Set Theory I

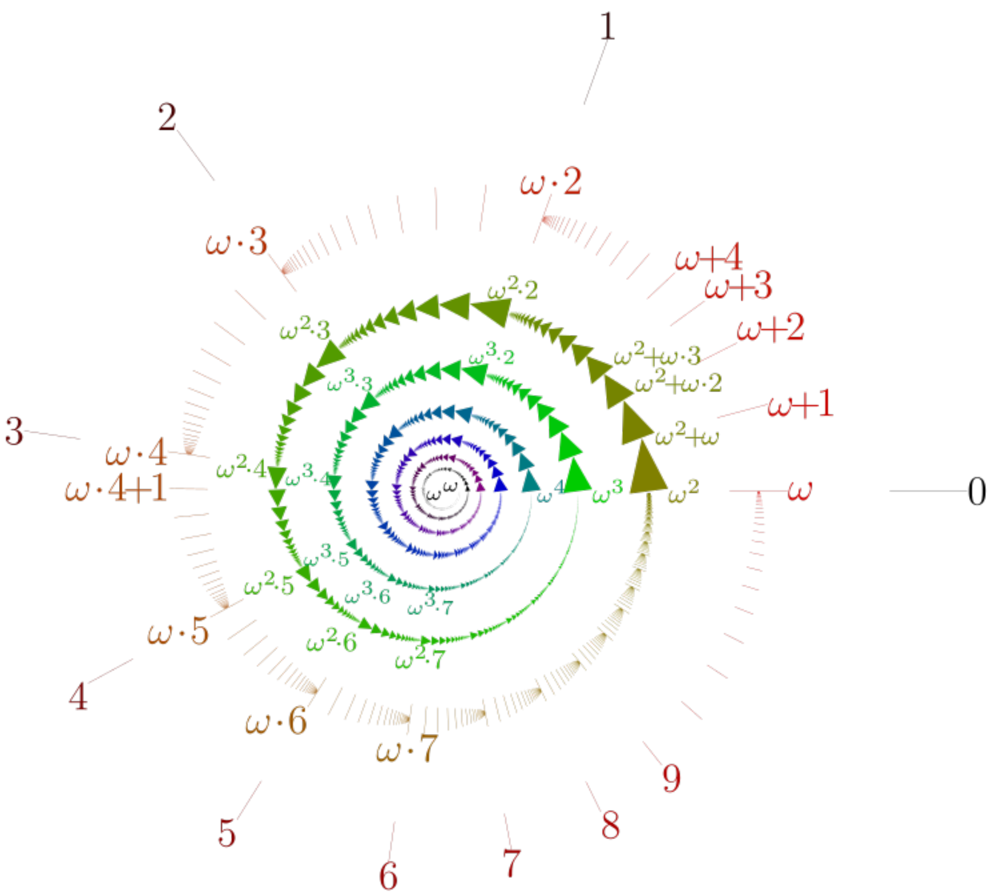
Based on lectures by

Notes taken by yehelip

Winter 2025

These notes are not endorsed by the lecturers. I have revised them outside lectures to incorporate supplementary explanations, clarifications, and material for fun. While I have strived for accuracy, any errors or misinterpretations are most likely mine.

$\omega + 1 \neq 1 + \omega$



Contents

1 Permutations **4**
1.1 Deinitions 4
1.2 The symmetric group 4

2 Hall's theorem **5**

3 Cantor's theorem **6**

4 Equivalence Relations **7**
4.1 Some terminology 7

5 Kőnig's Theorem **8**

6 Partial Orders **9**

7 Partially Ordered Sets **9**
7.1 Extrema 9
7.1.1 About lattices 9

8 Cardinals **10**

9 Schröder–Bernstein Theorem **11**

10 Homomorphism and Isomorphosm of Ordered Sets **12**
10.1 Homomorphisms 12
10.2 Isomorphisms 12
10.3 Lexicographic Order 12

11 Zorn's Lemma **13**
11.1 All vector spaces have a basis 13
11.2 Comparing Cardinals 13
11.3 Corollaries 14

12 More Axioms In ZF **15**

13 Axiom of Choice **15**
13.1 Nomenclature 15

14 Measure **16**
14.1 Lebesgue Measure 16

15 Well Order **17**
15.1 Some lemmas 17
15.2 A Lemma About Partial Orders 17

16 Comparison of Well Ordered Sets **18**

17 Ordinals **19**
17.1 Types of Ordinals 20
17.2 Ordinal Arthimetic 20
17.2.1 addition 20
17.2.2 multiplication 20
17.2.3 Powers 20

17.2.4 The Ordinals 2^ω and ω^ω . . . 21

18 The Well Ordering Theorem **22**
18.1 Proof From AC . . . 22
18.2 Proof of AC using WOT . . . 22

19 The Continuum Conjecture and the Alephs **23**
19.1 The Continuum Conjecture . . . 23
19.2 The Alephs . . . 23

20 Transfinite Induction **25**
20.1 Proof That The Only Isomorphism from a Well-Ordered Set to Itself is the Identity Isomorphism . . . 25

21 Extras **26**
21.1 A Bit About Constructions . . . 26
21.1.1 Construction of $\mathbb{N}$ . . . 26
21.1.2 Construction of $\mathbb{Z}$ . . . 26
21.1.3 Construction of $\mathbb{Q}$ . . . 26
21.1.4 About the Construction of $\mathbb{R}$ . . . 26
21.2 Discrete Functions . . . 26
21.3 More definitions . . . 27
21.3.1 Saturated dunctions . . . 27
21.3.2 Hasse diagrams . . . 27
21.3.3 Some denotions . . . 27

1 Permutations

1.1 Deinitions

Definition 1.1 (Permutation)**.** A permutation σ is a bijection from a set S onto iteself.

Example 1.1. We can denote permutations by a matrix of order $2 \times N$ where $N := |S|$. Let $S = \{1, 2, 3, 4, 5\}$. We interpret the permuatation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}$$

as $\sigma(1) = 2$, $\sigma(2) = 5$, $\sigma(3) = 4$, $\sigma(4) = 3$ and $\sigma(5) = 1$. We notice that the numbers (3 4) and (1 2 5) form what we call a "cycles". Thus, we can also write

$$\sigma = (3\ 4)(1\ 2\ 5).$$

Here are some more definitions related to permutations:

Definition 1.2 (Fixed point)**.** A cycle of one element is call a fixed point.

