similarly $\alpha \cap \beta = \alpha$ implies $\alpha \in \beta$. We can therefore assume that $\alpha \cap \beta \subsetneq \alpha$ and $\alpha \cap \beta \subsetneq \beta$. By (4) we have $\alpha \cap \beta \in \alpha$ and $\alpha \cap \beta \in \beta$. This implies that $\alpha \cap \beta \in \alpha \cap \beta$ which contradicts (1). $\square$

THEOREM 66. *If α is an ordinal then $\alpha^+ := \alpha \cup \{\alpha\}$ is an ordinal and there is no ordinal β with $\alpha < \beta < \alpha^+$.*

If A is a set of ordinals then $\operatorname{lub}(A) := \bigcup A$ *is an ordinal which is the* **least upper bound of** *A, i.e., (i) for all $\alpha \in A$ we have that $\alpha \leq \operatorname{lub}(A)$; (ii) if β is an ordinal with $\alpha \leq \beta$ for all $\alpha \in A$ then* $\operatorname{lub}(A) \leq \beta$.

PROOF. Clearly α^+ is an ordinal. The fact that there is no ordinal β with $\alpha < \beta < \alpha^+$ follows from (1) of the previous lemma.

For the second part of the statement, $\bigcup A$ is transitive since if $\alpha \in \beta \in \bigcup A$ then by the definition of $\bigcup$ we have that $\alpha \in \bigcup A$. By (2) of the lemma it consists of ordinals. Hence (b), (c), (d) follow either immediately or by (5) of the lemma. The fact that it is well ordered it follows from the next lemma (Lemma 67).

Finally, since $\alpha \in A$ implies $\alpha \subseteq \mathrm{lub}(A)$ by (4) of the previous lemma we have that $\alpha \leq \mathrm{lub}(A)$. Moreover, if β is an ordinal with $\alpha \leq \beta$ for all $\alpha \in A$ then $\alpha \subseteq \beta$ for all $\alpha \in A$. Hence $\bigcup A \subseteq \beta$, i.e., $\mathrm{lub}(A) \leq \beta$. □

LEMMA 67. *If* COLL $\neq \emptyset$ *is a collection of ordinals then* COLL *has a least element.*

PROOF. Similar to Theorem 60. □

Unfortunately it does not follow from the axioms Z+(AC) that ORD are representatives of every the $\simeq_{\mathrm{ISO}}$ classes of WELLORD since in the Zermelo universe the only ordinals are

$$0, 1, 2, \ldots, \omega, \omega^+, \omega^{++}, \ldots$$

The *axiom of replacement* which we will now add will allow us to form the set

$$\{0, 1, 2, \ldots, \omega, \omega^+, \omega^{++}, \ldots\},$$

and prove that ORD extends far enough to contain an isomorphic member to each wellordering.

14. Class-functions and axiom of replacement

By a **class** we mean a collection COLL of sets which satisfy a fixed definable property. Formally, COLL is a class if there is a formula $\psi(x, x_1, \ldots, x_n)$ in the language of set theory and sets $a_1, \ldots, a_n$ so that $x \in \mathrm{COLL} \iff \psi(x, \bar{a})$. Every set a is a class since $a = \{x \mid x \in a\}$ but there are classes which are not sets, e.g.:

$$V, \quad \mathrm{WELLORD}, \quad \text{and} \quad \mathrm{ORD}.$$

We have already seen that V is not a set. Notice that if WELLORD was a set then by the subset axiom ORD would also be a set. Hence it suffices to show that ORD is not a set. But if ORD was a set then $\alpha := \bigcup \mathrm{ORD} = \mathrm{ORD}$ is an ordinal by Theorem 66. By the same theorem we then have that α^+ is also an ordinal and therefore $\alpha^+ < \alpha$, a contradiction with (1) of the Lemma 65. If a class COLL is not a set then we say that it is a **proper class**. Given classes A, B we define

$$\mathrm{A} \cap \mathrm{B}, \quad \mathrm{A} \cup \mathrm{B}, \quad \mathrm{A} \setminus \mathrm{B}, \quad \mathrm{A} \times \mathrm{B}$$

in the obvious way. A **class relation** is a subclass of some product class $\mathrm{A} \times \mathrm{B}$.

DEFINITION 68. A **class function** if any relation F with

$$(x, y), (x, z) \in \mathrm{F} \implies y = z.$$

Any function is a class function but we also have proper class functions such as

$$x \mapsto \mathcal{P}(x), \quad x \mapsto x^+.$$

Given a class function F the classes

$$\mathrm{dom}(\mathrm{F}), \quad \mathrm{rng}(\mathrm{F}), \quad \mathrm{F}''\mathrm{C}$$

for all $\mathrm{C} \subseteq \mathrm{dom}(\mathrm{F})$. We write $\mathrm{F}\colon \mathrm{A} \to \mathrm{B}$ if $\mathrm{dom}(\mathrm{F}) = \mathrm{A}$ and $\mathrm{rng}(\mathrm{F}) = \mathrm{B}$.

The *axiom of replacement* states that for $\mathrm{F}\colon \mathrm{A} \to \mathrm{B}$ is a class function and $\mathrm{C} \subseteq \mathrm{dom}(\mathrm{F})$, happens to be a set then so is its image $\mathrm{F}''\mathrm{C}$. Formally we have:

Axiom of Replacement. (Scheme: one axiom for every formula $\varphi(x, y, \bar{x})$)

Replacement for $\varphi(x, y, \bar{x})$:

$$\forall \bar{x}\Big(\big(\forall x, y, z \varphi(x, y, \bar{x}) = \varphi(x, z, \bar{x}) \implies y = z\big) \implies \forall w \exists v \forall t \big(t \in v \iff \exists x (x \in w \wedge \varphi(x, t, \bar{x}))\big)\Big)$$

Recall Zermelo theory alone cannot show that sets such as the Zermelo universe V_Z exist. However, the following assignment is a class function on domain ω:

$$n \mapsto F(n) = \overbrace{\mathcal{P}(\cdots \mathcal{P}(}^{n} V_\omega) \cdots),$$

since $F(n) = x$ if and only if "*$n \in \omega$, and v is a function on domain $n+1$, and $v(0) = V_\omega$, and $v(n) = x$, and for all $m < n$ we have that $v(m+1) = \mathcal{P}(v(m))$*".

By axiom of replacement we have that

$$T_{\omega+\omega} := \{V_\omega, \mathcal{P}(V_\omega), \mathcal{P}(\mathcal{P}(V_\omega)), \cdots\}$$

is a set, and so is

$$V_Z := \bigcup T_{\omega+\omega}$$

Similarly we may formally define $\omega + \omega$ as the union of the set

$$\{\omega, \omega^+, \omega^{++}, \ldots\}$$

which also exists by replacement.

15. Ordinals II

From now on, we will always view an ordinal α as a wellordering $(\alpha, <) := (\alpha, \in)$ In Section 13 we developed some properties of ordinals. In particular,

$$\omega + \omega := \mathrm{lub}(\{\omega, \omega^+, \ldots\})$$

which exists by axiom of replacement, is an ordinal. In fact we can now show that ORD contains an element in every isomorphism class of WELLORD.

THEOREM 69. *Every wellordered set is isomorphic to a unique ordinal.*

PROOF. Uniqueness follows from the properties we developed in section 13, since if $\alpha \neq \beta$ then $\alpha < \beta$ or $\beta < \alpha$, and no ordinal is isomorphic to an initial segment of it.

Let $\mathcal{W} = (W, <)$ be a wellordering. Let $(x, \alpha) \in \mathrm{F}$ if and only if

$$x \in W \text{ and } (W_x, <) \simeq_{\text{iso}} (\alpha, \in) \text{ and } \alpha \in \mathrm{ORD}$$

Clearly F is a class function and $\mathrm{dom}(\mathrm{F}) \subseteq W$. Thus $\mathrm{dom}(\mathrm{F}) \cap W$ is a set.

Claim. $\mathrm{dom}(\mathrm{F}) = W$.

This is proved by an induction on $\mathcal{W}$: if $W_x \subseteq \mathrm{dom}(\mathrm{F})$ for some $x \in W$ then by axiom of replacement we have that $A_x := \mathrm{F}''W_x$ is a set. Notice that α_x is wellordered by $\in$ since it is a subset of the ordinal $(\bigcup \alpha_x)^+$. It is also transitive: if $\alpha \in \beta \in \alpha_x$ then since $\alpha_x \simeq_{iso} (W_x, <)$ there is $y < x$ with $\beta \simeq_{iso} (W_y, <)$, and similarly $\alpha \simeq_{iso} (W_z, <)$ for some $z < y$. Hence, $\alpha = \mathrm{F}(z)$ for some $z < x$ and therefore $\alpha \in \alpha_x$. It follows that $\alpha_x \in \mathrm{ORD}$. It is easy to see that $(W_x, <) \simeq_{iso} \alpha_x$ and therefore $\mathrm{F}(x) = \alpha_x$, i.e., $x \in \mathrm{dom}(\mathrm{F})$.

Let now $\gamma := F''W$. As before we have: γ is an ordinal and $(W, <) \simeq \gamma$. □

As in the case of wellorderings we have the following scheme of theorems

THEOREM 70 (scheme). *Let* $\mathrm{A} \subseteq \mathrm{ORD}$ *be a non-empty class of ordinals. Then:*

$$\exists \alpha \in \mathrm{A}\ \forall \beta \in \mathrm{A}\ \alpha \leq \beta.$$

PROOF. Exercise □

We also have the associated "definition by induction" schemes. Here is a simple form of this that is analogous to Theorem 10.

THEOREM 71. *Let* $\mathbb{G} \colon V \times \mathrm{ORD} \to V$ *be a class function. Then there is a unique class function* $\mathrm{F} \colon \mathrm{ORD} \to V$ *so that* $\mathrm{F}(\alpha) = \mathrm{G}(\mathrm{F}{\restriction}\alpha, \alpha)$.

PROOF. Similar to Theorem 10. □

Of course, for many constructions we may want to produce a class function $\mathrm{F} \colon \mathrm{ORD} \times V \to V$. In that case there are theorem schemes analogous to Theorem 18 which let as perform these constructions.

16. Ordinal Arithmetic

A **sequence of ordinals** (α_ξ) is either a function $f \colon \lambda \to \mathrm{ORD}$ where λ is an ordinal or class function $\mathrm{F} \colon \mathrm{ORD} \to \mathrm{ORD}$, e.g. defined by Theorem 71. We denote the sequence by $(\alpha_\xi)_{\xi < \lambda}$ or by $(\alpha_\xi)_{\xi \in \mathrm{ORD}}$ respectively, if we want to specify its length. The sequence (α_ξ) is **increasing** if for every $\xi < \eta$ in the domain of (α_ξ) we have that $\alpha_\xi < \alpha_\eta$. If λ is some limit ordinal with α_ξ defined for all $\xi < \lambda$ then the **limit of** (α_ξ) is simply the ordinal:

$$\lim_{\xi < \lambda} \alpha_\xi := \mathrm{lub}\{\alpha_\xi \mid \xi < \lambda\}.$$

For example, $\lim_{n<\omega} n^n = \omega$ and $\lim_{n<\omega} \omega + n = \omega + \omega$, where $\omega + n := (\omega + n - 1)^+$.

