MATH 316: SET THEORY

TOM BENHAMOU
RUTGERS UNIVERSITY

Set theory is the basic mathematical theory commonly accepted as the language to describe the mathematical universe. Our first goal in this course is to develop all regular mathematical objects (such as numbers, functions, vector spaces, graphs etc.) from the basic concept of a set. The second goal is to present some results in set theory which relates to the infinite.

We will use the first-order language of set theory (whatever that means) which includes the following symbols:

(1) Equality $=$.
(2) Negation $\neg A$ ("not A").
(3) Conjunction $A \wedge B$ ("A and B").
(4) Disjunction $A \vee B$ ("A or B").
(5) Implication $A \Rightarrow B$ (" If A then B").
(6) Equivalence $A \Leftrightarrow B$ (" A if and only if B").
(7) Universal quantifier $\forall x,\ \phi(x)$ ("For all x, $\phi(x)$").
(8) Existential quantifier $\exists x,\ \phi(x)$ ("there exists some x such that $\phi(x)$").
(9) The membership relation $\in$ ($a \in A$ means "a is a member in the set A")
(10) Uniqueness quantifier $\exists! x,\ \phi(x)$ ("There is a unique x such that $\phi(x)$")
(11) Bounded quantifiers: it will be convenient to use the notion of quantifiers which are bounded in a given set A:
    (a) $\forall x \in A, \phi(x)$ ("for every x in the set A, $\phi(x)$"). This is equivalent to $\forall x, x \in A \to \phi(x)$.
    (b) $\exists x \in A, \phi(x)$ ("there exists an element x in the set A such that $\phi(x)$"). This is equivalent to $\exists x,\ x \in A \wedge \phi(x)$.

    We think of these quantifiers as quatifiters which range over a given set.

1. SETS

In the next section we shall lay the formal foundations of set theory. Before that, we would like to gain intuition about sets. We shall give rules of thumb (which will later turn into Axioms) to describe sets and how one can think of and handle them.

Date: September 20, 2023.

Definition 1.1 (Non-formal)**.** A *set* is a collection of mathematical objects without repetitions and without ordering.

To understand this definition better, let us jump directly to the description of sets and through the example we will understand it better.

1.1. **Defining sets.** In general, there are exactly three ways to define a set.

1.1.1. *The list principle.*

$$\{a, b, c, ..., z\},\ \{1, 5, 17\},\ \{\{1, 2\}, \{2, 3\}\}$$

A set is always denoted with curly brackets $\{,\}$. Between the brackets we specify the *members* or *elements* of the set separated by commas.

Let us denote the set of *natural numbers* by:

$$\mathbb{N} = \{0, 1, 2, ...\}$$

Definition 1.2 (Non-formal)**.** The membership relation $a \in A$ is the statement that the object a is a member of the set A

Example 1.3. $1 \in \{1\}, \{2, 2\} \in \{\{1\}, \{2\}\}, \{3\} \notin \{3, 4, 5\}, \{1, 10, 100\} \ni 1, \frac{1}{2} \notin \mathbb{N}$

Formally, we can define the "List Principle" by

$$a \in \{a_1, ..., a_n\} \equiv a = a_1 \vee a = a_2 ... \vee a = a_n$$

Remark 1.4.
1. To explain the fact that sets have no order, we note that the sets $\{1, 2, 3\}, \{2, 3, 1\}$ represent the same set.
2. To explain the fact that sets have no repetitions, we note that $\{1, 1, 2, 3\}, \{1, 2, 3\}$ represent the same set.

Remember: The membership relation is always between a member of a set and a set

1.1.2. *The separation principle.* Given a set A and a predicate $p(x)$ (a first order formula) where x is a free variable in the set A, we can *separate* from A the elements $a \in A$ which satisfy $p(a)$ into a new set. This separated set is denoted by:

$$\{x \in A \mid p(x)\}$$

This reads as "the set of all x in A such that $p(x)$ holds true". The *Axiom* of separation states that such a set always exists.

Example 1.5.
1. $\{x \in \{1, 2, 6, 7\} \mid x > 3\} = \{6, 7\}$
2. $p(x)$ is the predicate $\exists k \in \mathbb{N}(3 \cdot k = x)$. Then we can separate from $\mathbb{N}$ the following set:
$$\{x \in \mathbb{N} \mid \exists k \in \mathbb{N}(3 \cdot k = x)\} = \{0, 3, 6, 9, ...\}$$
3. $A = \{1, 3, 6, 11, 21, 17\}$, $\{x \in A \mid x + 1 \textit{ is prime}\} = \{1, 6\}$

(4) $B = \{\{1\}, \{2\}, \mathbb{N}, \{\mathbb{N}\}, \{x \in \mathbb{N} \mid x \cdot x = x\}\}$

$$\{x \in B \mid 1 \notin x\} = \{\{2\}, \{\mathbb{N}\}\}$$

Define $a \in \{x \in A \mid p(x)\} \equiv a \in A \wedge p(a)$

1.1.3. *The replacement principle.* Let A be a set and $f(x)$ some operation/ function on the elements of A. We can *replace* every memeber a of the set A by the outcome of the operation $f(a)$ and collect all the outcomes into a new set. This new collection is denoted by:

$$\{f(x) \mid x \in A\}$$

This reads as "the set of all outcomes $f(x)$ where the parameter x runs in the set A".

Example 1.6.
- $f(x) = 2^x$, $\{2^x \mid x \in \mathbb{N}\} = \{2^0, 2^1, 2^2, ...\} = \{1, 2, 4, 8, 16, ...\}$
- $\{\{x\} \mid x \in \{1, 4, 3\}\} = \{\{1\}, \{3\}, \{4\}\}$. Sets of the form $\{a\}$ are called *singletons.*
- $\{x + 1 \mid x \in \mathbb{N}\} = \{x \in \mathbb{N} \mid x > 0\}$

Define $a \in \{f(x) \mid x \in A\} \equiv \exists x \in A.f(x) = a$

Important: a formula of the form $a \in A$ is a **statement** and should be proven by the definitions given above for each of the three principles.

Exercise 1. *Prove the following membership statements:*

(1) $2 + 5 \in \{1, 2, ..., 10\}$.

Solution 1. *By the list principle, we need to prove that*

$$(2 + 5 = 1) \vee (2 + 5 = 2) \vee ... \vee (2 + 5 = 10)$$

Indeed, $2 + 5 = 7$ *hence the* $\vee$*-statement holds.*

(2) $5 \in \{x \in \mathbb{N} \mid \exists y \in \mathbb{Z}.y + x = 5\}$.

Solution 2. *By the separation principle, we need to prove that* $5 \in \mathbb{N} \wedge \exists y \in \mathbb{Z}.y + 5 = 5$. *This is a* $\wedge$*-statement, so we need to prove two parts:*

(a) $5 \in \mathbb{N}$, *this is clear by the definition of the natural numbers.*
(b) We need to prove that $\exists y \in \mathbb{Z}.y + 5 = 5$. *Define* $y = 0$, *then* $y \in \mathbb{Z}$ *and* $y + 5 = 0 + 5 = 5$.

(3) $\{1\} \in \{\{n, 1\} \mid n \in \mathbb{N}\}$.

Solution 3. *By the replacement principle, we need to prove that* $\exists n \in \mathbb{N}.\{1\} = \{1, n\}$. *Define* $n = 1$, *indeed* $1 \in \mathbb{N}$ *and since there are no repetitions in sets we have that*

$$\{1, n\} = \{1, 1\} = \{1\}.$$

1.1.4. *Celebrity sets.*

(1) $\mathbb{N} = \{0, 1, 2, 3, ...\}$ you will not need to explain basic properties of the natural numbers which relates to addition, multiplication and power of natural numbers. Here are some other properties we assume about the natural numbers:
   - Every natural number has an immediate successor.
   - The natural numbers are *well-ordered*, which simply says that every set of natural numbers (finite or infinite) has a minimal element.
   - Every finite set of natural numbers has a maximal element.