Definition 1.3 (Derangement)**.** A permutation without fixed points is called a derangement.

Definition 1.4 (Transposition)**.** A permutation that is an orbit of 2 elements is called a transposition.

1.2 The symmetric group

Definition 1.5 (The symmetric group)**.** Let S be a set. The symmetric group defined over S is the group whose elements are all the permutation over S, and whose group operation is the composition of functions.

Remark 1.1. A group is an algebraic structure with the following characteristics:

- Associativity
- An idenentity permutation exists
- Every element has an inverse

2 Hall's theorem

Definition 2.1 (Neighborhood)**.** In a bipartite graph $G = (X, Y, E)$ the neighborhood of a subset X' of X denoted $N_G(X')$ is the set of all the vertices in Y that share an edge with some vertex from X'.

Theorem 2.1. (Hall's theorem). *In a finite bipartite graph $G(X, Y, E)$ a perfect matching exists if and only if for any subset W of X exists an injection from W to $N_G(W)$.*

Proof. First suppose there exists an X perfect matching. Since for any given W all vertices in W have a distinct matching vertex in Y, we get that the matching function is an injection from W to $N_G(W)$ which completes the easy part of the proof.

Next, assume by contradiction that an X-perfect matching doesn't exist. We can denote the maximal matching M, and the sets of vertices in X, Y that appear in M as S, T. Since an X-perfect matching doesn't exist we get $X \setminus S$ is not empty so we can choose a vertex $u_0 \in X \setminus S$ and consider all alternating paths of the form $P = (u_0, v_1, v_2, \ldots)$ such that odd edges are not in M and even edges are in M. Denote:

$$A = \big\{u \mid u \in P \wedge u \in X\big\}$$
$$B = \big\{v \mid v \in P \wedge v \in Y\big\}$$

We know every vertex in B is matched by M to a vertex in A because otherwise we could create a bigger matching by toggling whether each of the edges belong to M or not. It follows that

$$|B| \leq |A \setminus \{u_0\}| \Rightarrow |B| < |A|.$$

But also for any vertex $a \in A$, all of its neighbors are in B which implies that

$$N_g(A) \leq B.$$

We can also show that an alternating path to b exists either by removing the matched edge ab from the alternating path to a, or by adding the unmatched edge ab to the alternating path to a.

$$\Rightarrow B = N_g(A)$$
$$\Rightarrow |N_g(A)| < |A|$$

That's a contradiction so an X-perfect matching must exist, which completes the proof. $\square$

3 Cantor's theorem

Theorem 3.1. (Cantor's theorem). *Let A be a set. Then*

$$|A| < |\mathcal{P}(A)|.$$

We can define $f \colon A \to \mathcal{P}(A)$ as such:

$$f(a) = \{a\}.$$

This is clearly an injection so we get:

$$|A| \leq |\mathcal{P}(A)|.$$

Assume that $|A| = |\mathcal{P}(A)|$. That means there exists a bijection $g \colon A \to \mathcal{P}(A)$. Consider the following set:

$$D = \left\{a \colon a \notin g(a)\right\}.$$

Since g is a bijection exists $b \in A$ such that $f(b) = D$. Next consider the different possible cases:

$$\begin{cases} b \in D, & b \notin g(b) = D \implies \text{contradiction} \\ b \notin D = g(b), & b \in D \implies \text{contradiction} \end{cases}$$

Therefore $|A| \neq |\mathcal{P}(A)|$ which implies that $|A| < |\mathcal{P}(A)|$ and completes the proof.

4 Equivalence Relations

Definition 4.1. An equivalence relation on a set S is a binary relation R that satisfies the following for all $a, b, c \in S$.

- $(a, b) \in R$ (reflexivity);
- $(a, b) \in R$ implies that $(b, a) \in R$ (symmetricity);
- $(a, b) \in R$ and $(b, c) \in R$ implies that $(a, c) \in R$ (transitivity).

4.1 Some terminology

Suppose we have an equivalnce relation R on a set X.

Definition 4.2 (Equivalence class)**.** The equivalence class of a in R denoted $[a]_R$ is defined as:

$$\left\{b \in X \mid bRa = 1\right\}.$$

Definition 4.3 (Quotient set)**.** The quotient set denoted ${}^{X}/_{R}$ is defined as such:

$$\left\{[a]_R \mid a \in X\right\}$$

Definition 4.4 (Projection)**.** A projection of R is a function $\pi \colon X \to {}^{X}/_{R}$ such that:

$$\pi(x) = [x]_R.$$

Definition 4.5 (Cut)**.** A cut of X is a set with only one element of each equivalence class.

Equivalence relations can be defined by their quotient set. Thus they can also be defined by a function or a partition.

Definition 4.6 (Bell's numbers)**.** The numbers of partitions of a set $|X| = n$ are known as Bell's numbers and can be calculated recursivly as such:

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k.$$

Think why this is true.

5 Kőnig's Theorem

Theorem 5.1. (Kőnig's Theorem). *Let I be an index set. If for all $i \in I$ and κ_i, λ_i we know that $\kappa_i < \lambda_i$ then*

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i.$$

Proof. The idea of the proof is to prove that for any function

$$f \colon \sum_{i \in I} B_i \to \prod_{i \in I} C_i \quad \text{such that} \quad |B_i| = \kappa_i \text{ and } |C_i| = \lambda_i$$

that f is not surjective. Define the functions f_i for all $i \in I$ as such:

$$\begin{array}{rcrcl} f & : & B_i & \longrightarrow & C_i \\ & & f_i(x) & \longmapsto & f(x)_i \end{array}.$$

Then for all $i \in I$ we have that $|B_i| < |C_i|$ which implies that for all $i \in I$ that f_i is not surjective. Therefore for all $i \in I$ exist $c_i \in C_i \setminus \operatorname{im}(f_i)$. Consider the vector

$$\hat{c} := \langle c_i \mid i \in I \rangle.$$

If $\hat{c} \in \operatorname{im} f$ then exist $i \in I$ and $b \in B_i$ such that $f(b) = \hat{c}$. This implies that $f(b)_i = c_i$ so $f_i(b) = c_i$ but since $c_i \in C_i \setminus \operatorname{im} f_i$ we get a contradiction. Therefore:

$$\left| \sum_{i \in I} B_i \right| < \left| \prod_{i \in I} C_i \right| \implies \sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i$$

$\square$

6 Partial Orders

Definition 6.1 (Weak partial order)**.** A weak/non-strict partial order is a homogeneous relation $\leq$ on a set P such that for all $a, b, c \in P$ we have

- $a \leq a$ (Reflexivity);
- $a \leq b$ and $b \leq a$ imply that $a = b$ (Antisymmetricity);
- $a \leq b$ and $b \leq c$ imply that $a \leq c$ (Transitivity).