DEFINITION 72. A sequence (α_ξ) is **normal** if it is increasing and continuous, i.e., for every limit ordinal λ in the domain of (α_ξ) we have that:

$$\lim_{\xi<\lambda} \alpha_\xi = \alpha_\lambda.$$

Using Theorem 71 we may define arithmetic operations on ORD.

Addition on ORD**.** Using induction on β we define $\alpha + \beta$ for every fixed α:

$$\alpha + 0 := \alpha, \quad \alpha + \beta^+ := (\alpha + \beta)^+, \quad \alpha + \lambda := \lim_{\xi<\lambda} \alpha + \xi,$$

where λ above is a limit ordinal. The operation $\alpha + \beta$ simply takes α and "glues" at the end of it a copy of β

Notice that $\beta \mapsto \alpha + \beta$ is a normal sequence but $\alpha \mapsto \alpha + \beta$ is neither increasing nor continuous since

$$1 + \omega = 2 + \omega, \quad \lim_{n<\omega} n + \omega = \lim_{n<\omega} \omega + \omega = \omega \neq \omega + \omega.$$

Exercise. Show that every ordinal α can be uniquely written in form $\lambda + n$ where λ is a limit ordinal (or 0) and $n \in \omega$.

Multiplication on ORD**.** Using induction on β we define $\alpha \cdot \beta$ for every fixed α:

$$\alpha \cdot 0 := 0, \quad \alpha \cdot \beta^+ := (\alpha \cdot \beta) + \alpha, \quad \alpha \cdot \lambda := \lim_{\xi<\lambda} \alpha \cdot \xi,$$

where λ above is a limit ordinal. The operation $\alpha \cdot \beta$ may be read α β-many times, for the obvious reasons. Similarly, $\beta \mapsto \alpha \cdot \beta$ is a normal sequence but $\alpha \mapsto \alpha \cdot \beta$ is neither increasing nor continuous.

Exponentiation on ORD**.** Using induction on β we define α^β for every fixed α:

$$\alpha^0 := 1, \quad \alpha^{\beta^+} := (\alpha^\beta) \cdot \alpha, \quad \alpha^\lambda := \lim_{\xi<\lambda} \alpha^\xi,$$

where λ above is a limit ordinal. Similarly, $\beta \mapsto \alpha^\beta$ is a normal sequence (when $\alpha \neq 0$) but $\alpha \mapsto \alpha^\beta$ is neither increasing nor continuous.

A mental picture to have in mind when it comes to exponentiation is the following: let $\mathrm{Fin}(\alpha^\beta)$ be the collection of all functions from β to α which have the property that they are pointwise equal to 0 except in finitely many elements of β. We may order $\mathrm{Fin}(\alpha^\beta)$ with the anti-lexicographic order: $f \prec g$ iff $f \neq g$ and if $\xi < \beta$ is the largest ordinal with $f(\xi) \neq g(\xi)$ then $f(\xi) < g(\xi)$. Then we claim that

$$(\mathrm{Fin}(\alpha^\beta), \prec) \simeq_{\mathrm{iso}} (\alpha^\beta, \in)$$

This can be show easily by induction.

We have he following picture at the very beginning of the ORD:

$$0, 1, 2, \ldots, \omega, \omega + 1, \ldots, \omega + \omega, \omega \cdot 3, \omega \cdot 4, \ldots, \omega \cdot \omega, \omega^3, \ldots, \omega^\omega, \cdots, \omega^{\omega^\omega}, \cdots, \epsilon_0, \cdots\cdots$$

where ϵ_0 is simply the limit of $\{\omega, \omega^\omega, \omega^{\omega^\omega}, \ldots\}$. Notice that all these ordinals are countable and we are far from reaching the first uncountable ordinal denoted by ω_1.

To define ω_1 consider the set $\{A \subseteq \mathbb{Q} \mid (A, <_{\mathbb{Q}} \restriction A) \text{ is a wellordering}\}$. Let α_A be the unique ordinal associated to A. Then $\omega_1 := \{\alpha_A \mid A \text{ as above }\}$ is a set by axiom of replacement. Since $\mathbb{Q}$ embeds every countable linear order (it is the Fraïssé limit of linear orderings) we have that ω_1 is the set of all countable ordinals. It is not difficult to see that it is itself an ordinal. Moreover it is not countable because otherwise it would embed in $(\mathbb{Q}, <)$ and this would imply that $\omega_1 \in \omega_1$, a contradiction.

Not only has ω_1 different order type than all previous ordinals but also different cardinality. Is there an ordinal which has strictly larger cardinality than ω_1? The answer is yes as we will see in the next section.

17. Cardinals

By axiom of choice we have that every set x is wellorderable. By axiom of reflection (see Theorem 69) we then have $x \approx \alpha$ for some ordinal α. The ordinal

$$|x| := \min\{\alpha \in \mathrm{ORD} \mid x \approx \alpha\} \tag{3}$$

is called the **cardinality of** x. It is clear that $x \approx y$ if and only if $|x| = |y|$ and that

$$(x \lessapprox y) \wedge (x \not\approx y) \text{ if and only if } |x| < |y|.$$

Not every ordinal is of the form $|x|$ for some x. A **cardinal** or an **initial ordinal** is any ordinal of the form $|x|$, or equivalently, any $\alpha \in \mathrm{ORD}$ so that for all $\beta < \alpha$ we have $|\beta| < |\alpha|$. For example every $n \in \omega$ is a cardinal. Similarly ω and ω_1 are cardinals and, when viewed as such, they are denoted by

$$\aleph_0 \text{ and } \aleph_1.$$

We usually denote cardinals by the letters $\kappa, \mu, \nu, \cdots$

Theorem 73 (Hartogs). *For any ordinal α there is a least cardinal κ with $\alpha < \kappa$.*

Proof. We may assume that α is infinite. Let $\mathcal{WO}(\alpha)$ be the set of all relations $\prec$ which are wellorderings with $\mathrm{dom}(\prec) \subseteq \alpha$. For each $\prec$ as above let $\alpha_\prec$ be the unique ordinal isomorphic to $\prec$ and let

$$\kappa := \mathrm{lub}\{\alpha_\prec \mid \prec \in \mathcal{WO}(\alpha)\}.$$

It is clear that $\alpha < \kappa$ since $\alpha + 1$ is easily shown to be the order type of some element of $\mathcal{WO}(\alpha)$. To see that κ is an initial ordinal notice that if $\kappa \approx \beta$ for some $\beta < \kappa$ then by minimality of κ we have that $\beta \simeq_{\mathrm{iso}} \prec$ for some $\prec \in \mathcal{WO}(\alpha)$. But then composing the implicit bijections we can realize $(\kappa, \in)$ as an element of $\mathcal{WO}(\alpha)$ and by modifying this element a bit we can similarly realize $\kappa + 1$. By the choice of κ we then have $\kappa \in \kappa$, a contradiction.

Notice finally that κ is the least such cardinal since by minimality of the least upper bound, for every $\beta < \kappa$ is isomorphic (as an ordering) to some $\prec$ in $\mathcal{WO}(\alpha)$. $\square$

Warning. In the context of cardinals we use the notation κ^+ for a different operation than $\kappa \mapsto \kappa+1$. Namely, if κ is a cardinal then let

$$\kappa^+ := \mathrm{lub}\{\alpha_\prec \mid \prec \in \mathcal{WO}(\kappa)\} = \text{ the least cardinal } \mu \text{ with } \mu > \kappa.$$

THEOREM 74. *If X is a set of cardinals then* $\mathrm{lub}(X)$ *is also a cardinal.*

PROOF. Let $\alpha := \mathrm{lub}(X)$. We may assume without loss of generality that X has no max element, otherwise $\alpha \in X$ and we are done. If α is not a cardinal then let $\kappa < \alpha$ with $\kappa \approx \alpha$. Let now $\mu \in X$ with $\kappa < \mu < \alpha$. By CSB we have $\kappa = \mu$, a contradiction. $\square$

We define now inductively the ORD-length sequence $(\aleph_\alpha)$ of infinite cardinals by:

$$\aleph_0 := \omega, \quad \aleph_{\alpha+1} = (\aleph_\alpha)^+, \quad \aleph_\lambda := \lim_{\xi<\lambda} \aleph_\xi,$$

where λ is a limit ordinal. By the results above, every $\aleph_\alpha$ is a cardinal. Moreover it is easy to see that $\mathrm{CARD} = \omega \cup \{\aleph_\alpha \mid \alpha \in \mathrm{ORD}\}$. It is clear from the last Theorem that $(\aleph_\alpha)$ is a normal sequence. We will later see that $\alpha \mapsto \aleph_\alpha$ has a lot of fixed points. The first one being:

$$\aleph_{\aleph_{\aleph_{\dots}}} = \mathrm{lub}\{\aleph_0, \aleph_{\aleph_0},, \aleph_{\aleph_{\aleph_0}} \cdots\}$$

where the last set is the obvious countable set.

We define the operations of cardinal arithmetic as follows and we warn again the reader that these differ from the operations in ordinal arithmetic. If κ, μ are ordinals fix X, Y two disjoint sets with $\kappa \approx X$ and $\mu \approx Y$ and define:

$$\kappa + \mu := |X \cup Y|, \quad \kappa \cdot \mu := |X \times Y|, \quad \kappa^\mu := |X^Y|.$$

It is clear that $2^\kappa = |\mathcal{P}(\kappa)|$ and therefore Cantor's theorem says that for all κ:

$$\kappa < 2^\kappa.$$

This implies that $\kappa^+ \leq 2^\kappa$. The generalized continuum hypothesis conjectures that this is actually an equality. We will see in the next quarter that this cannot be decided from the axioms of ZFC.

There are many identities one can prove about cardinal arithmetic. Many of them however rely on the axiom of choice. The following theorem shows that infinite cardinal arithmetic for $+, \cdot$ is much simpler than finite cardinal arithmetic.

THEOREM 75 (The fundamental theorem of cardinal arithmetic.). *Let κ, μ be infinite cardinals then $\kappa + \mu = \kappa \cdot \mu = \max\{\kappa, \mu\}$.*

PROOF. Notice that it is enough to show that $\kappa^2 = \kappa$ since then, if $\kappa \leq \mu$, we have:

$$\mu \leq \kappa + \mu \leq \kappa \cdot \mu \leq \mu \cdot \mu = \mu.$$

So $\kappa + \mu = \kappa \cdot \mu = \mu = \max\{\kappa, \mu\}$. Hence the theorem follows from the next lemma. $\square$

THEOREM 76. *For every α we have that $\aleph_\alpha \times \aleph_\alpha \approx \aleph_\alpha$.*

PROOF. Consider the following (class) ordering $\prec$ on ORD $\times$ ORD:

$$\begin{aligned}(\alpha,\beta) \prec (\gamma,\delta) \iff &(\max\{\alpha,\beta\} < \max\{\gamma,\delta\}) \text{ or}\\ &(\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \text{ and } \alpha < \gamma) \text{ or}\\ &(\max\{\alpha,\beta\} = \max\{\gamma,\delta\} \text{ and } \alpha = \gamma \text{ and } \beta < \gamma).\end{aligned}$$

We will show that,for each α, the restriction of $\prec$ on $\aleph_\alpha \times \aleph_\alpha$ is a well ordering that is isomorphic to $(\aleph_\alpha, \in)$. It is easy to see that $\prec$ is a wellordering by computing the minimum of any given set $X \subseteq$ ORD $\times$ ORD coordinate-wise. Hence, $(\aleph_\alpha \times \aleph_\alpha, \prec)$ is a wellordering, as a subset of a wellordering. Let now λ_α be the unique ordinal which is isomorphic to $(\aleph_\alpha \times \aleph_\alpha, \prec)$.