(2) The set of positive natural numbers is: $\mathbb{N}_+ = \{x \in \mathbb{N} \mid x > 0\} = \{1, 2, 3, 4, ...\}$

(3) The set of integers is: $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$

(4) The set of fractions/ rational numbers is: $\mathbb{Q} = \{\frac{m}{n} \mid m, n \in \mathbb{Z} \wedge n \neq 0\}$

(5) The set of real numbers is denoted by $\mathbb{R}$. We will formally define the reals only later. We will simply describe them as numbers which have a (possibly infinite) decimal representation such as: 15.6755897847566372....... Among the real numbers, one can find $\sqrt{2}, \pi, e$. One of the most important properties of the reals is that the rational numbers are dense inside them:
$$\forall r_1, r_2 \in \mathbb{R}. r_1 < r_2 \Rightarrow (\exists q \in \mathbb{Q}. r_1 < q < r_2)$$
$$\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}.$$

(6) The intervals:
   - $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ denotes the *open interval* between a and b.
   - $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ the *closed interval.*
   - $[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$. Define similarly $(a, b]$.
   - $(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$ is the *infinite ray*. Similarly define $[a, \infty), (-\infty, a), (-\infty, a]$. Note that $(a, \infty]$ is not defined since ∞ is not a natural number.

(7) $\emptyset$ denoted the empty set, which is characterized by the following property: $\forall x. x \notin \emptyset$. Namely, the empty set is a set with no element. It is sometimes convenient to think of $\emptyset = \{\}$.

1.1.5. *Axiomatic development- Existence, Extensionality and Comprehension.* The reason the need of formal mathematics as emerged is the discovery of certain paradoxes.

Russel's Paradox The paradox emphasizes the fact that if we are not being careful regarding what might be considered a mathematical object (i.e. set) then we will run into paradoxes. More specifically, consider the collection of

$$\{x \mid x \notin x\}$$

Namely the collection of all sets which are not members of themselves. For example: $\emptyset \notin \emptyset$ and therefore $\emptyset$ is part of this collection. Also $\mathbb{N} \notin \mathbb{N}$ since $\mathbb{N}$ consists only of the natural numbers and $\mathbb{N}$ is not a natural number. In general, it is not clear if there is even a set x such that $x \in x$. Russel's paradox says that this collection cannot form a set:

Proof. Suppose otherwise, that $D = \{x \mid x \notin x\}$ is a set. Then there are exactly two options:

(1) $D \in D$.
(2) $D \notin D$.

In the first case, the set D itself (when thinking of $x = D$ in the definition of D) does not satisfy the condition $x \notin x$ and thus D is not a member of D, a contradiction.

In the second case, D satisfies the condition $x \notin x$ and therefore D must be a member of the set D, namely $D \in D$, contradiction. □

We shall see how the formal approach resolves this paradox.

We need to start with something.

Axiom (Ax0. Existence)**.** There exists some object x.

From the set existence we can only prove that the universe is non-empty and it can contain a single object (which is not enough for any interesting mathematics). Axioms which generate new sets will be presented later on.

Axiom (Ax1. Extensionality)**.** For every x, y we have that $x = y$ if and only if $\forall z, z \in x \leftrightarrow x \in y$.

The axiom of extensionality is not contributing to the existence of new sets. It is used usually to prove uniqueness.

Example 1.7. Let us claim that from extensionality, if there is a set x such that $\forall z.z \notin x$ then x is unique (this claim basically sais that the empty set is unique).

Proof. Suppose that x_1, x_2both satisfy that for all $\forall z.z \notin x_i$ $(i = 1, 2)$, then the antecedent $\forall z.z \in x_1 \leftrightarrow z \in x_2$ is satisfied and therefore $x_1 = x_2$. □

Once we prove that there is a unique set satisfying a certain property we may introduce a special notion for it and use it from now on.

Definition 1.8. The empty set, denoted by ϕ is the unique set satisfying $\forall z,\ z \notin \phi$.

Axiom (Ax3. Comprehension scheme)**.** For every set A and every first-order formula $\phi(x)$, there is a set B such that $\forall z.z \in B \leftrightarrow z \in A \wedge \phi(z)$.

Definition 1.9. We denote the set B from $Ax3$ by $B := \{x \in A \mid \phi(x)\}$.

By the axiom of extensionality, given a set A and a formula $\phi(x)$, the set B is unique and therefore the definition above is legitimate. Now Russel's paradox is just the theorem that a certain set does not exists

Theorem 1.10 (Russel's Paradox)**.** *There is no set A such that*

$$\text{For every } x,\ x \in A \text{ if and only if } x \notin x$$

Corollary 1.11. *There is no set A such that $\forall x,\ x \in A$.*

Proof. Just otherwise, from the set A, using the axiom of comprehension the set from Russel's paradox exists, contradicting the previous theorem. □

Remark 1.12. The general principle of replacement will be given only later. However, if C is a set defined by replacement and for some set B, $C := \{f(x) \mid x \in A\} \subseteq B$ then C can also be defined using comprehension $C = \{x \in B \mid \exists a \in A,\ f(a) = x\}$. This is of course not always possible, since otherwise, the axiom of replacement would have been redundant.

1.2. **Inclusion and set equality.**

Definition 1.13. Let A, B be any sets. We say that A *is included in* B and denote it by $A \subseteq B$ if

$$\forall x. x \in A \Rightarrow x \in B$$

In other words, if every element of A is an element of B. Using bounded quantifiers we can say that $A \subseteq B$ is the statement $\forall x \in A. x \in B$.

Example 1.14. $\{1, 5\} \subseteq \mathbb{N}_{odd} \subseteq \mathbb{N}_{+} \subseteq \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$.

1.3. **Proving sets inclusion.** Since $A \subseteq B$ is a universal implication, we have the following format:

(1) The proof starts with "Let $a \in A$".
(2) Then we should deduce from the assumption of that $a \in A$ usually that requires to interpret that assumption that $a \in A$. that $a \in B$ and the proof should terminate by "$a \in B$".

Of course, in special cases we can use the other methods of proving universal statements (such as proving $a \in B$ going over $a \in A$ one-by-one)

Example 1.15. Prove the following inclusions:

(1) $\{2, -1\} \subseteq \{x \in \mathbb{Z} \mid x^2 > x\}$.

   Proof. Let $a \in \{2, -1\}$. Since $\{2, -1\}$ includes only two elements, let us prove that $a \in B$ by going over the elements of $\{2, -1\}$ one-by-one:
   (a) For $a = 2$, we need to prove that $2 \in \{x \in \mathbb{Z} \mid x^2 > x\}$. By the separation principle we need to prove that $2 \in \mathbb{Z} \wedge 2^2 > 2$. Indeed 2 is an integer and $2^2 = 4 > 2$.
   (b) For $a = -1$, we need to prove that $-1 \in \mathbb{Z} \wedge (-1)^2 > -1$. Indeed, -1 is an integer and $(-1)^2 = 1 > -1$.

   □

(2) $\{n^2 + n \mid n \in \mathbb{N}\} \subseteq \mathbb{N}_{even}$.

Proof. Let $x \in \{n^2 + n \mid n \in \mathbb{N}\}$. We need to prove that $x \in \mathbb{N}_{even}$. By the replacement principle, there exist $n \in \mathbb{N}$ such that $x = n^2+n$, so let $n_0 \in \mathbb{N}$ be such that $x = n_0^2 + n_0$. In is an easy exercise to deduce now that x is even, namely that $x \in \mathbb{N}_{even}$. $\square$

(3) For every $a, b, c \in \mathbb{R}$. If $a < b < c$, then there is $\epsilon > 0$ such that $(a, b+\epsilon] \subseteq (a, c)$.