Definition 6.2 (Strong partial order)**.** A strong/strict partial Order is a homogeneous relation $<$ on a set P such that for all $a, b, c \in P$ we have

- $a \not< a$ (Irreflexivity);
- $a < b$ implies that $b \not< a$. (Asymmetricity);
- $a \leq b$ and $b \leq c$ imply that $a \leq c$ (Transitivity).

Remark 6.1. We have that for any strong partial order $<$ that the union $< \bigcup \leq_{\text{id}}$ is a weak partial order which we will denote $\leq$.

7 Partially Ordered Sets

Definition 7.1 (Partially ordered set)**.** A partially ordered set (also known as a poset) is an ordered pair of a set and a partial order on that set $(A, \leq)$.

Definition 7.2 (Comparable elements)**.** Two elements $a, b \in A$ are called comparable if and only if $a \leq b$ or $b \leq a$. If two elements are incomparable they are called linearly independent.

Definition 7.3 (Linear order)**.** A linear order (also known as a total order) is a partial order under which every pair of elements is comparable.

Definition 7.4 (Chain)**.** Let $(A \leq)$ be a partially ordered set. Then a subset $C \subset A$ is called a chain if and only if the restriction of $\leq$ of C is a linear order.

7.1 Extrema

Definition 7.5 (Greatest element)**.** A greatest element a in a set S is an element such that for all $b \in S$ we have $b < a$.

Definition 7.6 (Maximal element)**.** A maximal element a in a set S is an element such that there does not exist $b \in S$ such that $a < b$.

Remark 7.1. If the set S is linearly ordered then there could only exist a single greatest element, and if it exists, then it is also a maximal element.

Definition 7.7 (Upper bound)**.** An upper bound in A of $B \subseteq A$ is an element $a \in A$ such that for every $b \in B$ we have $b \leq a$.

Remark 7.2. Similarly, a lower bound in A of $B \subseteq A$ is an element $a \in A$ such that for every $b \in B$ we have $a \leq b$.

7.1.1 About lattices

Definition 7.8 (Lattice)**.** A partially ordered set A is a lattice if and only if for every $S \subseteq A$ with exactly two elements $\sup S$ and $\inf S$ exist.

8 Cardinals

Cardinal numbers are the "numbers" we use to represent the size of sets. We denote the cardinality of $\mathbb{N}$ as $\aleph_0$, and the cardinality of $\mathbb{R}$ as $\aleph$. To get good intuition for cardinals, I suggest trying to prove or find the following:

1. $|\mathbb{N}| < |\mathbb{R}|$.
2. $\aleph_0 = \aleph_0 + n = \aleph_0 \times n = \aleph_0 \times \aleph_0 = |\mathbb{Z}| = |\mathbb{Q}|$.
3. $\aleph = 2^{\aleph_0} = \left|(0,1)^{\aleph_0}\right| = \aleph \times \aleph_0 = \aleph \times \aleph = \big|(0,1)\big| = \big|[0,1]\big|$.
4. A plane can't be covered by $\aleph_0$ lines.
5. Let A be an infinte set. Then exists $S \subseteq A$ such that $|S| = \aleph_0$.
6. $\aleph = |\mathcal{P}(\mathbb{N})| = |\mathcal{P}(\mathbb{Q})|$.
7. let $A = \big\{A \subset \mathbb{N} \mid |A| < \infty\big\}$ prove that $|A| = \aleph_0$.
8. $\aleph_0^{\aleph} = \aleph$.
9. $\left|\mathbb{R}^{\mathbb{R}}\right| = \big|\mathcal{P}(\mathbb{R})\big| = 2^{\aleph}$.
10. A countable union of countable sets is countable.
11. $\aleph_0^{\aleph_0} = \aleph$.
12. $\left|\big\{f\colon \mathbb{R} \to \mathbb{R} \mid f \text{ is invertable}\big\}\right| = 2^{\aleph}$.
13. $|A| = \left|\big\{r \in \mathbb{R} \mid r \text{ is algebraic}\big\}\right| = \aleph_0$.
14. $|B| = |\mathbb{R} \setminus A| = \left|\big\{r \in \mathbb{R} \mid r \text{ is transcedental}\big\}\right| = \aleph$.
15. $\left|\big\{A \subset \mathbb{R} \mid |A| = \aleph_0\big\}\right|$.
16. $\left|\big\{A \subset \mathbb{R} \mid |A| = \aleph\big\}\right|$.
17. Let $\aleph_0$ people with a natural number of hats on their head guess how many hats they have. How many options are there, given only a finite number of people guessed right? How many people guessed wrong?

9 Schröder–Bernstein Theorem

Theorem 9.1. (Schröder–Bernstein theorem). *The Schröder–Bernstein theorem states that if $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.*

Proof. Suppose we have two injective functions:

$$f \colon A \to B$$
$$g \colon B \to A.$$

Without loss of generality, assume that A and B are disjoint. [1] Using the partially defined functions f^{-1}, g^{-1} we can construct a sequence for every element of $A \cup B$ in the following way:

$$\ldots \to f^{-1}g^{-1}(a) \to g^{-1}(a) \to a \to f(a) \to g(f(a)) \to \ldots$$

This sequence can keep going forever to the right, but to the left it may stop eventually since the inverse functions are partial [2]. We can see that every element in $A \cup B$ has a sequence and that if an element appears in two sequnces they'll be identical since they're injective and by our construction. Thus those sequences form a partition of $A \cup B$ so it is sufficient to create bijections for all partitions. Our bijection will be:

$$h(x) = \begin{cases} f(x), & \text{for } x \in A \text{ in an } A\text{-stop} \\ g^{-1}(x), & \text{for } x \in A \text{ in a } B\text{-stop} \\ f^{-1}(x), & \text{for } x \in B \text{ in an } A\text{-stop} \\ g(x), & \text{for } x \in B \text{ in a } B\text{-stop} \end{cases}$$

It's not hard to show that this function is well defined and a bijection for each sequence. Therefore h is a bijection from $|A|$ to $|B|$ as wanted which completes the proof. □

[1] Why can we do this?