CLAIM. *For all α we have that $\lambda_\alpha = \aleph_\alpha$.*

PROOF. We do this by induction. For $\alpha = 0$ it is easy to see this. Assume that it holds for all $\beta < \alpha$ and we show it for α. It is clear that $\lambda_\alpha \geq |\aleph_\alpha \times \aleph_\alpha| \geq \aleph_\alpha$. It therefore suffices to prove that $\lambda_\alpha \leq \aleph_\alpha$. It suffices to show that for all $\xi < \lambda_\alpha$ we have that $\xi < \aleph_\alpha$. Equivalently, if $f\colon \aleph_\alpha \times \aleph_\alpha \to \lambda_\alpha$ is the unique order isomorphism, it suffices to show that for any fixed (γ, δ) we have $f(\gamma, \delta) < \aleph_\alpha$. In fact it suffices to show that $|f(\gamma, \delta)| < \aleph_\alpha$ since $\aleph_\alpha$ is initial. But

$$|f(\gamma,\delta)| = |\{(\gamma',\delta') \mid (\gamma',\delta') \prec (\gamma,\delta)\}| \leq |(\varepsilon+1) \times (\varepsilon+1)|,$$

where $\varepsilon = \max\{\gamma, \delta\}$. But $\varepsilon + 1 < \aleph_\alpha$, so $\varepsilon + 1 = \aleph_\beta$ for some $\beta < \alpha$. By induction hypothesis we have that

$$|(\varepsilon+1) \times (\varepsilon+1)| = |(\varepsilon+1)| \times |(\varepsilon+1)| = |\aleph_\beta| \times |\aleph_\beta| = |\aleph_\beta \times \aleph_\beta| = \aleph_\beta < \aleph_\alpha.$$

$\square$

$\square$

COROLLARY 77. *If κ, μ are infinite cardinals with $2 \leq \kappa \leq \mu$, then $\kappa^\mu = 2^\mu$.*

PROOF. $2^\mu \leq \kappa^\mu \leq (2^\kappa)^\mu = 2^{\kappa\cdot\mu} = 2^\mu$. We leave to the reader to confirm the second equality from the end. $\square$

We point out at this point that one may define a notion of cardinality for each set without the use of axiom of choice (see next section). In that case not every set of the form $|x|$ is an ordinal and what we call here the "fundamental theorem of arithmetic" fails in general. Even more, without axiom of choice cardinals are not linearly ordered since there may be sets which are $\lessapprox$-incomparable. One may consult Tarski's theory of cardinal algebras and books in choiceless set theory for more on this.

18. Axiom of regularity and wellfounded class relations

Mow that we can use ORD as a formal notion of "stages" we can define by induction the cumulative hierarchy $(V_\alpha)_{\alpha\in\mathrm{ORD}}$ as follows:

$$V_0 := \emptyset, \quad V_{\alpha+1} := \mathcal{P}(V_\alpha), \quad V_\lambda := \bigcup_{\xi<\lambda} V_\xi,$$

where λ is a limit ordinal. It is clear that $V_\alpha \subseteq V_\beta$ if an $\alpha \leq \beta$ and if additionally $\alpha \neq \beta$ then $V_\alpha \in V_\beta$. We would like to be able to deduce that

$$V := \bigcup_{\alpha<\mathrm{ORD}} V_\alpha, \tag{4}$$

however this statement is not provable from the axioms. Indeed, using the current axioms in can be shown to be consistent that there are sets u with the property $u = \{u\}$ and such a set is clearly not in the above union: if it was, pick α the smallest ordinal with $u \in V_\alpha$ to get a contradiction with $u \in u$. We may now complete the collection of axioms which constitute ZFC by adding the following axiom that is equivalent to the above.

Axiom of Regularity (also know as **Axiom of Foundation**)

$$\forall X \text{ if } X \neq \emptyset\ \exists x \in X\ \forall y \in x\ (y \notin X).$$

It is clear that $V := \bigcup_{\alpha<\mathrm{ORD}} V_\alpha$ implies the axiom of regularity: if $X \in V$ then pick the least α so that $X \cap V_\alpha \neq \emptyset$. Clearly α is not a limit ordinal and if $X \neq \emptyset$ then $\alpha = \beta + 1$ is a successor. Pick any $x \in X \cap V_\alpha$ and notice that if $y \in x$ then $y \in V_\beta$ which, by minimality of α, implies $y \notin X$. In order to prove the converse we need to introduce the concept of induction on certain wellfounded class relations.

A class relation E is **wellfounded** if every set $X \subseteq \mathrm{dom}(E)$ has a minimal element. That is, there is $x \in X$ so that for all y with $y\mathrm{E}x$ we have $x \notin X$. By AC this is equivalent with saying that there is no infinite sequence $x_0, x_1, \ldots \in \mathrm{dom}(\mathrm{E})$ with

$$\cdots Ex_2Ex_1Ex_0$$

We say that E is **set-like** if for all $x \in \mathrm{dom}(\mathrm{E})$ we have that $\mathrm{E}_x := \{y \mid yEx\}$ is a set. It is clear that $\in$ is a set-like class relation. The axiom of regularity is equivalent to saying that $\in$ is also wellfounded. As usual we have the associated induction results

THEOREM 78. *If* E *is wellfounded and set-like class relation then for every subclass* $\mathrm{X} \subseteq \mathrm{E}$ *we have that*

$$\big(\forall x \in \mathrm{dom}(\mathrm{E})(\mathrm{E}_x \subseteq \mathrm{X} \implies x \in \mathrm{X})\big) \implies (\mathrm{X} = \mathrm{dom}(E)).$$

PROOF. We leave this as an exercise to the reader. Hint: use next lemma. □

LEMMA 79. *If* E *is wellfounded and set-like class relation then every subclass* X *of* $\mathrm{dom}(E)$ *has a least element.*

PROOF. Pick any $x \in \mathrm{dom}(E)$. If $\mathrm{E}_x = \emptyset$ then x is minimal. If not, use the fact that E_x is a set to find a minimal y for E_x. Then y is minimal for dom(E) as well. $\square$

We may now see why axiom of regularity implies $V := \bigcup_{\alpha<\mathrm{ORD}} V_\alpha$. Indeed it is enough to prove that for all x, if $\in_x := \{y \mid y \in x\} \subseteq \bigcup_{\alpha<\mathrm{ORD}} V_\alpha$, then $x \in \bigcup_{\alpha<\mathrm{ORD}} V_\alpha$. For each $y \in \in_x$ let $f(y)$ be the least ordinal with $y \in V_\alpha$. But then, $\mathrm{rng}(f) \subseteq \mathrm{ORD}$ is a set (by replacement) and we therefore may find an ordinal β with $\in_x \subseteq V_\beta$. This implies that $x \in \in V_{\beta+1}$, and we are done.

Exercise. We leave it to the reader to define and prove a "definition by induction" theorem for wellfounded set-like class relations.

DEFINITION 80. For every set x we define the **rank** of x, denoted by $\mathrm{rank}(x)$, to be the least ordinal α so that $x \subseteq V_\alpha$. Equivalently, the least α so that $x \in V_{\alpha+1}$. Under the axiom of regularity this is well defined for all x.

ZF and **ZFC**. From now on we write ZF for the collection of all axioms in Z together with axiom of replacement and axiom of regularity. When we add to this collection AC then we write ZFC.

We close this section by discussing the, so called, Scott's trick which allows as to define in ZF an notion of cardinality for sets. By a **cardinal assignment** we mean any class function $x \mapsto |x|_*$ from V to V with the property that

$$x \approx y \iff |x|_* = |y|_*.$$

Given such a cardinal assignment we say that a set y is a cardinal if $y = |x|_*$ for some x. Under AC we showed how one may define a cardinal assignment, namely the one we denote above by $x \mapsto |x|$. The obvious but naive attempt to define a cardinal assignment without any use of AC would be to let

$$|x|_* := \{y \mid y \approx x\}$$

This "assignment" is clearly invariant under $\approx$ but it fails to be a class function from V to V since $|x|_*$ is not set if $x \neq \emptyset$. In the context of ZF where we have the axiom of regularity we instead define

$$|x|_{\mathrm{Scott}} := \{y \in V_{\alpha(x)} \mid y \approx x\},$$

where $\alpha(x)$ is the least ordinal α for which there is a set $y \in V_\alpha$ so that $y \approx x$.

Remark. If x is well-orderable, then $|x|$ can be defined as is section 17 and it is easy to see that $|x| \in |x|_{\mathrm{Scott}}$. Hence, under axiom of choice the assignments $x \mapsto |x|_{\mathrm{Scott}}$ and $x \mapsto |x|$ are isomorphic. In the absence of AC there are cardinals $|x|_{\mathrm{Scott}}$ which cannot be identified to alephs and the cardinal arithmetic we defined in Section 17 is much more complicated. In Section 1 we will see some examples.

CHAPTER 2

Selected topics

1. Descriptive ergodic theory and cardinality without AC

There is an interesting relationship between ergodic theory and the theory of cardinals in ZF+¬AC which can be made precise within the framework of invariant descriptive set theory. We will first describe the framework and use it to differentiate between different orbit equivalence relations E^{Γ}_X of continuous group actions $\Gamma \curvearrowright X$. We will then connect this "complexity theory" of orbit equivalence relations with the behavior of cardinals in a very important model of ZF where AC fails, known as

$$\mathrm{L}(\mathbb{R}).$$

A curious property of $\mathrm{L}(\mathbb{R})$ is that it does not contain any injection from ω_1 to $\mathbb{R}$. As a consequence, within $L(\mathbb{R})$ we have that the cardinals $\aleph_1$ and $2^{\aleph_0}$ are incomparable.

A **Polish space** X is any separable space which admits a complete metric d inducing the same topology. In Section 8 we defined the continuum as any of the:

$$2^{\omega} \approx n^{\omega} \approx \omega^{\omega} \approx \mathcal{P}(\omega) \approx \mathbb{R}.$$

All these spaces are uncountable Polish spaces under a natural topology. The topology in ω^{ω} is given by basic open sets of the form $U_s = \{(a_n) \in \omega^{\omega} \mid (a_n)_{\omega} \text{ extends } s\}$ where $s\colon n \to \omega$ for some n and the topology in $\mathcal{P}(\omega)$ is inherited from 2^{ω} under the natural identification. The above spaces are equinumerous via maps which are definable in ZF and therefore, even in models such as $L(\mathbb{R})$ we have that

$$|2^{\omega}| = |n^{\omega}| = |\omega^{\omega}| = |\mathcal{P}(\omega)| = |\mathbb{R}| = 2^{\aleph_0}.$$

In fact it follows by Theorem 81 that any uncountable Polish space has cardinality $2^{\aleph_0}$ in any model of ZF which satisfy countable choice such as in $L(\mathrm{R})$. We will employ methods from ergodic theory to construct sets which, in $L(\mathrm{R})$, have cardinalities different than any cardinality from $\{\aleph_{\alpha}\} \bigcup \{2^{\aleph_0}\}$.