Proof. Let $a, b, c \in \mathbb{R}$ such that $a < b < c$. We need to prove that there is $\epsilon > 0$ such that $(a, b+\epsilon] \subseteq (a, c)$A moment of reflection reviles that we only need to find $0 < \epsilon$ such that $b + \epsilon < c$, hence $0 < \epsilon < c - b$. The following definition of ϵ is tailored to satisfy exactly these inequalities.. Define $\epsilon = \frac{c-b}{2}$. Since $c > b$, we have that $c - b > 0$ and also $\epsilon = \frac{c-b}{2} > 0$. Let us prove that[1] $(a, b+\epsilon] \subseteq (a, c)$. This is an inclusion, let $x \in (a, b+\epsilon]$. By definition of intervals, this means that $x \in \mathbb{R} \wedge (a < x \leq b+\epsilon)$. We need to prove that $x \in (a, c)$, namely, that $x \in \mathbb{R} \wedge (a < x < c)$. Indeed by the assumption, $x \in \mathbb{R}$, and $a < x$. To see that $x < c$, we use the definition of ϵ:

$$x \leq b + \epsilon = b + \frac{c-b}{2} < b + (c - b) = c$$

Hence $a < x < c$ and we conclude that $x \in (a, c)$. $\square$

Problem 1. *Prove that if $A \subseteq B \wedge B \subseteq C$, then $A \subseteq C$.*

Theorem 1.16. *For every set A, $\emptyset \subseteq A$.*

Proof. [2] Let A be a set. We need to prove that $\emptyset \subseteq A$. Note here the assumption "Let $a \in \emptyset$" is impossible. Instead, we recall that in order to prove that $\emptyset \subseteq A$ we need to prove that $\forall x. x \in \emptyset \Rightarrow x \in A$. Let x be any element, then $x \in \emptyset$ is false by the definition of $\emptyset$ and therefore the implication $x \in \emptyset \Rightarrow x \in A$ is vacuously true. $\square$

Definition 1.17. We denote by $A \nsubseteq B$ if $\neg(A \subseteq B)$, namely, if $\exists x \in A. x \notin B$.

Example 1.18. Prove that $\{n \in \mathbb{N} \mid n^2 - 7n + 12 = 0\} \nsubseteq \mathbb{N}_{odd}$

Proof. For example[3] $4 \notin \mathbb{N}_{odd}$ and also $4 \in \{n \in \mathbb{N} \mid n^2 - 7n + 12\}$, since $4 \in \mathbb{N}$ and $4^2 - 7 \cdot 4 + 12 = 0$. $\square$

1.4. **Set equality.** The extensionality axion expresses the fact the a set is determined by its elements.

Corollary 1.19. *For any two sets A, B:*

$$A = B \Leftrightarrow (A \subseteq B) \wedge (B \subseteq A)$$

[1] Recall that to prove an existential statement we give the example and prove it satisfy the desired property.

[2] Here is and example for the 0.1% of the cases where we prove that an implication is vacuously true.

[3] We need to prove an existential statement so we provide an example.

This means that when we wish to prove set equality $A = B$, we do so by proving a *double inclusion*:

(1) Prove $A \subseteq B$.
(2) Prove $B \subseteq A$.

Example 1.20. Prove that $\mathbb{N}_+ = \{x \in \mathbb{Z} \mid \exists y \in \mathbb{N}.y + 1 = x\}$.

Proof. Let us denote the set of the right-hand side by A. We want to prove $\mathbb{N}_+ = A$. This is sets equality and we prove it by a double inclusion:

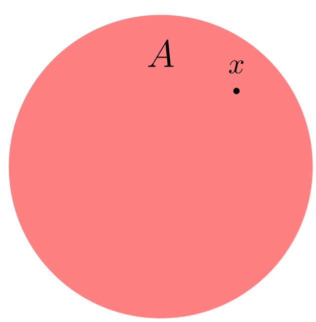
(1) $\underline{\mathbb{N}_+ \subseteq A}$: Let $n_0 \in \mathbb{N}^+$, then $n_0 \geq 1$ is an integer. We want to prove that $n_0 \in A$, by the separation principle, we want to prove that $n_0 \in \mathbb{Z} \wedge \exists y \in \mathbb{N}.y + 1 = n_0$. Clearly, $n_0 \in \mathbb{Z}$. Define $y = n_0 - 1$. Note that $y \geq 0$ is an integer, hence $y \in \mathbb{N}$ and clearly $y + 1 = n_0$, hence $n_0 \in A$.
(2) $\underline{A \subseteq \mathbb{N}_+}$: Let $a_0 \in A$. We want to show that $a_0 \in \mathbb{N}_+$. By the separation principle, we know that $a_0 \in \mathbb{Z}$ and that $\exists y \in \mathbb{N}.y + 1 = a_0$. Let $y_0 \in \mathbb{N}$ witness that $y_0 + 1 = a_0$. Since $y_0 \in \mathbb{N}$, we have that $y_0 \geq 0$ and therefore $a_0 = y_0 + 1 \geq 1$. It follows that $a_0 \in \mathbb{N}_+$.

$\square$

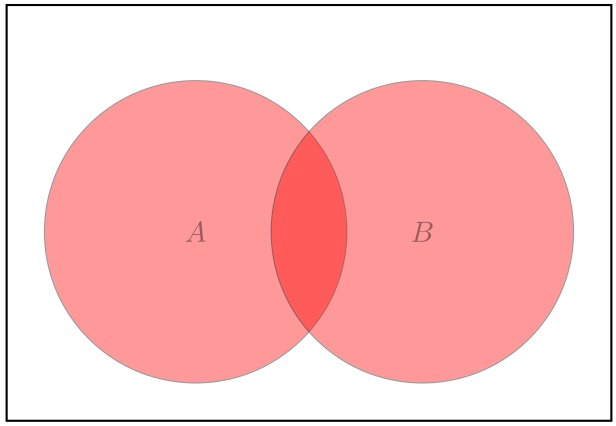
Definition 1.21. We denote $A \neq B$ is $\neg(A = B)$. This is equivalent to $\exists x.\ (x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)$. We denote by $A \subsetneq B$ if $(A \subseteq B) \wedge (A \neq B)$.

2. Operations on sets

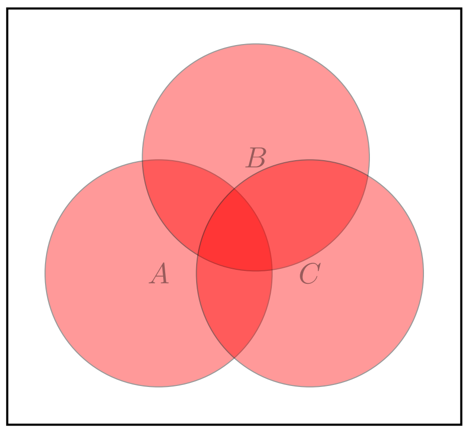
2.1. **Venn diagram.** The graphical representation of sets and elements is to think of a set A as an area and a member of it $x \in A$ as a point in that area:



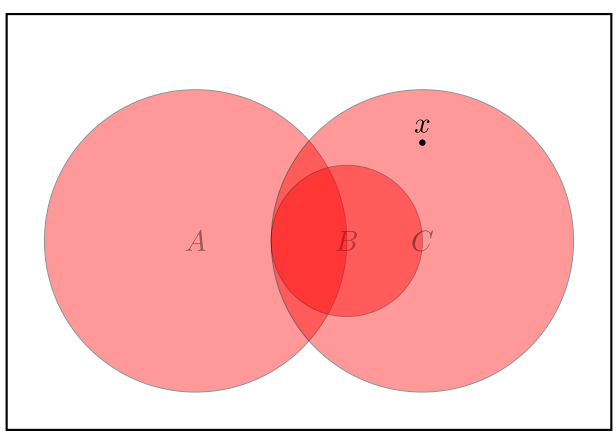
A Venn diagram of two or more sets, is graphial representation of general sets.



Three sets:



We can also add extra assumption to the diagram, for example if $B \subsetneq C$ we can express it as follows:

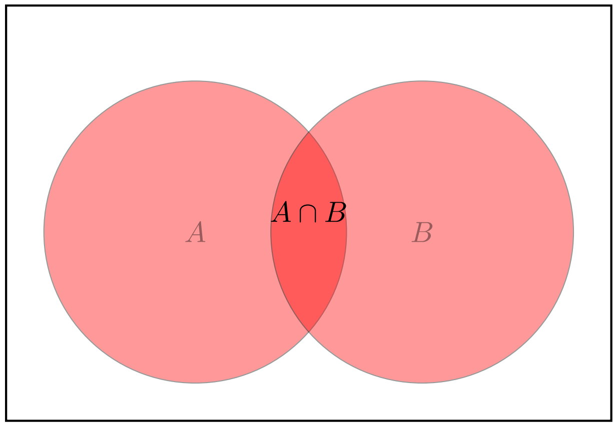


Note that x is a witness for a member of B which is not in C. A vigilant reader will notice that the picture is not fully accurate as we do not know if the witness x belongs to A.