[2] We'll call those who stop from the left on an element of A A-stops and the rest B-stops - even though they may not always stop!

10 Homomorphism and Isomorphosm of Ordered Sets

10.1 Homomorphisms

Let $(X, \leq_1), (Y, \leq_2)$ be partially ordered sets.

Definition 10.1 (Homomorphism)**.** A function $f\colon X \to Y$ is a called a homomorphism if and only if for every $x_1, x_2 \in X$ we have

$$x_1 \leq_1 x_2 \Rightarrow F(x_1) \leq_2 F(x_2).$$

10.2 Isomorphisms

Definition 10.2 (Isomorphism)**.** A function $f\colon X \to Y$ is called an isomorphism if and only if it is a bijection of sets and for every $x_1, x_2 \in X$ we have

$$x_1 \leq_1 x_2 \Rightarrow F(x_1) \leq_2 F(x_2).$$

Proposition 10.1. *If F is an isomorphism and the orders are total orders then F^{-1} is also an isomorphism.*

10.3 Lexicographic Order

Definition 10.3 (Lexicographical order)**.** A lexicographical order (also known as a left dictonary order) on $X \times Y$ is an order defined as such:

$$(x_1, y_1) \leq_L (x_2, y_2) \iff x_1 <_x x_2 \vee (x_1 = x_2 \wedge y_1 \leq_y y_2).$$

Remark 10.1. We can similarly define the right dictionary order as such

$$(x_1, y_1) \leq_L (x_2, y_2) \iff y_1 <_y y_2 \vee (y_1 = y_2 \wedge x_1 \leq_x x_2).$$

11 Zorn's Lemma

Lemma 11.1 (Zorn's lemma)**.** *Let F be a nonempty partially ordered set. If for every chain in F there exists an upper bound in F, then F has at least one maximal element.*

We will not prove this lemma. But we can see here a couple of its uses.

11.1 All vector spaces have a basis

Proposition 11.2. *All vector spaces have a basis*

Proof. Let V be a vector space. If $V = \{0\}$ then its basis is $\emptyset$. If V is finitely generated then we can add vectors from V to $\emptyset$ until it's spanning V. Suppose V is not finitely generated, let's define F as the set of all linearly independant sets of vectors. F is partially ordered by the order of inclusion of sets. Let $C = (A_i)_{i\in I}$ be a chain in F, $A = \bigcup_{i\in I} A_i$. A is clearly a maximal element of the chain. Let's prove it is in F. Assume A isn't in $F \Rightarrow$ there exists a finite series of linearly dependent vectors, each is an element of a finite series of elements of C. Since that series is finite, and linearly ordered as a subset of C, There exists a maximal element that must contain all the vectors in the linearly independent vector series, but that element is in F so it's both linearly dependent and independent at the same time! contradiction! We get that $A \in F$ so by Zorn's lemma F has a maximal element T. That T is our basis. $\square$

11.2 Comparing Cardinals

Proposition 11.3. *Let α, β be cardinals other than 0. Then either $\alpha \le \beta$ or $\beta \le \alpha$.*

Proof. Let A, B be two sets of cardinality α, β accordingly. Define F to be the set of all ordered pairs (X, f) such that $f\colon X \to B$ is an injective function ($X \subseteq A$). Now we'll define an order in the following way:

$$(X_1, f_1) \le (X_2, f_2) \iff X_1 \subseteq X_2 \wedge f_2|_{X_1} = f_1$$

Let $C = ((X_1, f_1), (X_2, f_2), \ldots)$ be a chain in F, and let $(X, g) = (\bigcup A_i, \bigcup f_i)$ this implies

$$\forall i((X_i, f_i) \le (X, g))$$

Assume g isn't a function, we get $(x, y), (x, z) \in G$

$$\exists i, j \text{ such that: } f_i(x) = y,\ f_j(x) = z$$

Since C is a chain so we without lose of generailty we get:

$$\begin{aligned} &f_i \le f_j \\ \Rightarrow &f_j|_{X_i} = f_i \\ \Rightarrow &f_i(x) = f_j(x) \\ \Rightarrow &y = z \end{aligned}$$

That means g is a function, and since it's a union of injective functions, it must also be injective. That means it's in F and using Zorn's lemma we get a maximal element in F, which we denote (D, h). If $D = X$ then h is injective and we get $A \le B$. If it's not, it must be surjective or we get a contradiction to (D, h)'s maximality and thus $B \le A$.
We can also prove that $\alpha + \alpha = \alpha$. We know that $\alpha + \alpha = 2\alpha$ so we will just prove $\alpha = 2\alpha$. We'll build F using bijections this time. Denote the maximal element $M = (X, g)$. If $|X| = 2\alpha$ We finished, else we get that there's a set of size $\aleph_0$ that can be mapped "bijectively" to the set of 2α contradicting M's maximality. $\square$

11.3 Corollaries

Here are a couple more coroallaries of Zorn's lemma.

For any cardinals α, β the following hold:

1. $\alpha + \beta = \max\{\alpha, \beta\}$
2. $|A \setminus B| = |A| \iff |B| \leq |A|$
3. $\alpha * \alpha = \alpha$ (not a direct corollary)
4. $\alpha^\alpha = 2^\alpha$

12 More Axioms In ZF

Axiom of extensionality:

$$\forall A \forall B (\forall x (x \in A \iff x \in B) \Rightarrow A = B)$$

Axiom of union:

$$\forall A \exists B \forall c (c \in B \iff \exists D (c \in D \wedge D \in A))$$

Axiom of infinity:

$$\exists I (\emptyset \in I \wedge \forall x \in I (x \cup \{x\} \in I))$$

Axiom of power set:

$$\forall x \exists y \forall z [z \in y \iff \forall w (w \in z \Rightarrow w \in x)]$$

Axiom of regularity:

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

Axiom of pairing:

$$\forall A \forall B \exists C \forall D [D \in C \iff (D = A \wedge D = B)]$$

Axiom schema of specification - *any definable subclass of a set is a set.*
Axiom schema of replacement - *the image of any set under any definable mapping is also a set*

13 Axiom of Choice

Definition 13.1 (Choice function)**.** A choice function f is a function from an indexed family of sets $(S_i)_{i \in I}$ to their union such that for every $i \in I$ f satisfies $f(S_i) \in S_i$.