A **Polish group** is a topological group G, i.e., multiplication and inversion are continuous, whose topology is Polish. We will consider continuous actions

$$G \curvearrowright X$$

of Polish groups G on Polish spaces X. These are continuous maps $G \times X \to X$ with $1x = x$ and $g(hx) = (g \cdot h)x$. We then say that X is a **Polish G-space**. Such an action induces always an equivalence relation E^G_X on X known as the **orbit equivalence relation** associated with the Polish G-space:

$$xE^G_X y \iff \exists g \in G \;\; gx = y.$$

Examples. Here are some natural examples of Polish G-spaces:

(1) Let $=$ be the orbit equivalence relation of the trivial group $\{1\}$ acting on any Polish space X;
(2) Consider the additive discrete group $\mathbb{Z}$ acting on the space $X := \{0,1\}^{\mathbb{Z}}$ by translation: $(k \cdot f)(n) = f(n-k)$ for all $f \in X$ and $k \in \mathbb{Z}$. We call this the **Bernoulli shift** of $\mathbb{Z}$ and let $E_X^{\mathbb{Z}}$ be the associated equivalence relation;
(3) Similarly for the free group F_2 in two generators acting on its Bernoulli shift $X := \{0,1\}^{F_2}$ by translation, $(\gamma \cdot f)(\delta) = f(\gamma^{-1}\delta)$ for all $f \in X$ and $\gamma \in F_2$, we let $E_X^{F_2}$ be the associated equivalence relation;
(4) Let $X = \mathrm{Graphs}(\omega)$ be the space of all symmetric and reflexive graph structures on domain ω which is clearly a Polish space as a closed subset of $2^{\omega\times\omega}$. The Polish group S_ω of all permutations of ω is a Polish group if endowed with the pointwise convergence topology and it acts continuously on X: every bijection $g\colon \omega \to \omega$ sends the graph (ω, R) to the new graph (ω, R^g), where $R^g(n,m) \iff R(g^{-1}(n), g^{-1}(m))$. Notice that the associated equivalence relation is the isomorphism relation $\simeq_{\mathrm{iso}}$ of graphs.

Assume now that your task is to classify all countable graphs up to isomorphism. One usually does that by finding enough "complete invariants." For example the assignments $(\omega, R) \mapsto \mathrm{maxdeg}((\omega, R))$ and $(\omega, R) \mapsto \mathrm{conn}((\omega, R))$ which map each element of $\mathrm{Graphs}(\omega)$ to its maximum degree and to its number of connected components, are both definable maps which are invariant under $\simeq_{\mathrm{iso}}$. However, the collection $\{\mathrm{maxdeg}(\cdot), \mathrm{conn}(\cdot)\}$ is not a *complete* set of invariants since there are non-isomorphic graphs with the same number of connected components and the same maximum degree. The obvious question is whether there is a *definable map*

$$f\colon \mathrm{Graphs}(\omega) \to Y$$

where Y is any Polish space so that $x \simeq_{\mathrm{iso}} y$ if and only if $x = y$. In the context of ZF definable would mean definable without axiom of choice. Here are two more notions of *definable* which are used in ergodic theory and descriptive set theory.

A subset N of a Polish space X is **nowhere dense** if the complement of its closure $(\overline{N})^c$ is dense in X. It is **meager** if $N = \bigcup_k N_k$ and each k is nowhere dense. By the Baire category theorem (see real analysis textbook). A Polish space X is never meager as a subset of itself, hence, meagerness can be thought of as a notion of (topological) smallness (see Proposition 22 for the formal definition of smallness) which is moreover closed under countable unions similar to the notion of measure 0 sets in measure theory. Let $\mathcal{B}(X)$ be the smallest family of subsets of X which contains all opens sets and which is closed under countable unions and countable intersections. A set is **Borel** if it is contained in $\mathcal{B}(X)$. Let $\mathcal{BP}(X)$ be the smallest family of subsets of X which contains all opens sets, all meager sets, and which is closed under countable unions and countable intersections. A set has the **Baire property** if it is contained in $\mathcal{BP}(X)$. A function $f\colon X \to Y$ between Polish spaces is **Borel** if $f^{-1}(U) \in \mathcal{B}(X)$ for every open set U and it is **Baire measurable** if

$f^{-1}(U) \in \mathcal{BP}(X)$ for every open set U. We view Borel sets and Borel maps as "*definable from the topology via countable operations*".

Assume now that $f\colon X \to Y$ is a map between Polish spaces. It follows that if f is Borel then f is in $L(\mathbb{R})$ and if f is in $L(\mathbb{R})$ then f is Baire-measurable[1]. Hence Borel maps and Baire-measurable maps can be seen as upper and lower estimates of the class of all maps in $L(\mathbb{R})$ between Polish spaces. The following theorem which we state without proof implies that the cardinality of every uncountable Polish space in $L(\mathbb{R})$ is $2^{\aleph_0}$

THEOREM 81. *If $X \approx Y$ are Polish spaces then $\approx$ is witnessed by a Borel map.*

Using the following anti-classification result regarding for the examples (2),(3),(4) above and the fact that all maps in $L(\mathbb{R})$ between Polish spaces are Baire-measurable, one may produce new cardinalities which are not in $\{\aleph_\alpha\} \bigcup \{2^{\aleph_0}\}$ such as:

$$|\{0,1\}^{\mathbb{Z}}/E_X^{\mathbb{Z}}|, \quad |\{0,1\}^{F_2}/E_X^{F_2}|, \quad |\mathrm{Graphs}(\omega)/\simeq_{\mathrm{iso}}|$$

Notice, for example that if $\{0,1\}^{\mathbb{Z}}/E_X^{\mathbb{Z}} \approx 2^\omega$ was true in $L(\mathbb{R})$ then this would contradict the following theorem.

THEOREM 82. *Let X be either $\{0,1\}^{\mathbb{Z}}$ or $\{0,1\}^{F_2}$ or $\mathrm{Graphs}(\omega)$ and let E be either $E_X^{\mathbb{Z}}$ or $E_X^{F_2}$ or $\simeq_{\mathrm{iso}}$ respectively. Then there is no Baire-measurable map $f\colon X \to 2^\omega$ so that $xEy \iff f(x) = f(y)$.*

Before we prove the theorem lets point out that every two uncountable Polish spaces X and Y are equinumerous by a Borel map $f\colon X \to Y$. Hence we can replace $(2^\omega, =)$ with $(Y, =)$ in the above theorem where Y is any Polish space. We will need the following lemma which can be thought of as a 0-1 law.

LEMMA 83. *Let X be a Polish G-space which has some dense orbit. If $A \subseteq X$ is Baire-measurable and G-invariant, then either A or A^c is meager.*

PROOF. Notice first that for every non-empty open $U, U' \subseteq X$ there is $g \in G$ with $gU \cap U' \neq \emptyset$. To see this, if $[x]$ is dense then $gx \in U$ and $g'x \in U'$ for some $g, g' \in G$. Hence $g'g^{-1}U \cap U'$ contains $g'x$.

Assume that there is an invariant Baire measurable set A so that both A and $B = A^c$ are non-meager. By Baire-measurability we can find U_A and U_B so that A is comeager in U_A and B is comeager in U_B. Let $g \in G$ with $gU_A \cap U_B \neq \emptyset$ and let U be this non-empty and open common intersection. Since gA is comeager in gU_A it is comeager in U as well. But $A = gA$ so A is comeager in U which contradicts the assumption that its complement B is comeager in U. □

LEMMA 84. *If X is a Polish G-space with a dense orbit and $f\colon X \to 2^\omega$ is a Baire measurable map with $xE_X^G y \implies f(x) = f(y)$ then there is some $\alpha \in 2^\omega$ so that $f^{-1}(\alpha)$ is comeager.*

[1]Here we assume, as it is often the case with $L(\mathbb{R})$, axiom of determinacy (which follows by assuming that certain large cardinals exist)

PROOF. Assume that $f\colon X \to 2^\omega$ be a Borel reduction. For every $s \in 2^{<\omega}$ let N_S be the basic open set of 2^ω consisting of all sequences extending s.

Notice that $\{f^{-1}(N_{(0)}), f^{-1}(N_{(1)})\}$ forms a partition of invariant Borel subsets of X. By Lemma 84 one of them has to be comeager. Continuing inductively we build a sequence $\alpha \in 2^\omega$ so that for all $n > 0$ we have that $f^{-1}(N_{\alpha|n})$ is comeager. But then $C = \cap_n f^{-1}(N_{\alpha|n})$ is a comeager subset of X that is mapped under f to the singleton $\{\alpha\}$. $\square$

We may now finish the proof of Theorem 82 for the case $\mathbb{Z} \curvearrowright 2^{\mathbb{Z}}$. For the other cases see HW.

PROOF OF THEOREM 82 CASE: $\mathbb{Z} \curvearrowright 2^{\mathbb{Z}}$. It is easy to construct a dense orbit of $2^{\mathbb{Z}}$. Simply enumerate all possible finite sequences $s\colon n \to \omega$ and construct some $x \in 2^{\mathbb{Z}}$ which realizes every s somewhere, i.e., for every s as above there is $k \in \mathbb{Z}$ so that $s(i) = x(i+k)$ for all $i \leq n$. By lemma 84 there is some comeager set C of $2^{\mathbb{Z}}$, on which the map f is constant. But since the orbits are countable they are meager. Hence C has at least 2 distinct orbits which map to the same element of 2^ω. $\square$

2. Combinatorics on ω_1

We show that ω satisfies two very important properties which are reflections of some short of compactness. Namely:

- ω has the **tree property** (aka König's lemma): if T is a finitely branching tree of cardinality $\aleph_0$ then there is a branch of length ω.
- ω has the **Ramsey property** (i.e., $\omega \to (\omega)^2$): if we color the set of all two-sets $[\omega]^2 := \big\{\{a,b\}a, b \in \omega, a \neq b\big\}$ with finitely many colors then there is a set $A \subseteq \omega$ of size $\aleph_0$ so that $[A]^2$ is monchromatic.

The obvious question is if the analogous properties hold for ω_1. We will see here that both properties fail for ω_1. This difference between ω and ω_1 can be seen as a lack of "compactness" for ω_1. Many large cardinal axioms which are often added in ZF or ZFC are used to reproduce this ω-like compactness behaviour in larger than ω_1 uncountable cardinalities.