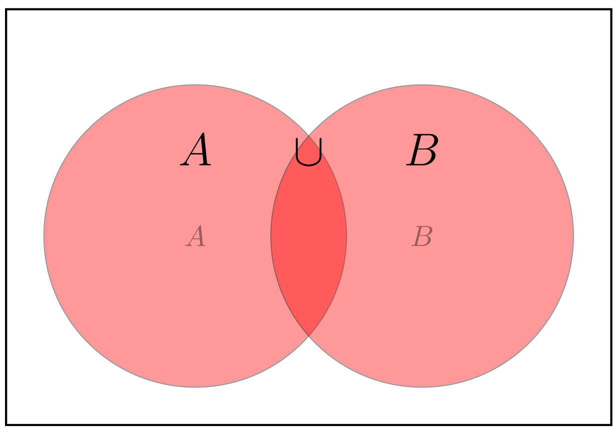
2.2. **Operation between sets.**

Definition 2.1. Let A, B be sets

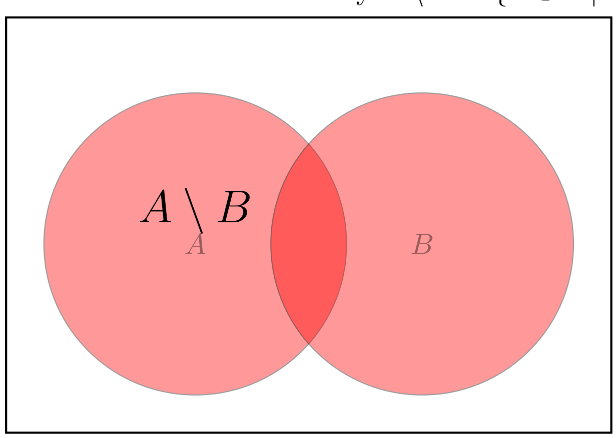
(1) The *intersection* of the sets is defined by $A \cap B = \{x \mid x \in A \wedge x \in B\}$.



(2) The *union* of the two sets is denoted by $A \cup B = \{x \mid x \in A \vee x \in B\}$

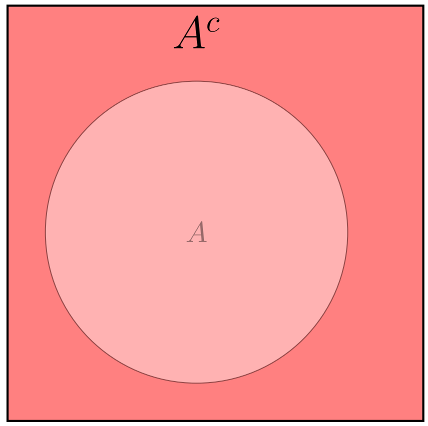


(3) The *difference* of the sets is defined by $A \setminus B = \{x \in A \mid x \notin B\}$

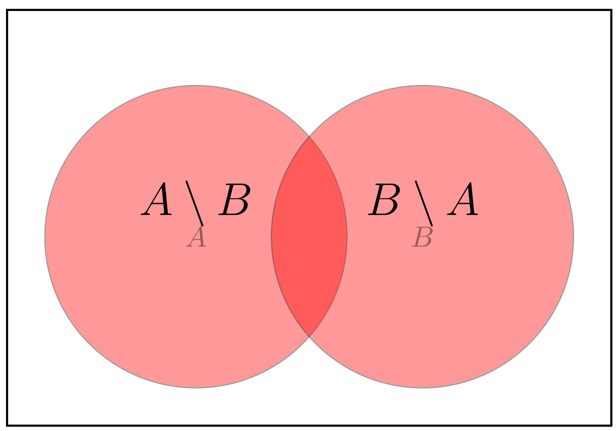


In the literature, difference of sets is sometimes denoted by $A - B$.

(4) The *complement* of A inside a supset U of A is denoted by $A^c = U \backslash A$. This is conceptually different from difference since we assume that U is some framework set and then A^c is an operation on a single set.



(5) The *symmetric difference* of the sets is denoted by $A \Delta B = (A \setminus B) \cup (B \setminus A)$



Example 2.2. (1) $\{1,2\} \cup \{2,3\} = \{1,2,3\}$, $\{1,4,5\} \cap \{2,4,4\} = \{4\}$, $[0,\infty) \cap (-\infty,1) = [0,1)$
(2) $\{1,2,6\} \setminus \{2,7,8\} = \{1,6\}$, $A \cap A = A \cup A = A$, the set of irrational numbers is the set $\mathbb{R} \setminus \mathbb{Q}$

2.3. **General union and the axiom of union.**

Axiom (Ax5. Union)**.** For any family of sets $\mathcal{F}$ there is a set U such that

$$\forall x, (\exists z. z \in \mathcal{F} \wedge x \in z) \to x \in Y.$$

The set Y only includes the union of the sets in $\mathcal{F}$ but with comprehension we my form the (unique) set:

Definition 2.3. For any set $\mathcal{F}$ we define

$$\bigcup \mathcal{F} := \{x \in Y \mid \exists z \in \mathcal{F}. x \in z\}$$

where Y is the set guaranteed from Ax5.

Problem 2. *Prove that the definition of $\bigcup\mathcal{F}$ does not depend on the choice of Y. Namely, if Y, Y' are two sets witnessing the union axiom for $\mathcal{F}$, then the resulting definition $\bigcup\mathcal{F}$ is the same.*

Example 2.4.
(1) $\bigcup\{A, B\} = A \cup B$.
(2) $\bigcup\Big\{\{0,1\},\{0,2\},\{0,3\}\Big\} = \{0,1,2,3\}$.
(3) $\bigcup\{[0,n) \mid n \in \mathbb{N}\} = [0,\infty)$.
(4) $\bigcup\{(n,n+1) \mid n \in \mathbb{Z}\} = \mathbb{R} \setminus \mathbb{Z}$.

Remark 2.5. In many situations (for example items (3),(4) above) the set $\mathcal{F}$ will be defined by replacement $\mathcal{F} := \{A(x) \mid x \in B\}$. This we write

$$\bigcup \mathcal{F} = \bigcup_{x\in B} A(x).$$

Exercise 2. *Compute* $\bigcup_{n\in\mathbb{N}_+}(\frac{1}{n}, n)$

Solution 4. *We claim that $\bigcup_{n\in\mathbb{N}_+}(\frac{1}{n}, n) = (0,\infty)$. We shall prove it by a double inclusion:*

- $\subseteq$*: Let $x \in \bigcup_{n\in\mathbb{N}_+}(\frac{1}{n}, n)$. By definition of union, there is $n \in \mathbb{N}_+$ such that $x \in (\frac{1}{n}, n)$. By definition of interval this means that $\frac{1}{n} < x < n$ and in particular $0 < x$. By definition of $(0,\infty)$ this means that $x \in (0,\infty)$.*
- $\supseteq$*: Let $x \in (0,\infty)$. To prove that $x \in \bigcup_{n\in\mathbb{N}_+}(\frac{1}{n}, n)$, we need to find some $n \in \mathbb{N}_+$ such that $\frac{1}{n} < x < n$. This means that $x < n$ and also, since $x > 0$, the inequality $\frac{1}{n} < x$ is equivalent to $\frac{1}{x} < n$. So n should be greater then both x and $\frac{1}{x}$. There exists such a natural number n (for example $n = \lceil \max\{x, \frac{1}{x}\}\rceil$).*

The following definition does not require the union axiom:

Definition 2.6. Let $\mathcal{F} \neq \emptyset$, define the intersection

$$\bigcap \mathcal{F} := \{x \mid \forall z \in \mathcal{F}. x \in z\}$$

Example 2.7.
(1) $\bigcap\Big\{\{1,2,3\},\{2,3,5\},\{1,2,7\}\Big\} = \{2\}$.
(2) $\bigcap\big\{(-\frac{1}{n}, \frac{1}{n}) \mid n \in \mathbb{N}_+\big\} = \{0\}$.
(3) $\bigcap_{n\in\mathbb{N}_+}(0, \frac{1}{n}) = \emptyset$.