Definition 13.2 (Axiom of Choice)**.** The axiom of choice (also known as AoC) is described in first order logic as:

$$\forall X \left[\emptyset \notin X \to \exists f \colon X \to \bigcup X \text{ s.t. } \forall A \in X (f(A) \in A) \right]$$

In other words, every family of sets that does not include the empty set has a choice function.

13.1 Nomenclature

Z - *The first seven axioms*
ZF - *Zermelo-Fraenkel set theory. Z + Axiom of replacement*
AC - *Axiom of Choice*
ZFC - *ZF extended to include AC*

14 Measure

Measure theory is complex and goes well beond what I can show in this section but let's talk about it anyway. A measure is a way to generalize the length, volume, and such for sets. Let X be a set and Σ a σ-algebra over X. A set function μ from Σ to the extended real number line is called a measure if:

- for all $E \in \Sigma$ we have $\mu(E) \geq 0$;
- $\mu(\emptyset) = 0$;
- (σ-additivity) for all countable collections $\{E_k\}_{k=1}^{\infty}$ of pairwise disjoint sets in Σ,

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

If the condition of non-negativity is dropped then μ is called a signed measure. The pair (X, Σ) is called a measurable space, and the members of Σ are called measurable sets. A triple (X, Σ, μ) is called a measure space. A probability measure is a measure with total measure $\mu(X) = 1$. A probability space is a measure space with a probability measure.

14.1 Lebesgue Measure

Now we have to simplify things so we'll consider only Lebesgue measure of bounded sets on the real number line. First, if the set is of the form $X = (a, b)$ or $X = [a, b]$ and such the measure must satisfy $\mu(X) = b - a$, if $Y = f(X)$ and f is an isometric function then $\mu(X) - \mu(Y)$. Denote $\mathcal{Y} = (Y_i)_{i \in I}$ such that $|I| < \aleph_0$ and $X \subseteq \bigcup_{i \in I} Y_i$ and each Y_i is an interval on R. Denote $s(\mathcal{Y})$ The sum of lengths of intervals in $\mathcal{Y}$. The outer measure of X is

$$\mu^*(X) = \inf_{\mathcal{Y}} s(\mathcal{Y})$$

And the inner measure is defined with an interval $X \subseteq [a, b]$ and $X' = [a, b] \setminus X$

$$\mu_*(X) = (b - a) - \mu^*(X')$$

It's easy to show that the outer measure is always greater than the inner measure. The Lebesgue measure is defined if they are equal and is equal to

$$\mu_*(X) = \mu^*(X) = \mu(X)$$

15 Well Order

Definition 15.1 (Well ordered set)**.** A partially ordered set $(X, \leq)$ is said to be well ordered if for every nonempty subset $S \subseteq X$ there exists a minimal element $b \in S$.

Proposition 15.1. *Every finite totally ordered set is well ordered.*

Proposition 15.2. *If $\leq$ is a well order then it's a linear order as well.*

Proposition 15.3. *Let $(X, \leq)$ be a linearly ordered set. It's well ordered if and only if it doesn't include an infinite decreasing series.*

We now proceed to define two very similar terms.

Definition 15.2 (Risha)**.** Let X be a set. Then $A \subseteq X$ is said to be a risha of X if $x \in A$ and $y < x$ implies that $y \in A$.

Definition 15.3 (Initial segment)**.** Let X be a set, let $a \in X$. We say that

$$I_X(a) := \left\{x \in X \mid x < a\right\}$$

is the initial segement of a in X.

Remark 15.1. The interval $[0, 0.5]$ in $[0, 1] \in \mathbb{R}$ is a risha in $\mathbb{R}$ but not an initial segment.

Proposition 15.4. *Let X be a well oredered set. Prove that $A \subset X$ is a risha of X if and only if it is an initial segement of X.*

15.1 Some lemmas

1. let X be a woset, $f : X \to X$ a one-to-one homomorphism $\to \forall x \in X(x \leq f(x))$
2. let $(X, \leq_x) \cong (Y, \leq_y)$ be isomorphic wosets, there's only one unique isomorphism betweem them (proof using previous theorem)
3. in a woset X a risha can't be have an isomorphism with X
4. in wosets $I_x(a) \cong I_x(b) \Rightarrow a = b$
5. let $f : X \to Y$ be an isomorphism between wosets s.t. $y_0 = f(x_0) \Rightarrow I_x(x_0) = I_y(y_0)$

15.2 A Lemma About Partial Orders

If $(X, \leq_x), (Y, \leq_y)$ are partial orders, and $\leq_x$ is a total order, then if f is an inversible homomorphism it is also an isomorphism, and $\leq_y$ is a total order.

16 Comparison of Well Ordered Sets

If X, Y are wosets then exatcly one of the following is true:

1. $(X, \leq_x) \cong (Y, \leq_y)$
2. $\exists y_0 \in Y : (X, \leq_x) \cong (I_y(y_0), \leq_y)$
3. $\exists x_0 \in X : (Y, \leq_y) \cong (I_x(x_0), \leq_x)$

If $X = \emptyset$ or $Y = \emptyset$ the proof is trivial. Assuming they're not empty we'll define:

$$\begin{aligned}
&A = \{x \in X : \exists y \in Y(I_X(x) \cong I_Y(y))\}\\
&B = \{y \in Y : \exists x \in X(I_X(x) \cong I_Y(y))\}\\
&\phi\colon A \to B\\
&\phi(x) = y\colon I_X(x) \cong I_Y(y)
\end{aligned}$$

It is clear why ϕ is a bijection, we will show it's an isomorphism. Consider $a_1 < a_2 \in A$ and $\phi(a_1) = b_1, \phi(a_2) = b_2$. Since $I_X(a_2) \cong I_Y(b_2)$ we'll denote their isomorphism α. We get $a_1 < a_2 \Rightarrow a_1 \in \text{Dom}\alpha \Rightarrow \alpha(a_1) \in \text{Im}\alpha = I_Y(b_2) \Rightarrow \alpha(a_1) < b_2$. By one of our previous lemmas [0] $I_X(a_1) \cong I_Y(\alpha(a_1))$ and we know $I_X(a_1) \cong I_Y(b_1) \Rightarrow b_1 = \alpha(a_1)$. Recall that $\alpha(a_1) < b_2 \Rightarrow b_1 < b_2$. Since ϕ is a bijection and a homomorphism it's an isomorphism $\Rightarrow A \cong B$. By cases we'll get:

1. If $A = X, B = Y \Rightarrow (1)$.
2. If $B = Y \wedge A \subset X \neq \emptyset$ denote $A \setminus X$'s minimal element c and then $I_X(c) = A$ [1] $\Rightarrow (3)$[2]
3. If $A = X \wedge B \subset Y \neq \emptyset$ denote $Y \setminus B$ minimal element d and then $I_X(d) = B \Rightarrow (2)$
4. If $A \subset X \wedge B \subset Y \Rightarrow I_X(c) \cong A \wedge I_Y(d) \cong B. A \cong B \Rightarrow I_X(c) \cong I_Y(d) \Rightarrow c \in A$ but $c \notin A$ by our construction $\Rightarrow$ contradiction.