We adopt the following definition of a tree which permits trees of arbitrary long heights. A **tree** is a partial order $(T, <)$ with the property that for each $x \in T$ the set $\{y \in T \mid y < x\}$ is well ordered by $<$. The **height** $\mathrm{ht}(x,T)$ **of** x **in** T is the unique ordinal isomorphic to the above set. The α-th level $\mathrm{Lv}_\alpha(T)$ of T is the set

$$\{x \in T \mid \mathrm{ht}(x,T) = \alpha\}.$$

The **height** $\mathrm{ht}(T)$ **of** T is the least α with $\mathrm{Lv}_\alpha(T) = \emptyset$. A **subtree** T' **of** T is any subset of T that is $<$-downward closed. A **chain of** T is any subset C of T that is linearly ordered by inclusion. The set $\{y \in T \mid y < x\}$ is a chain for all $x \in T$. More generally if C is a chain then

$$C \subseteq \bigcup_{x \in C} \{y \in T \mid y \leq x\}$$

DEFINITION 85. Let κ be a cardinal. A **κ-Aronszajn tree** is any tree T with

(1) $\mathrm{ht}(T) = \kappa$;
(2) $\mathrm{Lv}_\alpha(T) < \kappa$ for all α;
(3) T has no chain of size κ.

We say that κ has the **tree property** if there exists no κ-Aronszajn tree, i.e., any tree which satisfies (1), (2), fails (3).

It is clear that König's lemma shows that ω has the tree property. However:

THEOREM 86. *There exists an ω_1-Aronszajn tree.*

PROOF. Consider the set

$$\mathrm{Inj}(\omega^{<\omega_1}) := \{s \colon \alpha \to \omega \mid \alpha < \omega_1, s \text{ is injective}\}.$$

Then $(\mathrm{Inj}(\omega^{<\omega_1}), <)$ is a tree, where $<$ is just the subset relation $\subset$. It clearly satisfies (1) and (3) for $\kappa = \omega_1$: since every $\alpha < \omega_1$ is countable there is $s \colon \alpha_\omega$ injective and it is clear that $\mathrm{ht}(s, \mathrm{Inj}(\omega^{<\omega_1})) = \alpha$; if there was a chain C of size $\aleph_1$ then the union $f := \bigcup C$ would be an injection from ω_1 to ω, contradicting that ω_1 is uncountable. However, T does not satisfy (2) even for $\alpha = \omega < \omega_1$. We will remedy this by constructing some "sparse" subtree T of $\mathrm{Inj}(\omega^{<\omega_1})$.

For $s, t \in \omega^\alpha \subseteq \omega^{<\omega_1}$ we write $s \sim_{\mathrm{fin}} t$ if $\{\xi < \alpha \mid s(\xi) \neq t(\xi)\}$ is finite. We will construct a sequence $(s_\alpha)_{\alpha<\omega_1}$ in $\mathrm{Inj}(\omega^{<\omega_1})$ with the following properties:

(i) $\mathrm{dom}(s_\alpha) = \alpha$;
(ii) if $\alpha < \beta$ then $s_\alpha \sim_{\mathrm{fin}} s_\beta{\restriction}\alpha$.
(iii) $\omega \setminus \mathrm{range}(s_\alpha)$ is infinite.

Assuming that we have defined this sequence we may finish the proof by setting $T := \{s \in \mathrm{Inj}(\omega^{<\omega_1}) \mid s \sim_{\mathrm{fin}} s_\alpha\}$ for some α. As above it is clear that T satisfies (1) and (3). It also satisfies (2) since for every α the $\sim_{\mathrm{fin}}$-equivalence class of any $t \in \omega^\alpha$ is countable. So we are left to construct $(s_\alpha)_{\alpha<\omega_1}$. We do this by induction.

Assume that s_α has been constructed so that it satisfies the above. Pick any $a \in \omega \setminus \mathrm{range}(s_\alpha)$, which exists by (iii) and set $s_{\alpha+1} := s_\alpha \cup \{(\alpha, a)\}$. Suppose now that $\lambda < \omega_1$ is a limit ordinal and assume that s_α has been defined for all $\alpha < \lambda$. Pick some increasing sequence α_n with $\alpha_n \to \lambda$. We would like to define s_λ as some sort of a union of the $(s_{\alpha_n})_{n\in\omega}$. Of course, this sequence does not cohere so we first define some injective $t_n \sim_{\mathrm{fin}} s_{\alpha_n}$ so that $t_{n+1}{\restriction}\alpha_n = t_n$ first. This can be done by an easy induction. Set $t := \bigcup_n t_n$. Then t is injective and if we where to set s_λ equal to t then we would have (i),(ii) above. However this choice of s_λ may not satisfy (iii). To fix this, define s_λ so that $s_\lambda(\alpha_n) = t(\alpha_{2n})$ and $s_\lambda(\alpha) = t(\alpha)$ otherwise. It is easy to check now that all (i),(ii),(iii) hold. $\square$

We point out that it is independent of ZFC whether $\aleph_2$ has the tree property. We will explore what happens at the "higher infinite" in the next section.

THEOREM 87. *There is a 2-coloring of $[\omega_1]^2$ for which there is no uncountable $A \subseteq \omega_1$ with $[A]^2$ monochromatic. That is $\omega_1 \not\to (\omega_1)^2$.*

PROOF. It suffices to show that $2^{\aleph_0} \not\to (\omega_1)^2$, since $2^{\aleph_0} \geq \aleph_1$.

Let $<_{\mathrm{R}}$ be the usual ordering of $\mathbb{R}$ and notice that there are no $<_{\mathrm{R}}$-increasing or $<_{\mathrm{R}}$-decreasing sequences $(a_\xi)_{\xi\in\omega_1}$ of reals of length ω_1. To see this notice that if (a_ξ) is strictly increasing then we can assign to each a_ξ so rational q_ξ with the property that $\xi \neq \xi' \implies q_\xi \neq q_{\xi'}$: let q_ξ be any rational between a_ξ and $a_{\xi+1}$. Hence any increasing sequence has countable length.

Since $\mathbb{R} \approx 2^{\aleph_0}$ we may transfer $<_{\mathrm{R}}$ on $2^{\aleph_0}$ via any bijection. So, on the ordinal $2^{\aleph_0}$ we have two orderings: the usual $<$ ordering and $<_{\mathrm{R}}$. Consider the coloring $f\colon [2^{\aleph_0}]^2 \to 2$ with $f(\{\alpha,\beta\}) = 0$, if $<$ and $<_{\mathrm{R}}$ agree on $\{\alpha,\beta\}$; and $f(\{\alpha,\beta\}) = 1$ otherwise. But then, there is no f-monochromatic subset of $2^{\aleph_0}$ because this would give either a decreasing or an increasing sequence $(a_\xi)_{\xi\in\omega_1}$ of reals. □

Finally consider the following strengthening of the notion of an Aronszajn tree.

DEFINITION 88. Let κ be a cardinal. A **κ-Suslin tree** is any Aronszajn tree T with property (2) in Definition 88 replaced with:

(2′) $|A| < \kappa$ for every antichain $A \subseteq T$.

By an antichain we mean any subset A of T with $a, b \in A \implies (a \not< b$ and $b \not< a)$

The existence of some κ-Suslin tree indicates that κ very far from having compactness properties. We point out that ZFC cannot decide if there are ω_1-Suslin trees. However, under the assumption $V = L$, that is, in the Gödel universe that we are going to discuss later, there are κ-Suslin trees for many κ, including ω_1. This shows indicates that Gödel's constructible universe significantly lacks compactness.

Suslin trees are related to the following problem. A linear ordering $(L, <)$ is **separable** if it has a countable dense subset (in the order topology). It has the **countable chain condition** if every collection of disjoint open intervals is countable. The reals $(\mathbb{R}, <)$ with their ordering is the unique separable complete dense linear ordering without endpoints. This can be checked easily by a back and forth argument (using Cantor's characterization of $(\mathbb{Q}, <)$). Suslin asked if the same is true after replacing "separable" above with "having the countable chain condition". It turns out that this question is equivalent to the existence of a ω_1-Suslin tree, and therefore, undecidable from ZFC.

3. Large cardinals

By large cardinals we mean some subclass $\mathrm{C} \subseteq \mathrm{CARD}$ of cardinals which cannot usually shown to be non-empty by ZFC and if it is assumed to be non-empty, this has consequences for the set theoretic universe V, but frequently also (and more interestingly) for the reals. Let us list some usual large cardinal assumptions in increasing order/strength (this list is far from complete):

$$\text{weakly inaccessible} < \text{inaccessible} < \text{Mahlo} < \text{weakly compact} < 0^{\#} \text{ exists} <$$

$$< \text{measurable} < \text{Woodin} < \text{strongly compact} < \text{supercompact} < \cdots$$

We will focus only on the inaccessible and the weakly compact cardinals.

Let λ be a limit ordinal. The **cofinality** $\operatorname{cof}(\lambda)$ **of** λ is the least β for which there is $f\colon \beta \to \lambda$ with $\sup_{\xi<\beta} f(\xi) = \lambda$. Any such map f is called **cofinal in** λ.

Examples.

(1) $\operatorname{cof}(\aleph_\omega) = \omega$;
(2) $\operatorname{cof}(\alpha) = \omega$ if $\alpha < \omega_1$;
(3) $\operatorname{cof}(\omega_1) = \omega_1$.

LEMMA 89. *Some properties of cofinality are*

(1) for all λ *we have that* $\operatorname{cof}(\lambda) = \aleph_\alpha$ *for some* α*;*
(2) if $\operatorname{cof}(\lambda) = \kappa$ *then there is cofinal map* $f\colon \kappa \to \lambda$ *which is also normal;*
(3) $\operatorname{cof}(\aleph_\lambda) = \operatorname{cof}(\lambda)$ *for every limit ordinal* λ*.*

PROOF. (1) if $f\colon \kappa \to \lambda$ is cofinal and $g\colon \kappa \to |\kappa|$ is its cardinality then $f \circ g^{-1}$ is also cofinal.

(2) Let $g\colon \kappa \to \lambda$ be cofinal and define by induction $f\colon \kappa \to \lambda$ as follows:

$$f(0) = g(0)$$

$$f(\xi+1) = \min\{\alpha \in \mathrm{ORD} \mid \alpha > f(\eta), \text{ and } \alpha > g(\eta) \text{ where } \eta \leq \xi\}$$

$$f(\theta) = \lim_{\xi<\theta} f(\xi).$$

We leave the reader to confirm that this works.

(3) If $f\colon \kappa \to \lambda$ is cofinal then $g\colon \kappa \to \aleph_\lambda$, with $g(\xi) = \aleph_{f(\xi)}$ is also cofinal. Conversely, If $g'\colon \kappa \to \aleph_\lambda$ is cofinal then $f'\colon \kappa \to \lambda$, with $f'(\xi) := \min\{\alpha \mid g'(xi) < \aleph_\alpha\}$ is cofinal. $\square$

Examples. We compute using the above lemma:

(1) $\operatorname{cof}(\aleph_\omega) = \omega$
(2) $\operatorname{cof}(\aleph_{\aleph_{\aleph_1}}) = \operatorname{cof}(\aleph_{\aleph_1}) = \operatorname{cof}(\aleph_1)$
(3) $\operatorname{cof}(\aleph_{\aleph_{\aleph_{\dots}}}) = \omega$.