Note that the intersection exists by comprehension since

$$\bigcap \mathcal{F} = \{x \in B \mid \forall z \in \mathcal{F}. x \in z\}$$

where B is any member of $\mathcal{F}$.

Proposition 2.8. *Sets operations identities:*

(1) Associativity:

   (a) $A \cap (B \cap C) = (A \cap B) \cap C$.
   (b) $A \cup (B \cup C) = (A \cup B) \cup C$.
   (c) $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.
(2) *Commutativity:*
   (a) $A \cap B = B \cap A$.
   (b) $A \cup B = B \cup A$.
   (c) $A\Delta B = B\Delta A$.
(3) *Distributivity:*
   (a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
   (b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(4) *Identities of difference and De-Morgan low's for sets:*
   (a) $A \setminus B = A \cap B^c$.
   (b) $(A \cup B)^c = A^c \cap B^c$.
   (c) $(A \cap B)^c = A^c \cup B^c$.
   (d) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$
   (e) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$.
(5) *Identities of the empty set:*
   (a) $A \cap \emptyset = \emptyset$.
   (b) $A \cup \emptyset = A$.
   (c) $A \setminus \emptyset = A$.
   (d) $\emptyset \setminus A = \emptyset$.
   (e) $A\Delta\emptyset = A$.
(6) *Identities of a set and itself:*
   (a) $A \cap A = A$.
   (b) $A \cup A = A$.
   (c) $A \setminus A = \emptyset$.
   (d) $A\Delta A = \emptyset$.

As examples, we will prove some of the items. We encourage the readers to write the proof for the other items.

Proof of 3.(b). We need to prove sets equality. We do so by proving a double inclusion.

$\underline{(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)}$: Let $x \in (A \cap B) \cup (A \cap C)$. By definition of $\cup$, we can split into cases:

(1) If $x \in A \cap B$, then by definition of $\cap$, $x \in A \wedge x \in B$. Hence $x \in B \cup C$ and $x \in A \cap (B \cup C)$.
(2) If $x \in A \cap C$ then $x \in A \wedge x \in C$. Once again, $x \in B \cup C$, thus $x \in A \cap (B \cup C)$.

In both cases we conclude that $x \in A \cap (B \cup C)$.

$\underline{(A \cap B) \cup (A \cap C) \supseteq A \cap (B \cup C)}$: Exercise.

□

Proof of 4.(e). Let us prove it using the other items.

$$A \setminus (B \cup C) \overset{4.(a)}{=} A \cap (B \cup C)^c \overset{4.(b)}{=} A \cap (B^c \cap C^c) \overset{6.(a)}{=} (A \cap A) \cap (B^c \cap C^c) =$$
$$\overset{2.(a)+1.(a)}{=} (A \cap B^c) \cap (A \cap C^c) \overset{4.(a)}{=} (A \setminus B) \cap (A \setminus C)$$

$\square$

Proof of 4.(b): We will prove $4.(b)$ in is generalized form, i.e.

$$(\bigcup \mathcal{F})^c = \bigcap \{B^c \mid B \in \mathcal{F}\}$$

$\subseteq$: Let $x \in (\bigcup \mathcal{F})^c$. Then $x \notin \bigcup \mathcal{F}$. By definition of union, it follows that there is no $B \in \mathbb{F}$ such that $x \in B$. In other words, for every $B \in \mathcal{F}$, $x \notin B$, or equivalently, $x \in B^c$. By definition of intersection, $x \in \bigcap \{B^c \mid B \in \mathcal{F}\}$.

$\supseteq$: similar to the first direction.

$\square$

Proposition 2.9. *The following are equivalent:*

(1) $A \subseteq B$
(2) $A \cap B = A$
(3) $A \setminus B = \emptyset$
(4) $A \cup B = B$

Proof. We shall prove: (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (1).[4]

(1) $\Rightarrow$ (2): Suppose $A \subseteq B$. We need to prove that $A \cap B = A$. We need to prove a double inclusion: Clearly, $A \cap B \subseteq A$. As for the other inclusion, let $x \in A$, since $A \subseteq B$ we conclude $x \in B$ and therefore $x \in A \cap B$ thus $A = A \cap B$.

(2) $\Rightarrow$ (3): Suppose that $A \cap B = A$ and suppose toward a contradiction that $AB \neq \emptyset$. By the definition of $\emptyset$, we conclude that there is $x \in A \setminus B$. By definition of sets difference, $x \in A \wedge x \notin B$. By definition of $\cap$, $x \notin A \cap B$. Thus $x \in A$ and $x \notin A \cap B$. By extensionality, $A \neq A \cap B$, contradicting the assumption.

(3) $\Rightarrow$ (4) and (4) $\Rightarrow$ (1) are left as exercises.

$\square$

2.4. **The power set.**

Axiom (Ax8 Power set)**.** For every set x there is a set y such that

$$\forall z, z \subseteq x \Rightarrow z \in y.$$

Definition 2.10. Let A be any set. Define the *power set* of A as the set of all possible subsets of A. We denote it by

$$P(A) := \{x \mid x \subseteq A\}$$

[4]This is a standard trick to prove equivalence between several statements. The order is not important as long as we close a circle of implications.

The definition above is justified by the power set axiom and comprehension (to establish existence) and extensionality (for uniqueness).

Example 2.11. (1) $P(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$
(2) $P(\{\{1\},2\}) = \{\emptyset, \{\{1\}\}.\{2\}, \{\{1\},2\}\}$
(3) $\emptyset, A \in P(A)$

Exercise 3. *Prove that $A \subseteq B$ If and only if $P(A) \subseteq P(B)$.*

Solution 5. $\underline{\Rightarrow}$: *Suppose that $A \subseteq B$. We want to prove that $P(A) \subseteq P(B)$. To prove the inclusion, let $X \in P(A)$, we want to prove that $X \in P(B)$. By definition of power set, $X \in P(A)$ implies that $X \subseteq A$. By the assumption $A \subseteq B$ and by a transitivity of inclusion we conclude that $X \subseteq B$. Again by definition of the power set, we have that $X \in P(B)$.*

$\underline{\Leftarrow}$: *Suppose that $P(A) \subseteq P(B)$. We want to prove that $A \subseteq B$. Usually, we would take an element $a \in A$ and try to prove that $a \in B$. However, there is a "trick" here which simplifies the proof. We have that $A \in P(A)$ and by the assumption, $P(A) \subseteq P(B)$, hence $A \in P(B)$. By definition of power set this means that $A \subseteq B$, as wanted.*

Definition 2.12. For a finite set A, we denote be $|A|$ the number of elements in the set A. For example $|\{1,2,3,18,-3\}| = 5$ and $|(-5,5) \cap \mathbb{Z}| = 9$.

Theorem 2.13. *Let A be a finite set then $|P(A)| = 2^{|A|}$.*

"Proof". Suppose that $A = \{a_1, ..., a_n\}$.

Every subset $X \subseteq A$, defines a sequence of n yes/no answers in the following way: for each $i = 1, ..., n$, we ask the question, is $a_i \in X$? For example suppose that:

- a_1 yes
- a_2 no
- a_3 no
- $a_4, ..., a_n$ yes

Then the sequence of answers would be

$$yes, no, no, yes, yes, yes, ..., yes$$

Note that from this sequence of answers we can reproduce the subset $X = \{a_1, a_4, ..., a_n\}$. This means that we are left to count the number of possible sequences of answers. Since typically there are n answers, with 2 possibilities for each answer we conclude that there are

$$\underbrace{2 \cdot 2....2}_{ntimes} = 2^n$$

many subsets of A. □

Problem 3. *What is the sequence of answers which corresponds to $\emptyset, A$?*

2.5. **The pairing axiom, Ordered pairs and Cartesian product.** Many mathematical objects involve order and repetitions. For example, the coordinates of a point in the plane is an object for which the order is important (since the point $P = (1,2)$ is not the same point as $Q = (2,1)$) and repetition is allowed (there is the point $(1,1)$). We shall aim to define objects which allow order and repetition. They will be denoted by $\langle x, y\rangle$ and the point is that we allow $x = y$ and $\langle x, y\rangle \neq \langle y, x\rangle$ in case $x \neq y$.