Now we'll show only one of $(1), (2), (3)$ can be true for any X, Y: $(2) + (3) \Rightarrow \exists \delta : X \to I_Y(d)$ isomorphism $\Rightarrow^0$ $I_X(c) \cong I_Y(\delta(c))$ and since we know $Y \cong I_X(c)$ we get that $Y \cong I_Y(\delta(c))$ which we know can't be. $(1) + (3)/(1) + (2)$ imply an initial segment of X/Y is isomorphic to X/Y and that can't be!

[0] Refer to 15.1.5
[1] Think why(two sided inclusion).
[2] since $A \cong B$

17 Ordinals

Ordinals are the generalization of ordinal numerals aimed to extend enumeration to infinite sets. Since sets don't have an innate order we only define ordinals on well ordered sets. It's easy to see that all well orders on finite sets are isomorphic, thus the following definition for ordinals makes sense:

Definition 17.1. Define the finite ordinals as such:

$$\begin{aligned}&k = ord(\{0, 1, \dots, k-1\}) = ord(I_{\mathbb{N}}(k))\\&ord(\emptyset) = 0\end{aligned}$$

Remark 17.1. We denote the ordinal of the natural numbers with the standard order as:

$$ord((\mathbb{N}, \leq_{\text{std}})) = \omega$$

By the comparability of wosets we can define an order on the ordinals as such:

$$\begin{aligned}ord(A) = ord(B) &\iff A \cong B\\ ord(A) < ord(B) &\iff A < B\\ ord(A) > ord(B) &\iff A > B\end{aligned}$$

Now we will define a new set function on ordinals $W(\alpha)$ which is defined as:

$$W(\alpha) = \{\beta \colon \beta < \alpha\}$$

Proposition 17.1. *The set $W(\alpha)$ is a well ordered set with ordinal comparison, and its ordinal is α.*

Proof. The proof's idea is by constructing an isomorphism from a set A with $ord(A) = \alpha$ as such:

$$\begin{aligned}&\phi \colon A \to W(\alpha)\\&\phi(a) = W(ord(I_A(a)))\end{aligned}$$

We get that ϕ is a bijection that preserves order and thus an isomorphism. By definition we get that the ordinal of $W(\alpha)$ is $ord(A) = \alpha$ as wanted. $\square$

Proposition 17.2. (ZFC) *Every set of ordinals X is well ordered.*

Proof. Let $\emptyset \neq A \subseteq X$. Since it is not empty we can choose $a \in A$. If a is minimal we are done. Otherwise, exists $\beta \in A$ such that $\beta < \alpha$ and thus $\beta \in W(\alpha) \cap A$. Since $W(\alpha) \cap A$ is a subset of a well ordered set A it is also well ordered and thus exists $\gamma \in W(\alpha) \cap A$ a first element. It is clear that γ is first in A and thus we are done. $\square$

Cesare Burali-Forti Paradox

The paradox states that the set of all ordinals it not well defined in ZFC. Suppose by contradiction it were a set O, then as a set of ordinals by the previous proposition it will be well ordered. Now denote $ord(O) = \alpha$ then $\alpha \in O$ which implies $W(O) \subseteq O$ thus an initial segment of the set is isomorphic to it which is a contradiction.

Russell's Paradox

Let R be the set than contains all the sets that don't contain themselves. If R contains itself, it must not contain itself. If R doesn't contain itself, then it must contain itself.

17.1 Types of Ordinals

There are two kinds of ordinals:

Definition 17.2. Successor Ordinals - Ordinals that immediatly success another ordinal

Definition 17.3. Limit Ordinals - Ordinals that don't immediatly success another ordinal

17.2 Ordinal Arthimetic

The following definitions are based on our motivation to extend enumeration to infinite sets.

17.2.1 addition

Let $(X, \leq_x), (Y, \leq_y)$ be two **disjoint** well ordered sets such that $(ord(A), ord(B)) = (\alpha, \beta)$. We denote $\left(X \bigcup Y, \leq\right)$ as:

$$a \leq b \iff \begin{Bmatrix} a, b \in X & a \leq_x b \\ a, b \in Y & a \leq_y b \\ a \in X & b \in Y \end{Bmatrix}$$

As $X \oplus Y$, and by definition $\alpha + \beta = ord(X \oplus Y)$. Notice that oridnals are associative but not commutative under addition:

- $n + \omega = \omega$
- $\alpha + 0 = \alpha$
- $\omega < \omega + 1 < \omega + 2 < \ldots < \omega + k < \ldots$

17.2.2 multiplication

Let $(X, \leq_x), (Y, \leq_y)$ be two **disjoint** wosets such that $(ord(A), ord(B)) = (\alpha, \beta)$. We denote:

$$(X \times Y, \leq_{dictionary})$$

As $X \odot Y$, and by definition $\alpha * \beta = ord(X \odot Y)$. It is possible to show $\omega = k\omega$ for $k \in \mathbb{N}$ by constructing an isomorphism:

$$\phi \colon \mathbb{N} \to \{0, 1, \ldots, k-1\} \times \mathbb{N}$$
$$\phi(n) = (\lfloor n/k \rfloor, n \bmod k)$$

- $\omega * 0 = 0$
- $\alpha * 1 = \alpha$
- $k\omega = \omega < \omega 2 = \omega + \omega < \omega 3 < \ldots < \omega k < \ldots < \omega^2$

Notice that ordinals are left distribuitive but not right distributive.