We are interested in cardinals that cannot be exhausted by smaller cardinals:

DEFINITION 90. A cardinal κ is **regular** if it is infinite and $\operatorname{cof}(\kappa) = \kappa$. Otherwise, κ is **singular**.

Notice that both $\aleph_0$ and $\aleph_1$ are regular while $\aleph_{\aleph_0}$ is singular.

LEMMA 91. *Every "successor" cardinal* $\aleph_{\alpha+1}$ *is regular.*

PROOF. If $f\colon \kappa \to \aleph_{\alpha+1}$ is cofinal with $\kappa < \aleph_{\alpha+1}$ then we have that

$$\aleph_{\alpha+1} = |\aleph_{\alpha+1}| \leq |\bigsqcup_{\xi<\kappa} f(\xi)| \leq |\bigsqcup_{\xi<\kappa} \aleph_\alpha| = \kappa \cdot \aleph_\alpha = \aleph_\alpha,$$

a contradiction. $\square$

Are there any regular cardinals which are not successor cardinals except $\aleph_0$? These are precisely the weakly inaccessible cardinals below.

DEFINITION 92. An uncountable cardinal κ is called

(1) **weakly inaccessible** if κ is regular and $\mu^+ < \kappa$ for all $\mu < \kappa$;
(2) **strongly inaccessible** if κ is regular and $2^\mu < \kappa$ for all $\mu < \kappa$;
(3) **weakly compact** if it has the Ramsey property, i.e., $\kappa \to (\kappa)^2$.

We may now justify the relative position of these cardinals in the ordering we sketch at the beginning of this section.

LEMMA 93. *κ is weakly compact $\implies$ κ is strongly inaccessible $\implies$ κ is weakly inaccessible.*

PROOF. Since $\mu^+ \leq 2^\mu$ we have that κ is strongly inaccessible $\implies$ κ is weakly inaccessible.

Let now κ be weakly compact. We will show that κ is strongly inaccessible. First notice that κ is regular: if not, then κ is the disjoint union $\bigcup_{\xi<\lambda}\{A_\xi \mid \xi < \lambda\}$ with $\lambda < \kappa$ and $|A_\xi| < \kappa$. Define the coloring $f\colon [\kappa]^2 \to 2$ with $f(\{\alpha,\beta\}) = 0$ if and only if α, β are in the same A_ξ. Clearly there is no monochromatic set supported on a set of size κ.

To see that $\lambda < \kappa \implies 2^\lambda < \kappa$ assume the contrary. Then by a similar argument as in Theorem 87 one shows that $2^\lambda \not\to (\lambda^+)^2$. This implies that $\kappa \not\to (\lambda^+)^2$ by assumption, and since $\kappa \leq \lambda^+$, we have $\kappa \not\to (\kappa)^2$. The pertinent result needed for generalizing the argument from Theorem 87) is the following claim:

CLAIM. *The lexicographically ordered set $\{0,1\}^\lambda$ has no increasing nor decreasing sequence of length λ^+.*

PROOF OF CLAIM. Let $(f_\alpha)_{\alpha<\lambda^+}$ be a sequence in $\{0,1\}^\lambda$ with $\alpha < \beta \implies f_\alpha <_{\text{lex}} f_\beta$. Let also $\gamma \leq \lambda$ be the least γ so that $\{(f_\alpha\restriction\gamma) \mid \alpha < \lambda^+\}$ has size λ^+. By dropping some of the entries in $(f_\alpha)_\alpha$ and reindexing if necessary we may assume that $\big((f_\alpha\restriction\gamma)\big)_{\alpha<\lambda^+}$ is (strictly) increasing.

For every $\alpha < \lambda^+$ let ξ_α be so that $f_\alpha\restriction\xi_\alpha = f_{\alpha+1}\restriction\xi_\alpha$ but $f_\alpha(\xi_\alpha) = 0 < 1 = f_{\alpha+1}(\xi_\alpha)$. Clearly $\xi_\alpha < \gamma$, hence there is some $\xi < \gamma$ so that $\xi = \xi_\alpha$ for λ^+-many α. But then, the set $\{(f_\alpha\restriction\xi) \mid \alpha < \lambda^+\}$ has size λ^+ since if $\xi = \xi_\alpha = \xi_\beta$ and $f_\alpha\restriction\xi = f_\beta\restriction\xi$ we have that $f_\beta < f_{\alpha+1}$ and $f_\alpha < f_{\beta+1}$ which holds only for $\alpha = \beta$. This contradicts the minimality assumption on γ. $\square$

$\square$

Going back to the tree property, if κ is singular then it is easy to see that there are κ-Aronszajn trees: let $\kappa \setminus \{0\}$ be a disjoint union $\bigcup_{\alpha<\lambda} X_\alpha$ with and consider the tree $(\kappa, <_T)$ with $\xi <_T \zeta$ if and only if $\xi, \zeta \in X_\alpha \bigcup\{0\}$ for some α and $\xi < \zeta$. The situation with regular cardinals is a bit more complicated and cannot be decided from ZFC even for $\aleph_2$. However, it is known that under the assumption $V = L$ there is a κ-Aronszajn tree for each infinite successor cardinal $\kappa = \aleph_{\alpha+1}$. In other words, under the assumption $V = L$ the only place we can look for "compactness properties" above $\aleph_0$ is inaccessible cardinals and above. With the next two theorems we establish that the next place after $\aleph_0$ where we have compactness "with certainty" is from weakly compact cardinals and above.

THEOREM 94. *If κ is weakly compact then κ has the tree property.*

PROOF. Let $(T, <_T)$ be a tree with κ-many levels, each of size $< \kappa$. Since $\kappa \leq |T| \leq \kappa \cdot \kappa = \kappa$ we can assume without loss of generality that $T = \kappa$. So we have on κ two orderings so far: the usual total ordering $<$, and the partial ordering $<_T$. We extend this partial ordering $<_T$ to a total ordering $\ll$ as follows: if α, β are $<_T$-incomparable then set $\alpha \ll \beta$ if and only if $\alpha_\xi < \beta_\xi$, where ξ is the first level on which the predecessors α_ξ and β_ξ of α and β, respectively, differ.

Consider now the coloring $f\colon [\kappa]^2 \to \{0,1\}$ with $f(\{\alpha,\beta\}) = 1$ if and only if $<$ and $\ll$ agree on $\{\alpha, \beta\}$. Since κ is weakly compact there is $A \subseteq \kappa$ of size κ so that f is constant on $[A]^2$. Consider the set

$$C := \{x \in \kappa \mid \text{ there are } \kappa\text{-many } \alpha \in A \text{ with } x <_T \alpha\}$$

We claim that this is a $<_T$-chain of size κ. To see that C has size κ notice that every level of the tree T contains at least one $x \in A$ (otherwise $|T| < \kappa$). To see that C is a chain assume towards contradiction that there are $x, y \in C$ which are $<_T$-incomparable. Assume without loss of generality that $x \ll y$. But then we pick

$$\alpha < \beta < \gamma \text{ within } A$$

so that $x <_T \alpha$, $x <_T \gamma$, and $y <_T \beta$. It follows that $\alpha \ll \beta$ and $\gamma \ll \beta$, and therefore $f(\{\alpha,\beta\}) \neq f(\{\beta,\gamma\})$, contradicting that $f{\restriction}[A]^2$ is constant. □

The following theorem shows this is the best we can provably do relative to ZFC.

THEOREM 95. *If κ is strongly inaccessible and has the tree property then it is weakly compact. Moreover, if $V = L$ (and weakly compact cardinals exist), then there is no cardinal between $\aleph_0$ and the first weakly compact which has the tree property.*

PROOF. Let $f\colon [\kappa] \to \{0,1\}$. We will find $A \subseteq \kappa$, with $|A| = \kappa$, so that $f{\restriction}[A]^2$ is constant.

We first construct a tree $(T, \subset)$ whose elements are maps $t\colon \gamma \to \{0,1\}$ where $\gamma < \kappa$. We will do this by induction on κ. For each $\alpha < \kappa$ we will add precisely one new element in (what eventually will become) T. Let $t_0 := \emptyset$. Assume we have constructed $(t_\beta)_{\beta<\alpha}$ we construct t_α by induction on ξ: assume that $t_\alpha{\restriction}\xi$ has been constructed. If $t_\alpha{\restriction}\xi = t_\beta$ for some β then set $t_\alpha(\xi) = f(\{\alpha,\beta\})$. If $t_\alpha{\restriction}\xi$ was a new element to start with ($\neq t_\beta$ for any $\beta < \alpha$) then consider the construction of t_α finished,$t_\alpha := t_\alpha \restriction \xi$.

Notice that each level $\gamma < \kappa$ of the tree has at most $2^{|\gamma|}$ elements and since κ is inaccessible each level has size $< \kappa$. This additionally implies that there are κ-many levels (since we added κ-many t_α). Since κ has the tree property there is a chain $C \subseteq T$ of size κ. Let $A_0 := \{\alpha \mid t_\alpha \in C, t_\alpha^\frown 0 \in C\}$ and $A_1 := \{\alpha \mid t_\alpha \in C, t_\alpha^\frown 1 \in C\}$. Since A_0, A_1 partitions $\{\alpha \mid t_\alpha \in C\}$, one of the two, say A_i, has size κ. Let then $A := A_i$ and notice that for every $\alpha < \beta$ in A we have that $f(\{\alpha, \beta\}) = t_\beta(\text{length}(t_\alpha)) = i$. $\square$

One may intuitively argue that there ought to be inaccessible cardinals in V. For example, in any model $\mathcal{M}$ of set theory $\text{ORD}^{\mathcal{M}}$ would be an inaccessible cardinal if it was a set. One can then, presumably extend $\mathcal{M}$ to a new model by adding $\text{ORD}^{\mathcal{M}}$ as a set. Then $\text{ORD}^{\mathcal{M}}$ in the new model would still be inaccessible but moreover also a cardinal. However, ZFC cannot prove the existence of inaccessible cardinals as we will now see.

THEOREM 96. *Let κ be a strongly inaccessible cardinal. Then $V_\kappa \models$ ZFC.*

PROOF. The proof is similar to the one needed (in HW1) for showing that $V_{\omega+\omega}$ is a model of Z. It is not difficult to see that all $V_\omega, V_{\omega+\omega}, V_\kappa$ satisfy AC and Foundation. Inaccessibility is used for showing that V_κ satisfies the axiom of replacement.