Definition 2.14. Let x, y be two objects, the *ordered pair* of x and y is defined by $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$.

but how do we justify this definition? we need to first be able to define sets using the list principle $\{a_1, ..., a_n\}$.

The following axiom ensures that some of the most basic concepts in set theory exist, and in particular, prove the existence of non-empty sets.

Axiom (Ax4. Pairing)**.** For every sets x, y there is a set w such that $x \in w \wedge y \in w$.

So using comprehension we can now prove the existence of the set $\{x, y\}$ and the set $\{x\}$ by applying paring to x, x.

We can now justify the definition of order pairs by applying pairing and comprehension to $\{x\}, \{x, y\}$.

Exercise 4. *Define (and prove the existence and uniqueness using pairing and other axioms) of the following objects:* $A \cup B$, $A \cap B$, $A \setminus B$, $A \Delta B$, $\{a_1, ..., a_n\}$.

Theorem 2.15 (Pairs equality)**.** *For every* a, b, c, d *we have*

$$\langle a, b\rangle = \langle c, d\rangle \leftrightarrow a = xc \wedge b = d$$

Proof. $\underline{\Rightarrow:}$ Suppose $a = c, b = d$ then

$$\langle a, b\rangle = \{\{a\}, \{a, b\}\} = \{\{c\}\{c, d\}\} = \langle c, d\rangle$$

$\underline{\rightarrow:}$ Suppose that $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$ then

I $\{a\}, \{a, b\} \in \{\{c\}, \{c, d\}\}$.

II $\{c\}, \{c, d\} \in \{\{a\}, \{a, b\}\}$.

Let us split into cases:

(a) If $a = b$ then
$$\{\{a\}, \{a, b\}\} = \{\{a\}, \{a, a\}\} = \{\{a\}, \{a\}\} = \{\{a\}\}$$
and therefore, since $\{c\}, \{c, d\}\{\{a\}\}$ we have that $\{c\} = \{a\} = \{c, d\}$. It follows that $a = b = c = d$, in which case we are done.

(b) The case $c = d$ is symmetric to the one above.

(c) Suppose that $a \neq b$ and $c \neq d$. Then $\{a\} = \{c\}$, since otherwise, by I, $\{a\} = \{c, d\}$ and therefore $a = c = d$, contradicting our assumption. Hence $a = c$. Also $\{a, b\} = \{c, d\}$ since otherwise, again by I, $\{a, b\} = \{c\}$ resulting in $a = c = b$, contradiction. This means that $b \in \{c, d\}$. Since $a = c$ and $b \neq a$ we conclude that $b = d$.

□

Definition 2.16. Let A, B be two sets. The *Cartesian product* of the sets (named after René Descartes) is defined by $A \times B = \{\langle a, b\rangle \mid a \in A, B \in B\}$

Also define the *square* of a set A is to be $A \times A$.

The existence of the cartesian product is justified by the powerset axiom, union, pairing, comprehension and extensionality:

Problem 4. *Prove that* $A \times B \subseteq P(P(A \cup B))$.

Remark 2.17. In practice the power set axiom can be replaced by the so-called replacement theorem which is usually assumed before the powerset axiom.

Example 2.18. (1) $\{1,2\} \times \{3,4\} = \{\langle 1,3\rangle, \langle 1,4\rangle, \langle 2,3\rangle, \langle 2,4\rangle\}$
(2) $\{2,3\}^2 = \{\langle 2,2\rangle, \langle 2,3\rangle, \langle 3,2\rangle, \langle 3,3\rangle\}$
(3) The *Real plane* is defined to be the set $\mathbb{R}^2$.

Definition 2.19. Let us define by recursion an n-tuple. A 1-tuple is defined by $\langle a\rangle = a$. Given we have defined an n-tuple, we define $n+1$-tuples using n-tuples and ordered pairs we have already defined.:

$$\langle a_1, ..., a_n, a_{n+1}\rangle = \langle\langle a_1, ...a_n\rangle, a_{n+1}\rangle$$

Example 2.20. (1) $\langle a_0\rangle = a_0$.
(2) Note that a 2-tuple is the same as an ordered pair. Indeed, let us denote momentarily the 2-tuple by $\langle a_0, a_1\rangle^*$, then we have

$$\langle a_0, a_1\rangle^* = \langle\langle a_0\rangle, a_1\rangle = \langle a_0, a_1\rangle$$

.

(3) $\langle a_0, a_1, a_2\rangle = \langle\langle a_0, a_1\rangle, a_2\rangle =$
$\{\{\langle a_0, a_1\rangle\}, \{\langle a_0, a_1\rangle, a_2\}\} = \{\{\{\{a_0\}, \{a_0, a_1\}\}\}, \{\{\{a_0\}, \{a_0, a_1\}\}, a_2\}\}$
(4) $\langle a_0, a_1, a_2, a_3\rangle = \langle\langle\langle a_0, a_1\rangle, a_2\rangle, a_3\rangle$

Theorem 2.21. *For all* $n \in \mathbb{N}_+$ *and* $a_1, .., a_n, b_1, .., b_n$,

$$\langle a_1, .., a_n\rangle = \langle b_1, ..., b_n\rangle \iff \forall 1 \leq i \leq n . a_i = b_i$$

Proof. We will use Theorem 2.15, that for every a_1, a_2, b_1, b_2

$$\langle a_1, a_2\rangle = \langle b_1, b_2\rangle \Leftrightarrow a_1 = b_1 \wedge a_2 = b_2$$

The induction is of the variable n, which is the length of the n-tuple.

The induction base: For $n = 1$, we need to prove that for every a_1, b_1

$$(\star) \quad \langle a_1\rangle = \langle b_1\rangle \iff a_1 = b_1$$

Recall that by definition of 1-tuple, $\langle a\rangle = a$, hence the equivalence $(\star)$ is clear.

The induction hypothesis: Suppose that for a general n, for every $a_1, .., a_n, b_1, ..., b_n$,

$$\langle a_1, .., a_n\rangle = \langle b_1, ..., b_n\rangle \iff \forall 1 \leq i \leq n . a_i = b_i$$

The induction step: We need to prove that for every $a_1, .., a_{n+1}, b_1, ..., b_{n+1}$,

$$\langle a_1, .., a_{n+1}\rangle = \langle b_1, ..., b_{n+1}\rangle \iff \forall 1 \leq i \leq n+1. a_i = b_i$$

Let $a_1, ..., a_{n+1}, b_1, ..., b_{n+1}$. We need t prove that

$$\langle a_1, .., a_{n+1}\rangle = \langle b_1, ..., b_{n+1}\rangle \iff \forall 1 \leq i \leq n+1. a_i = b_i$$

We will prove this equivalences with a chain of equivalences which we already know.

$$\langle a_1, .., a_{n+1}\rangle = \langle b_1, ..., b_{n+1}\rangle \underset{\text{Recursive definition of } n\text{-tuples}}{\iff} \langle\langle a_1, .., a_n\rangle, a_{n+1}\rangle = \langle\langle b_1, ..., b_n\rangle, b_{n+1}\rangle$$

$$\underset{\text{Pairs equality}}{\iff} \langle a_1, .., a_n\rangle = \langle b_1, ..., b_n\rangle \wedge a_{n+1} = b_{n+1} \underset{\text{I.H.}}{\iff}$$

$$\forall 1 \leq i \leq n. a_i = b_i \wedge a_{n+1} = b_{n+1} \iff \forall 1 \leq i \leq n+1. a_i = b_i$$

$\square$

Definition 2.22. $A_1 \times ... \times A_n = \{\langle a_1, ..., a_n\rangle \mid a_1 \in A_1, ..., a_n \in A_n\}$

Definition 2.23. For $n \geq 1$, $A^n = \underbrace{A \times ... \times A}_{n \text{ times}}$.