17.2.3 Powers

We define powers of ordinals as such:

$$\alpha^\gamma = \begin{pmatrix} 1 & \gamma = 0 \\ \alpha^{\gamma - 1} * \alpha & \gamma \text{ is a succesor ordinal} \\ \min_{\delta < \gamma}\{\mu : \alpha^\delta < \mu\} & \gamma \text{ is a limit ordinal} \end{pmatrix}$$

Notice that ordinals are usually expressed as polynomials in ω with natural coefficients. We have that $p(\omega) < q(\omega)$ for polynomials p, q if and only if $\deg p < \deg q$ or $\deg p = \deg q$ and the leading coefficient of $(q - p)(\omega)$ is strictly positive.

17.2.4 The Ordinals 2^ω and ω^ω

By our previous definition we can conclude that 2^ω is:

$$\begin{aligned}&\min_{\delta<\omega}\{\mu : 2^\delta < \mu\}\\ =&\min\{2^1, 2^2, \ldots, 2^k, \ldots\}\end{aligned}$$

Since this sequence doesn't have an upper bound the result is the smallest infinite ordinal or $2^\omega = \omega$.
By our previous definition we can conclude that ω^ω is:

$$\begin{aligned}&\min_{\delta<\omega}\{\mu : \omega^\delta < \mu\}\\ =&\min\{\omega^1, \omega^2, \ldots, \omega^k, \ldots\}\end{aligned}$$

Now consider the family of disjoint sets $\{X_n\}_{n\in\mathbb{N}}$ such that for any $n \in \mathbb{N}$ we have $ord(X_n) = \omega^n$. We define:

$$X := \bigoplus_{n\in\mathbb{N}} X_n$$

We can prove by definition that $ord(X) = \omega^\omega$ but also since this is a countable sum of countable sets we have that $|X| = \aleph_0$ and since we construct ω^{ω^ω} in the same way we can see that this is the case for any sets with ordinals ω^{ω^ω} and so on...

Remark 17.2. Notice that:

$$\omega^\omega < \omega^\omega + 1 < \omega^\omega + 2 < \ldots < \omega^{\omega^\omega} < \ldots$$

A good way to think about ω^ω for example is all finite sequences of natural numbers.

18 The Well Ordering Theorem

The Well Ordering Theorem [3] states that any set can be well ordered and is equivalent to Zorn's lemma and AC.

18.1 Proof From AC

Let the set we're trying to well-order be A and let f be a choice function for the family of non-empty subsets of A. Now for every ordinal α define:

$$\begin{cases} a_\alpha = f(A \setminus \{a_\xi | \xi < \alpha\}), & A \setminus \{a_\xi | \xi < \alpha\} \neq \emptyset \\ a_\alpha = \text{UNDEFINED}, & \text{otherwise} \end{cases}$$

Then

$$\langle a_\alpha | a_\alpha \text{ is defined} \rangle$$

Is a well order on A.

18.2 Proof of AC using WOT

To make a choice function for a collection of non-empty sets E, take the union of the sets in E and call it X. There exists a well-ordering of X; let R be such an ordering. The function that to each set S of E associates the smallest element of S, as ordered by (the restriction to S of) R, is a choice function for the collection E. [4]

It's worth noting the difference between choosing this one choice function here (R), and applying the WOT to all the sets $S \in E$ seperately, and choosing the minimal element in each set seperately. While the first is allowed under ZF since we're only making a single choice, the latter is not allowed when there are infinitely many elements in E without assuming the axiom of choice itself, and thus is not a valid way to prove AC.

[3] sometimes shortened to WOT

[4] This proof was taken straight from wikipedia.

19 The Continuum Conjecture and the Alephs

19.1 The Continuum Conjecture

Let X be a set of cardinality $\aleph$ endowed with a well order from the well order theorem, and consider the set of all initial segments with cardinality greater than $\aleph_0$. If this set is empty we can move the first element in X to the "end" of the set and make it nonempty. Since X is well ordered, we know that exists a minimal element $a \in X$ such that $\aleph_0 < |I_X(a)|$. We denote $|I_A(a)| = \aleph_1$ and $ord(I_A(a)) = \Omega$, the first uncountable ordinal. The conjecture is that $\aleph = \aleph_1$.

Remark 19.1. This conjecture was proven to be independent of the axioms of ZFC.

19.2 The Alephs

Similarly we can construct a set X with cardinality $2^{\aleph_1}$ with a well order and choose the first element $a \in X$ with $|I_X(a)| > \aleph_1$ and denote $|I_X(a)| = \aleph_2$. In this way we can define $\aleph_n$ for every $n \in \mathbb{N}$. To construct $\aleph_\omega$ we can take the family of disjoint sets $\{X_n\}_{n\in\mathbb{N}}$ such that $|X_n| = \aleph_n$ and define:

$$X := \bigcup_{n\in\mathbb{N}} X_n$$

We can define on X a well order such that exist elements $x \in X$ such that $I_X(x) > \aleph_n$ for all $n \in \mathbb{N}$. Let $a \in X$ be the first element with this property. We denote $|a| = \aleph_\omega$ and we can see that:

$$\aleph_\omega \leq \sum_{n\in\mathbb{N}} \aleph_n$$

As expected we can keep defining in this way $\aleph_{\omega+1}, \aleph_{\omega+2}, \ldots, \aleph_{\omega 2}, \ldots, \aleph_{\omega^2}, \ldots, \aleph_{\omega^\omega}, \ldots$ and after all countable ordinals we will reach $\aleph_\Omega$, the $\aleph_1$-th aleph, and the first one with an uncountable ordinal. This construction allows to consider these cardinals as a subclass of ordinals with the association rule $\alpha \to \aleph_\alpha$ from ordinals to cardinals. This also implies that the set of these cardinals and the finite cardinals is also well ordered.

Remark 19.2. The generalized Continuum Conjecture is:

$$2^{\aleph_\alpha} = \aleph_{\alpha+1}$$

Results About Alephs

Proposition 19.1. *The cardinal $\aleph$ is not a countable sum of cardinals that are smaller than itself.*

Proof. Using König's inequalities for the family $\{a_n\}_{n\in\mathbb{N}}$ such that $a_n < \aleph$ for every $n \in \mathbb{N}$ we get:

$$\sum_{n\in\mathbb{N}} a_n < \prod_{n\in\mathbb{N}} \aleph = \aleph^{\aleph_0} = \aleph$$

$\square$

Corollary 19.2. $\aleph_\omega \neq \aleph$.