CLAIM. *If $x \in V_\kappa$ then $|x| < \kappa$.*

PROOF OF CLAIM. If $x \in V_\kappa$ then $x \subseteq V_\alpha$ for some α. Hence $|x| < |V_\alpha|$. So it suffices to show that $|V_\alpha| < \kappa$ for all $\alpha < \kappa$. This is done by induction. The 0-case is trivial. If $|V_\alpha| = \mu < \kappa$ then $|V_{\alpha+1}| = 2^\mu < \kappa$ since κ is strongly inaccessible. Finally, if $\lambda < \kappa$ is a limit ordinal then $|V_\lambda| = |\bigcup_{\xi<\lambda} V_\xi| = \bigcup_{\xi<\lambda} |V_\xi|$. Since by inductive assumption $|V_\xi| < \kappa$ we have by inaccessibility (by $\text{cof}(\kappa) = \kappa$, in particular) that $|V_\lambda| < \kappa$. $\square$

To see that V_κ satisfies the axiom of replacement, let F be a class function (from the perspective of V_κ) and $x \in V_\kappa$ with the property that $x \subseteq \text{dom}(\text{F})$ (again from the perspective of V_κ). We will show that $\text{F}''x \in V_\kappa$. When we say "from the perspective of V_κ" we mean that $\text{F} \cap V_\kappa \times V_\kappa$ is a function and for every $a \in x$ there is b in V_κ with $\text{F}(a) = b$.

It suffices to show that $\text{F}''x \subseteq V_\alpha$ for some $\alpha < \kappa$. But if this was to fail then the map $y \mapsto \text{rank}(\text{F}(y))$ with $y \in x$ would be cofinal in κ. This gives a contradiction by the previous claim and the fact that $\text{cof}(\kappa) = \kappa$. $\square$

COROLLARY 97. *If ZFC is consistent, so is ZFC+"there is no strongly inaccessible cardinal".*

PROOF. Assume the contrary, that is ZFC$\vdash$"inaccessible cardinals exist". Let κ_0 be the least inaccessible cardinal. Then $V_{\kappa_0} \models$ZFC, hence there is $\kappa \in V_{\kappa_0}$ that is inaccessible from the perspective of V_{κ_0}. It is not difficult to see however that κ would also be inaccessible in V, contradicting the minimality assumption on κ_0.

Indeed, for every cardinal λ in V with $\lambda < \kappa$ we have that $\lambda \in V_{\kappa_0}$ and $\lambda < \kappa$ from the perspective of V_{κ_0}. Since κ is inaccessible in V_{κ_0} we have that $2^\lambda < \kappa$ from the perspective of V_{κ_0}. But this implies $2^\lambda < \kappa$ from the perspective of V since if $a, b \in V_{\kappa_0}$ then "$a \in b$ from the perspective of V_{κ_0}" if and only if "$a \in b$ from the perspective of V." Similarly one argues that $\mathrm{cof}(\kappa) = \kappa$ from the perspective of V_{κ_0} implies $\mathrm{cof}(\kappa) = \kappa$ from the perspective of V.

□

COROLLARY 98. *If ZFC is consistent, so is ZFC+"there is no weakly inaccessible cardinal".*

PROOF. We will see when we study L that if ZFC is consistent then so is ZFC+"$\lambda^+ = 2^\lambda$ for infinite λ". But then, a weakly inaccessible cardinal is strongly inaccessible and we may apply the previous argument. □

4. Relativization, absoluteness and reflection

We have seen already that $V_\omega, V_{\omega+\omega}$ satisfy fragments of ZFC and V_κ satisfies all of ZFC, when κ is strongly inaccessible. More generally, let M be any subclass of V (given by some formula $\pi(x)$). We may want to study the structure $(M, \in^{\mathrm{M}})$, where $\in^{\mathrm{M}}$ is just the restriction of $\in$ to $\mathrm{M} \times \mathrm{M}$. Any formula φ has an interpretation in M which generally disagrees with the interpretation of φ in V. Often we are interested in the interplay between the interpretations of φ in V and in M and we need to have the appropriate notation to distinguish between the two, as well as conditions which guarantee when the two interpretations agree.

Lets consider the example of $\mathrm{M} = V_\kappa$ in Theorem 96 and Corollary 97. When we proved axiom of replacement holds in M we considered a class function F and some $x \in \mathrm{M}$ with" $x \subseteq \mathrm{dom}(\mathrm{F})$ from the perspective of M". This just meant that for every $a \in x$ there is b **within** M so that $\mathrm{F}(a) = b$. For example, any subset x of V_κ is a subset of the domain of the class function

$$\mathrm{F} := \{(a, \kappa) \mid a \in \mathrm{V}\}$$

from the perspective of V but not from the perspective of $\mathrm{M} = \mathrm{V}_\kappa$, if $x \neq \emptyset$. We will write $x \subseteq \mathrm{dom}(\mathrm{F})$ and $x \not\subseteq \mathrm{dom}^{\mathrm{M}}(\mathrm{F})$ to distinguish between the two statements.

DEFINITION 99. Let M be a class and let φ be any formula. We define inductively the **relativization** φM of φ on M to be:

(1) $(x = y)^{\mathrm{M}}$ is $x = y$;
(2) $(x \in y)^{\mathrm{M}}$ is $x \in y$;
(3) $(\varphi \wedge \psi)^{\mathrm{M}}$ is $\varphi^{\mathrm{M}} \wedge \psi^{\mathrm{M}}$;
(4) $(\neg\varphi)^{\mathrm{M}}$ is $\neg\varphi^{\mathrm{M}}$;
(5) $(\exists x\, \varphi)^{\mathrm{M}}$ is $\exists x\, \big((x \in \mathrm{M}) \wedge \varphi^{\mathrm{M}}\big)$;

Hence, if φ is a formula and $\bar{a}$ is a tuple in M we have that $\varphi^{\mathrm{M}}(\bar{a})$ holds if and only if $(\mathrm{M}, \in^{\mathrm{M}}) \models \varphi(\bar{a})$.

One has to be careful from now on when using abbreviations introduced above. For example $(x \subseteq y)^{\mathrm{M}}$ stands for $\forall z \in \mathrm{M}(z \in x \implies z \in y)$ and the powerset $\mathcal{P}^{\mathrm{M}}(x)$ of $x \in \mathrm{M}$, if it exists in M, is the unique set y with the property that $\forall z \in \mathrm{M}(z \in y \iff z \subseteq^{\mathrm{M}} x)$. Here however, we will only consider classes M for which the relativization of many simple formulas takes a simple form. For example we have the following lemma whose consequence is that we can often take $\mathcal{P}^{\mathrm{M}}(x)$ to simply be $\mathcal{P}(x) \cap \mathrm{M}$.

LEMMA 100. *If* M *is transitive then the axiom of extensionality holds. The powerset axiom also holds if additionally we have that* $\forall x \in \mathrm{M} \exists y \in \mathrm{M}(y = \mathcal{P}(x) \cap \mathrm{M})$.

PROOF. Left to the reader. □

In the context of the next definition, extensionality is *absolute* for transitive M.

DEFINITION 101. Let $\mathrm{M} \subseteq \mathrm{N}$ be two classes. Call φ **absolute for** M, N if

$$\forall \bar{x} \in \mathrm{M}(\varphi^{\mathrm{M}}(\bar{x}) \iff \varphi^{\mathrm{N}}(\bar{x})).$$

We say that φ is **absolute for** M if it is absolute for M, V.

Recall from model theory that a $\exists$-formula is "upward absolute" and a $\forall$-formula is "downward absolute". Hence a formula equivalent to both an existential and a universal formula is absolute for model-theoretic reasons. In set theory we have the a very convenient form of absoluteness for formulas with "bounded quantification"

DEFINITION 102. The collection of Δ_0-formulas is the build inductively by

(1) $(x \in y) \in \Delta_0$ and $(x = y) \in \Delta_0$;
(2) if $\varphi, \psi \in \Delta_0$ then $\neg\varphi, \varphi \wedge \psi \in \Delta_0$;
(3) if $\varphi \in \Delta_0$ then so is $\exists x(x \in y \wedge \varphi)$

The following theorem is very useful since many formulas are equivalent relative to (fragments of) ZFC to Δ_0-formulas. Some examples are

$$y = x \cup \{x\}, \quad z = \{x, y\}, \quad y = \bigcup X, \quad x \text{ is an ordinal}$$
$$x \text{ is a limit ordinal}, \quad x \text{ is a successor ordinal}, \quad x = \omega$$

THEOREM 103. *In* M *is transitive and* $\varphi \in \Delta_0$ *then* φ *is absolute for* M.

PROOF. Exercise. □

We are interested in classes $\mathrm{M} \subseteq \mathrm{V}$ which model ZF. Ideally we would like M to be a set, in which case we would like to argue, ideally, that ZF $\vdash$ "M models ZF". However, because of Gödel's second incompleteness theorem this is hopeless. In fact, this gives another proof as to why ZFC cannot prove that there exists a strongly inaccessible κ since, otherwise, $\mathrm{M} = V_\kappa$ is a set and ZFC $\vdash$ "M models ZFC". That being said, the next theorem shows that we can approximate this goal from below...

THEOREM 104 (Reflection Theorem). *If $\varphi_1, \ldots, \varphi_n$ are finitely many formulas:*

$$\mathrm{ZF} \vdash \forall\alpha\ \exists\beta > \alpha\ (\varphi_1, \ldots, \varphi_n \text{ are absolute for } V_\beta)$$

In fact, one may prove the following more general version of this theorem which is very useful for the study of L, as we will see in the next section.

THEOREM 105 (General Reflection Theorem). *Assume that* ZF *proves that for some class* M *and some sequence* $(\mathrm{M}_\alpha)_{\alpha\in\mathrm{ORD}}$ *that:*

(1) $\alpha < \beta \implies \mathrm{M}_\alpha \subseteq \mathrm{M}_\beta$*;*
(2) $\mathrm{M}_\lambda = \bigcup_{\xi<\lambda} \mathrm{M}_\xi$*;*
(3) $\mathrm{M} = \bigcup_{\alpha\in\mathrm{ORD}} \mathrm{M}_\alpha$.

Then for any finitely many formulas $\varphi_1, \ldots, \varphi_n$ we have that ZF proves

$$\forall\alpha\ \exists\beta > \alpha\ (\varphi_1, \ldots, \varphi_n \text{ are absolute for } \mathrm{M}_\beta, \mathrm{M})$$

PROOF. A subformula of a formula φ is any formula ψ which appears as a node in the syntactic tree of φ. By expanding the list $\varphi_1, \ldots, \varphi_n$ to a larger finite list we may assume without loss of generality that every subformula of each φ_i appears somewhere in the list. Call such list **subformula closed**. By the same inductive argument as the one in the Tarski-Vaught test (see model theory) it suffices to prove the following claim (notice both superscripts are M in φ_j below).