3. Relations

Relations are a wide class of important mathematical objects such as functions, orders and equivalence relation.

3.1. Non-formal functions.

Functions are among the most common mathematical objects and appear in almost every mathematical theory. Intuitively speaking, a function is just a machine which assigns to **every** element a (the input) in a given set A (the domain of the function) a **unique** element $f(a)$ (the output/ the image of a) in a set B (the range of the function). To illustrate these ideas, here are some day-to-day examples:

(1) The function which attaches to every person its height. The domain of the function is the set of humans and the range of the functions is the set of real numbers (theoretically, a person can be 5 feet and $\sqrt{2}$ inches tall).
(2) If we attach to every person, its siblings, the result is not a function and there are two reasons for that. The first is that there are people with no siblings (and therefore the function is not defined for **every** person), also there are people with more than one sibling and for those people, we do not attach a **unique** person).

We will formally define function only later and steak with a non-formal definition for now. We will later have to justify this non-formal definition.

Definition 3.1 (Non formal)**.** Let A, B be any sets. A function from A to B is an object f, such that:

(1) f is *total* on A: for every $a \in A$, $f(a)$ is defined.
(2) f is univalent: for every $a \in A$, $f(a)$ is a unique element of B.

We denote it we $f : A \to B$. The set A is the *domain f the function* f which is denoted by $\mathrm{dom}(f)$ and B is the *range of the function* f which we denote by $\mathrm{rng}(f)$.

3.1.1. *How to define functions?* Usually, we declare what A and B are in advance by saying we are about to define a function $f : A \to B$. Then we provide some formula with a free variable a which we think of as a general element in the set A. This formula prescribes what element $f(a) \in B$ is assigned to a.

Example 3.2. (1) Define $f : \{1, 2, 3\} \to \{1, 2, 3, 4\}$ by $f(x) = x + 1$. Then $f(1) = 1 + 1 = 2$, $f(3) = 3 + 1 = 4$.
(2) Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$.
(3) define $f : \mathbb{R} \to \mathbb{R}$ by $f(r) = 2$, this is the constant function which for every real r returns the value 2.
(4) $f : \big\{\{1, 2, 3\}, \{1, 3, 5\}\big\} \to \mathbb{N}$, $f(X) = \max(X)$.
(5) $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(\langle x, y\rangle) = \langle x^2 + y^2, x - y + 1\rangle$.
(6) $f : \mathbb{N} \to P(\mathbb{Z})$ $f(n) = \{n\} \cup \{1, -1\}$.
(7) Here are some non-examples:
    (a) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{x}$.
    (b) $f : P(\mathbb{N}) \to \mathbb{N}$, $f(X) = \min(X)$.
    (c) $f : [0, \infty) \to [0, \infty)$, $f(x) = x - 1$.
    (d) $f : [0, \infty) \to \mathbb{R}$ $f(x) = y$ for y such that $y^2 = x$.
(8) Definition of a function by cases: Suppose we which to define a function on a set A, and for some of the elements of A we want one formula and for the another part of A we want to use a different formula. We can do that the following way: "Define $f : A \to B$ by

$$f(a) = \begin{cases} \text{(first formula)} & \text{(first condition on } a\text{)} \\ \text{(second formula)} & \text{(second condition on a)} \\ \ldots \end{cases}$$

where the conditions on a describe the element for which you would like to use the formula. When we check that a function defined by cases is well defined, we also have to check the condition on a covers all possible a and that they are "disjoint" in the sense that no a satisfy two of the condition.
    (a) Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(a) = \begin{cases} \sqrt{a} & a > 0 \\ a + 1 & -1 < a \leq 0 \\ |a|^3 - 15 & a \leq -1 \end{cases}$$

    We can also use "otherwise" if we would like to take care of the remaining cases.
    (b) If we have a "small" number of elements in the domain we can use the definition by cases above to explicitly assign to every

element a value, without worrying about a formula which describes that assignment. For example $f : \{1,2,3\} \to \{a,b,c,d\}$

$$f(x) = \begin{cases} b & x = 3 \\ a & x = 2 \\ c & x = 1 \end{cases}$$

Important: If we define $f : A \to B$ by a formula $f(a) =$(some formula) we **must** always make sure that the functions we define are well defined in the sense that:

(1) The function is total. Practically, this means that we should make sure that the formula for $f(a)$ is defined for every $a \in A$.
(2) The function is univalent. This means that for every $a \in A$, the formula for $f(a)$ points to a single element. (This is trivial in most cases)
(3) for every $a \in A$ the formula for $f(a)$ describes an element of B. So the range we declared when we wrote $f : A \to B$ is indeed correct.

Here are further examples:

(1) $f : \mathbb{N} \to \mathbb{N}$ defined by $f(x) = x^2$ satisfies $f(4) = 16$.
(2) $g : \mathbb{N} \to P(\mathbb{N})$ defined by $g(x) = \{x, x+1\}$ satisfies $g(5) = \{5,6\}$.
(3) $t : \mathbb{N} \to \mathbb{N}$ defined by $t(n) = \begin{cases} 0 & n \in \mathbb{N}_{even} \\ 1 & n \in \mathbb{N}_{odd} \end{cases}$.
satisfies that $t(1) = 1, t(14) = 0$. $s(f)(3) = \{-2\}$.
(4) $F : P(\mathbb{N})^2 \to \mathbb{N}$ defined by $F(\langle A, B\rangle) = \begin{cases} 0 & A \cap B = \emptyset \\ \min(A \cap B) & else \end{cases}$
satisfies that $F(\langle \{1,2,3,4\}, \mathbb{N}_{even}\rangle) = 2$.
(5) $f : \mathbb{N}^2 \to P(\mathbb{N})$ defined by $f(\langle x, y\rangle) = \{n \in \mathbb{N} \mid x < n < y\}$ satisfies $f(\langle 1,4\rangle) = \{2,3\}$ and $f(\langle 4,1\rangle) = \emptyset$.

When formally working with functions we will only need the following criterion for equality of functions. This is exactly what we will have to justify once we will give the formal definition of a function:

Theorem 3.3. *Let $f, g : A \to B$ be two function. Then the following are equivalent:*

(1) $\forall x \in A. f(x) = g(x)$.
(2) $f = g$.

The theorem says that two functions with the same domain and range are equal if and only if for every x in this domain, the functions assign the same value to x. From this point, our proofs will be completely formal relaying in this theorem.

Remark 3.4. The function equality theorem indicated that a function is **not** the same as a formula defining it.

For example the functions: $f_1, f_2 : \{-1, 0, 1\} \to \mathbb{R}$ defined by $f_1(x) = |x|$ and $f_2(x) = x^2$ have different formulas but they define the same function since $f_1(-1) = f_2(-1),\ f_1(0) = f_2(0),\ f_1(1) = f_2(1)$.

Remember! Different formulas can define the same function.

3.1.2. *Operations on functions.*

Definition 3.5. Let $f : A \to B$ be a function and $X \subseteq A$. We define the *restriction of* f *to* X, denoted by $f \restriction X : X \to B$, to be the function with domain $dom(f \restriction X) = X$ and for every $x \in X$, $(f \restriction X)(x) = f(x)$.

Intuitively, the restriction of a function acts the same way that the original function did, the only difference is that the domain restricts to the new set X.

Definition 3.6. Let A be any set. We define the *Identity function* on A as the function $Id_A : A \to A$ defined by $Id_A(a) = a$.