Proof. Assume by contradiction that $\aleph_\omega = \aleph$ then since $\aleph_n < \aleph_\omega = \aleph$ for every $n \in \mathbb{N}$ we have that $\sum_{n\in\mathbb{N}} \aleph_n < \aleph = \aleph_\omega$ which is a contradiction by the construction of $\aleph_\omega$ so the assumption is false and we are done. $\square$

Consider the inequlity:

$$q < q^{\aleph_0}$$

It is clear that every countable satisfies it, but cardinals like $\aleph, 2^\aleph$ don't. Let c be a infinite cardinal. We define the sequence:

$$a_1 = c < \underbrace{2^c}_{a_2} < \underbrace{2^{c^c}}_{a_3} < \dots$$

We see that:

$$d := \sum_{n\in\mathbb{N}} a_n < \prod_{n\in\mathbb{N}} c_n$$

For example by adding 0 to the sum before the sequence. On the other hand we have $c_n \leq d$ for every $n \in \mathbb{N}$ so we have:

$$\prod_{n\in\mathbb{N}} c_n \leq d^{\aleph_0}$$

And thus finally we have $d < d^{\aleph_0}$ as wanted.

20 Transfinite Induction

Transfinite induction is an extension of mathematical induction to well-ordered sets, for example to sets of ordinal numbers or cardinal numbers. Its correctness is a theorem of ZFC. When given X a set, and P the set of elements who have a certain property, the principle that's derived by the WOT can be written like this:

$$\forall x \in X(\forall y < x(y \in P)) \Rightarrow x \in P$$

The steps of transfinite induction:

1. the 0 case ($0 \in P$)
2. The succesor ordinal case ($\alpha \in P \Rightarrow \alpha + 1 \in P$)
3. The limit ordinal case case ($\forall \beta < \gamma(\beta \in P) \Rightarrow \gamma \in P$)

20.1 Proof That The Only Isomorphism from a Well-Ordered Set to Itself is the Identity Isomorphism

Consider the propety P that "*this element is transformed to itself under all iso- morphisms*". Now consider an element a such that all elements that are lesser than a are in P . This can always be done by choosing the minimal element by WOT. a can't be transformed to an element lesser than a because than the isomorphism won't be injective, and also not to an element greater than it, because a must also have a source, since the isomorphism is surjective, but then we get a contradiction to the the fact the isomorphism is a homomorphism.

Remark 20.1. transfinite induction works because of WOT but there are of course sets like $\mathbb{R}$ with normal ordering that isn't a woset so we can't use transfinite induction on it. A counter example for our proof may be $f(x) = x + 1$

21 Extras

21.1 A Bit About Constructions

Constructions of sets are the way to formally define sets like $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

21.1.1 Construction of $\mathbb{N}$

There are multiple ways [5] to define $\mathbb{N}$ one in ZF is recursivly defining the natural numbers as such:

$$\begin{aligned} 0 &= \{\} = \emptyset \\ 1 &= \{0\} = \{\{\}\} \\ 2 &= \{0, 1\} = \{\emptyset, \{\emptyset\}\} \end{aligned}$$

And so on defining numbers unsing the succesor function $S(n) = n \cup \{n\}$. $\mathbb{N}$ is the smallest set containing 0 and closed under $S(n)$

21.1.2 Construction of $\mathbb{Z}$

$\mathbb{Z}$ was constructed as ${}^{\mathbb{N}\times\mathbb{N}}/_{R}$

$$\langle a, b\rangle \, R \, \langle c, d\rangle \iff c - d = a - b$$

For example $-1 = \{\langle 2, 3\rangle, \langle 5, 6\rangle, \ldots, \langle n, n+1\rangle\}$
Define $\mathbb{Z}_+, \mathbb{Z}_*$

21.1.3 Construction of $\mathbb{Q}$

$\mathbb{Z}$ was constructed as $\mathbb{Z} \times \mathbb{Z}'$[6]$/R$

$$\langle a, b\rangle \, R \, \langle c, d\rangle \iff ad = cb$$

For example $\frac{1}{2} = \{\langle 1, 2\rangle, \langle -2, -4\rangle, \ldots, \langle n, 2n\rangle\}$
Try defining $\mathbb{Q}_+, \mathbb{Q}_*$

21.1.4 About the Construction of $\mathbb{R}$

The construction of $\mathbb{R}$ is more difficult than you may expect. It should be studied in a number theory course, and is usually very unrigoursly defined as all numbers in the interval $(-\infty, \infty)$

21.2 Discrete Functions

Discrete Function - A function that is defined only for a set of numbers that can be listed, such as the set of whole numbers or the set of integers.

[5] One of them is by isomorphism classes of finite sets
[6] $\mathbb{Z}' = \mathbb{Z} \setminus \{0\}$

21.3 More definitions

21.3.1 Saturated dunctions

For a function $f : X \to Y$

$$\forall A \subseteq X, f_\star(A) = \{f(x) : x \in A(A \in P(X))\}$$
$$\forall B \subseteq X, f^\star(A) = \{x : f(x) \in B(B \in P(Y))\}$$

21.3.2 Hasse diagrams

Hasse diagrams represent posets. For example the Hasse Diagram of the the set $\{1,2,3\} \times \{1,2\}$ with the standard order is:

$$\begin{tikzcd}[column sep=tiny, row sep=small]
 & & (3,2) \arrow[dl, dash] \arrow[dr, dash] & \\
 & (2,2) \arrow[dl, dash] \arrow[dr, dash] & & (3,1) \arrow[dl, dash] \\
(1,2) \arrow[dr, dash] & & (2,1) \arrow[dl, dash] & \\
 & (1,1) & &
\end{tikzcd}$$

21.3.3 Some denotions

- A *singleton* is a set containing only one element.
- $\mathcal{P}(A) := \{B \colon B \subseteq A\}$.
- $A \triangle B := \{A \cup B\} \setminus \{A \cap B\}$.
- $|\mathbb{R}| = c = \beth_1 = \aleph$.
- $A^c = \big\{b \colon b \notin A\big\}$.
- $\prod_{i \in I} X_i = \big\{f \colon I \to \bigcup_{i \in I} X_i \mid f(i) \in X_i \text{ for all } i \in I\big\}$.
- A pairing function is a bijection $\pi \colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.
- The indicator function of $A \subseteq X$ is defined as $1_A(x) = I_A(x) = \chi_A(x) = 1 \iff x$ is in A and equals 0 otherwise.