CLAIM. *For all α there is $\beta > \alpha$ so that if φ_i is of the form $\exists x\varphi_j(x, \bar{y})$ then*

$$\forall \bar{y} \in \mathrm{M}_\beta \left(\exists x \in \mathrm{M}\ \varphi_j^{\mathrm{M}}(x, \bar{y}) \iff \exists x \in \mathrm{M}_\beta\ \varphi_j^{\mathrm{M}}(x, \bar{y})\right)$$

PROOF OF CLAIM. For each formula φ_i of the form $\exists x\varphi_j(x, \bar{y})$ consider the class function $\mathrm{G}_i\colon \mathrm{M} \times \cdots \times \mathrm{M} \to \mathrm{ORD}$ with $\mathrm{G}_i(\bar{a}) = 0$, if $\neg\exists x \in \mathrm{M}\varphi_j(x, \bar{a})$; and $\mathrm{G}_i(\bar{a}) =$"least ξ so that $\exists x \in \mathrm{M}_\xi\varphi_j(x, \bar{a})$", otherwise. Consider the associated map

$$\mathrm{F}_i\colon \mathrm{ORD} \to \mathrm{ORD}, \quad \text{with} \quad \mathrm{F}_i(\eta) = \sup\{\mathrm{G}_i(\bar{a}) \mid \bar{a} \in \mathrm{M}_\xi\}.$$

Start with α and by induction define increasing $(\alpha_k)_{k\in\omega}$ with

$$\alpha_{k+1} := \max\{\mathrm{F}_1(\alpha_k), \ldots \mathrm{F}_n(\alpha_k), \alpha_k + 1\}.$$

Set $\beta = \sup_k \alpha_k$ and notice that for all $\xi < \alpha_k$ and all i we have that $\mathrm{F}_i(\xi) < \beta$. □

□

5. Gödel's constructible universe L

We will define here a class L of sets. Every universe of ZF contains within it its own version of L. We will show that for every single axiom σ of ZF we have that ZF proves that σ holds in L. We will then show that ZF proves that L satisfies the Generalized Continuum Hypothesis as well as AC. This implies that ZF cannot prove $\neg$AC, as well as ZFC cannot prove $\neg$GCH. We will finally address the compatibility of the axiom V=L with large cardinal axioms. For example V=L is compatible with "weakly compact cardinals exist". However, we will see that it is not compatible with "measurable cardinals exist." Hence, it is consistent that the universe is "tall and thin" but not consistent that the universe is "very tall and thin."

Let Φ be the collection of all first order formulas in the language of set theory. If A is any set then we write Φ^A for the collection of all relativized formulas φ^A where $\varphi \in \Phi$. For any set A consider the collection:

$$\mathcal{D}(A) := \big\{X \subseteq A \mid \text{there is } \varphi^A(\bar{x}, y) \in \Phi^A \text{ and } \bar{a} \in A^n \text{ s.t. } X = \{b \in A \mid \varphi^A(\bar{a}, b)\}\big\}$$

$$= \big\{X \subseteq A \mid X \text{ is a definable (with parameters) subset of } A \text{ in the structure } (A, \in)\big\}.$$

Then $\mathcal{D}(A)$ is set since $\mathcal{D}(A) \subseteq \mathcal{P}(A)$. We may now define **Gödel's constructible universe** L to be the union $\bigcup_{\alpha \in \mathrm{ORD}} \mathrm{L}_\alpha$, where L_α is inductively as follows:

(1) $\mathrm{L}_0 := \emptyset$
(2) $\mathcal{L}_{\alpha+1} := \mathcal{D}(\mathrm{L}_\alpha)$;
(3) $\mathrm{L}_\lambda := \bigcup_{\xi<\lambda} \mathrm{L}_\xi$, if λ is a limit ordinal.

Notice that with the next lemma we establish that L and $(\mathrm{L}_\alpha)_\alpha$ satisfy the assumptions of the general reflection theorem (Theorem 105) above.

Lemma 106. *Let $\alpha < \beta$ be ordinals. Then*

(1) $\mathrm{L}_\alpha \subseteq \mathrm{L}_\beta$*;*
(2) L_α*, as well as* L*, are a transitive sets.*

Proof. First notice that if A is any transitive set then $A \subseteq \mathcal{D}(A)$. Indeed if $a \in A$ then the formula $\varphi(x, y) \iff y \in x$ with parameter $x = a$ introduces some element of $\mathcal{D}(A)$ which happens to be a by transitivity of A.

Assume now by induction that L_α is transitive. by the above we have that

$$\mathrm{L}_\alpha \subseteq \mathrm{L}_{\alpha+1} \subseteq \mathcal{P}(\mathrm{L}_\alpha)$$

If $x \in X \in \mathrm{L}_{\alpha+1}$ then $x \in \mathrm{L}_\alpha$ by the second $\subseteq$ above and then $x \in \mathrm{L}_{\alpha+1}$ by the first $\subseteq$ above. We have therefore shown the successor case in the inductive argument which proves both (1), (2). The rest follows by straightforward properties of unions. $\square$

We now describe briefly the "geometry" of L.

Definition 107. For every $x \in \mathrm{L}$ we denote by $\mathrm{rank}_{\mathrm{L}}(x)$ the **L-rank of** x, i.e., least ordinal α so that $x \in L_{\alpha+1}$.

Lemma 108. *The universe* L *has the following properties:*

(1) for all $\alpha \leq \omega$ we have that $\mathrm{L}_\alpha = V_\alpha$*;*
(2) (assuming AC) for all $\alpha \geq \omega$ we have that $|\mathrm{L}_\alpha| = |\alpha|$*.*
(3) $\mathrm{ORD} \cap \mathrm{L}_\alpha = \alpha$*, and hence,* $\mathrm{ORD}^{\mathrm{L}} = \mathrm{ORD}$*;*

Proof. (1) is immediate. For (2), notice that $|\mathrm{L}_{\alpha+1}| \leq |\Phi \times (\mathrm{L}_\alpha)^{<\omega}| = |\mathrm{L}_\alpha|$, and since $|\alpha + 1| = \alpha$ we see by induction that $|\mathrm{L}_\alpha| \leq |\alpha|$. The converse, $|\alpha| \leq |\mathrm{L}_\alpha|$, follows from property (3). For (3), we argue by induction. The only non-trivial case is the successor case, so assume that $\mathrm{ORD} \cap \mathrm{L}_\alpha = \alpha$. Since $\mathrm{L}_\alpha \subseteq \mathrm{L}_{\alpha+1} \subseteq \mathcal{P}(\mathrm{L}_\alpha)$ we have that $\alpha \subseteq \big(\mathrm{L}_{\alpha+1} \cap \mathrm{ORD}\big) \subseteq \alpha + 1$. From HW there is a Δ_0-formula φ so that

$$\forall x (x \in \mathrm{ORD} \iff \varphi(x)).$$

But then, the set $a := \{x \in \mathrm{L}_\alpha \mid \varphi^{\mathrm{L}_\alpha}(x)\}$, which is contained in $\mathrm{L}_{\alpha+1}$, is equal to α by Lemma 103 and since L_α is transitive. It follows that $\alpha \in \mathrm{L}_{\alpha+1}$ and so $\big(\mathrm{L}_{\alpha+1} \cap \mathrm{ORD}\big) = \alpha + 1$. $\square$

Clearly, by (2) above, we have that $\mathrm{L}_{\omega+1} \neq \mathcal{P}(\mathrm{L}_\omega)$. In fact we moreover have:

$$\mathrm{L}_{\omega+1} \neq \mathcal{P}^{\mathrm{L}}(\mathrm{L}_\omega).$$

The reason is that the bijection $f \colon \mathrm{L}_{\omega+1} \to \omega$ witnessing that $|\mathrm{L}_{\omega+1}| = \aleph_0$ in the previous lemma can be seen to be an element of $\mathrm{L}_{\omega+2}$ and therefore of L. So L is "aware" that $\mathrm{L}_{\omega+1}$ is countable but since L satisfies ZFC it is also "aware" that $\mathcal{P}^{\mathrm{L}}(\mathrm{L}_\omega)$ is uncountable. Which means that not every element of $\mathcal{P}^{\mathrm{L}}(\mathrm{L}_\omega)$ has L-rank ω. For the same reasons $\mathcal{P}^{\mathrm{L}}(\omega) \cap \mathrm{L}_{\omega+1} \neq \mathcal{P}^{\mathrm{L}}(\omega)$. One can say that the elements $\mathcal{P}^{\mathrm{L}}(\omega) \cap \mathrm{L}_{\omega+1}$ are precisely the arithmetical subsets of ω, or equivalently, the subsets of ω which are computable from the n-th Turing jump for some $n \in \omega$. The question remains: what is the smallest ordinal α so that $\mathcal{P}^{\mathrm{L}}(\omega) \cap \mathrm{L}_\alpha = \mathcal{P}^{\mathrm{L}}(\omega)$? We will come back to this question after we establish the following theorem.

Theorem 109 (ZF). L *is a model of ZF.*

Proof. Extensionality follows by Lemma 100, since L is transitive. Foundation is clear since $\mathrm{L} \subseteq \bigcup_\alpha \mathrm{V}_\alpha$. Infinity axiom is also clear since $V_\omega \in \mathrm{L}_{\omega+1}$

For the subset axiom, let $\varphi(z, \bar{x})$ be a formula, we want to show that:

$$\forall x, x_1, \dots x_n \in \mathrm{L}\, \exists y \in \mathrm{L}\, \forall z \in \mathrm{L}\, \big((z \in y)^{\mathrm{L}} \iff ((z \in x)^{\mathrm{L}} \wedge \varphi^{\mathrm{L}}(z, x_1, \dots, x_n))\big).$$

Of course, the superscript L in $(z \in y)$ and $(z \in x)$ is redundant. Fix $x, \bar{x} \in \mathrm{L}$ and let α be large enough so that $x, \bar{x} \in \mathrm{L}_\alpha$. Notice that in stage $\mathrm{L}_{\alpha+1}$ we added:

$$y' := \{z \in x \mid \varphi^{\mathrm{L}_\alpha}(z, x_1, \dots, x_n))\}$$

but this in not the same as the desired y since $\varphi^{\mathrm{L}} \neq \varphi^{\mathrm{L}_\alpha}$ in general. However, by Theorem 105 we may pick some $\beta > \alpha$ so that $(z \in x) \wedge \varphi$ is absolute for $\mathrm{L}_\alpha, \mathrm{L}$. But then we can take y to simply be

$$y := \{z \in \mathrm{L}_\beta \mid (z \in x) \wedge \varphi^{\mathrm{L}_\beta}(z, x_1, \dots, x_n))\} \in \mathcal{D}(\mathrm{L}_\beta) \subseteq \mathrm{L}$$

which was constructed by Theorem 105 to be equal to the desired:

$$\{z \in x \mid \varphi^{\mathrm{L}}(z, x_1, \dots, x_n)\}$$

The pairing axiom, union axiom, powerset axiom, and replacement axiom all follow by establishing existence in L of certain large enough sets. We just show replacement. So let $\varphi(x, y, \bar{w})$ be a formula and let $X, \bar{a} \in \mathrm{L}$ so that:

$$\big(\varphi(x, y, \bar{a}) \text{ is a class function F and } X \subseteq \mathrm{dom}(\mathrm{F})\big)^{\mathrm{L}}$$

This implies that "from the perspective of V" we have

$$\varphi^{\mathrm{L}}(x, y, \bar{a}) \bigcap (\mathrm{L} \times \mathrm{L}) \text{ is a class function } \mathrm{F}' \text{ and } X \subseteq \mathrm{dom}(\mathrm{F}' \bigcap (\mathrm{L} \times \mathrm{L})),$$

So by replacement in V we have that $\{\mathrm{rank}_{\mathrm{L}}(\mathrm{F}'(x)) \mid x \in X\} \subseteq \mathrm{ORD}$ is a set. Pick any β ordinal which bounds this set from above. The rest follows by the subset axiom applied to the formula $\mathrm{F}''X$ on the set L_β.

$\square$