Example 3.7. Let $f : \mathbb{Z} \to \mathbb{Z}$ be defined as $f(z) = |z|$. Prove that $f \restriction \mathbb{N} = Id_{\mathbb{N}}$

Proof. We want to prove equality of functions. First we want to prove that $\text{dom}(f \restriction \mathbb{N}) = \text{dom}(Id_{\mathbb{N}})$. Indeed by definition of restriction and the identity function, both of the functions have domain $\mathbb{N}$. Next we want to prove that $\forall x \in \mathbb{N}.(f \restriction \mathbb{N})(x) = Id_{\mathbb{N}}(x)$. Let $x \in \mathbb{N}$, then by definition of restriction and since $n \geq 0$ we have

$$(f \restriction \mathbb{N})(x) = f(x) = |x| = x$$

and by definition of the identity function we have

$$Id_{\mathbb{N}}(x) = x$$

Hence

$$(f \restriction \mathbb{N})(x) = x = Id_{\mathbb{N}}(x)$$

as wanted □

Definition 3.8. Let $f : A \to B$ and $g : B \to C$ be two functions. We define the composition of g in f as $g \circ f : A \to C$, to be the function with domain f and range C such that for each $a \in A$, $(g \circ f)(a) = g(f(a))$.

Example 3.9.
(1) $f(x) = x^2$ and $g(x) = x + 1$, then $g \circ f(x) = x^2 + 1$ and $f \circ g(x) = (x + 1)^2$.
(2) $f : P(\mathbb{N}) \setminus \{\emptyset\} \to \mathbb{N} \times \mathbb{N}$, $f(X) = \langle \min(X), \min(X) + 1\rangle$ and $g : P(\mathbb{N}) \to P(\mathbb{N}) \setminus \{\emptyset\}$, $g(X) = X \cup \{0\}$. Then $f \circ g(X) = f(X \cup \{0\}) = \langle \min(X \cup \{0\}), \min(X \cup \{0\}) + 1\rangle = \langle 0, 1\rangle$.

Proposition 3.10. *Suppose that* $f : A \to B$, $g : B \to C$ *and* $h : C \to D$. *Then:*

(1) $f \circ Id_A = f$, $Id_B \circ f = f$.
(2) $h \circ (g \circ f) = (h \circ g) \circ f$.

Proof. Let us prove for example that $f \circ Id_A = f$. We need to prove function equality, the domain of both functions is A. Let $a \in A$, then $(f \circ Id_A)(a) = f(Id_A(a)) = f(a)$ hence $f \circ Id_A = f$. □

3.1.3. *Properties of functions.*

Definition 3.11. Let $f : A \to B$ be a function we sat that f is:

(1) One to one/ injective: if for every $a_1, a_2 \in A$, if $f(a_1) = f(a_2)$ then $a_1 = a_2$.
(2) Onto/ surjective: if for every $b \in B$ there is $a \in A$ such that $f(a) = b$.

Example 3.12. (1) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not injective as $1 \neq -1$ and $f(-1) = (-1)^2 = 1 = 1^2 = f(1)$.

(2) $f : \mathbb{N} \to \mathbb{Z}$ defined by $f(n) = n - 1$ is injective.

Proof. Let $n_1, n_2 \in \mathbb{N}$. Suppose that $f(n_1) = f(n_2)$, we want to prove that $n_1 = n_2$. By definition of f, $n_1 - 1 = n_2 - 1$, adding 1 to both sides of the equation we conclude that $n_1 = n_2$. □

(3) $g : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ defined by $g(\langle n, m\rangle) = \langle 2n+m, n+m\rangle$ is injective.

Proof. Let $\langle n_1, m_1\rangle, \langle n_2, m_2\rangle \in \mathbb{N} \times \mathbb{N}$ and assume that $g(\langle n_1, m_1\rangle) = g(\langle n_2, m_2\rangle)$ we want to prove that $\langle n_1, m_1\rangle = \langle n_2, m_2\rangle$. By the assumption we know that $\langle 2n_1 + m_1, n_1 + m_1\rangle = \langle 2n_2 + m_2, n_2 + m_2\rangle$ and by equality of pair we get that

$$2n_1 + m_1 = 2n_2 + m_2 \text{ and } n_1 + m_1 = n_2 + m_2$$

Subtracting the second equation from the first we get:

$$2n_1 + m_1 - (n_1 + m_1) = 2n_2 + m_2 - (n_2 - m_2)$$

$$n_1 = n_2$$

Hence by the equality $n_1 + m_1 = n_2 + m_2$, we have that $n_1 = n_2$ cancels so $m_1 = m_2$. By equality of pairs we conclude that $\langle n_1, m_1\rangle = \langle n_2, m_2\rangle$. □

(4) $F : P(\mathbb{N}) \to P(\mathbb{N})$ defined by $F(X) = \{x + 1 \mid x \in X\}$ is injective.

Proof. Let $X_1, X_2 \in P(\mathbb{N})$, suppose that $F(X_1) = F(X_2)$ we want to prove that $X_1 = X_2$. By definition of F,

$$)*) \quad \{x + 1 \mid x \in X_1\} = \{x + 1 \mid x \in X_2\}$$

Let us prove $X_1 = X_2$ by a double inclusion:

(a) $\underline{X_1 \subseteq X_2}$: Let $x_0 \in X_1$ we want to prove that $x_0 \in X_2$. By definition $x_0 + 1 \in \{x + 1 \mid x \in X_1\}$ and by $(*)$, $x_0 + 1 \in \{x + 1 \mid x \in X_2\}$. By the replacement principle, there exists $y \in X_2$ such that $x_0 + 1 = y + 1$, hence $x_0 = y \in X_2$, which implies that $x_0 \in X_2$ as wanted.
(b) $\underline{X_2 \subseteq X_1}$: Symmetric to the first inclusion.

□

(5) $F_1 : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by $F(\langle n, m\rangle) = 2^n \cdot 3^m$ is injective.

Proof. Let $\langle n_1, m_1\rangle, \langle n_2, m_2\rangle \in \mathbb{N}\times\mathbb{N}$. Suppose that $F_1(n_1,m_1) = F_1(n_2,m_2)$ we want to prove that $\langle n_1, m_1\rangle = \langle n_2, m_2\rangle$. By definition of F_2 we have that $(*)\quad 2^{n_1}3^{m_1} = 2^{n_2}3^{m_2}$. By the fundamental theorem of arithmentics, each positive natural number has a unique factorization into primes. The equality $(*)$ provides two factorization into primes of the same numbers, hence it must be the same, namely $n_1 = n_2$ and $m_1 = m_2$. By the basic property of pairs, $\langle n_1, m_1\rangle = \langle n_2, m_2\rangle$. □

Definition 3.13. Let $f : A \to B$ be a function. The image of f, denoted by $Im(f) = \{f(x) \mid x \in A\}$.

Exercise 5. *For the function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$ Prove that* $\operatorname{dom}(f) = Rng(f) = \mathbb{R}$ *while* $\operatorname{Im}(f) = [0,\infty)$.

Solution 6. *Since the last equality if a set equality, we should prove it by a double implication:*

(1) $\subseteq$*: Let $r \in \operatorname{Im}(f)$, we need to prove that $r \in [0,\infty)$. By definition of* $\operatorname{Im}(f)$*, there is $x \in \mathbb{R}$ such that $f(x) = r$. Those $r = x^2 \geq 0$ and by definition of* $[0,\infty)$*,* $r \in [0,\infty)$.

(2) $\supseteq$*: Let $r \in [0,\infty)$. we need to prove that $r \in \operatorname{Im}(f)$. By definition, $r \geq 0$ and therefore we have $\sqrt{r}$ defined. Define (This is an existential proof) $x = \sqrt{r}$, then $f(x) = x^2 = r$.*

Remark 3.14. f is surjective if and only if $Im(f) = Range(f)$.

Example 3.15. (1) The function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) = 2n$ is not surjective.

Proof. For example $1 \in \mathbb{N}$ and for every $n \in \mathbb{N}$, $f(n) \neq 1$. Otherwise, there exists $n \in \mathbb{N}$ such that $f(n) = 1$ then by definition of f, $2n = 1$ which implies that 1 is even, contradiction. □

Note also that $Im(f) = \mathbb{N}_{even}$ and that f is injective.

(2) The function $g : P(\mathbb{Z}) \to P(\mathbb{N})$ defined by $g(X) = X \cap \mathbb{N}$ is surjective.

Proof. Let $Y \in P(\mathbb{N})$ we want to prove that there is $X \in P(\mathbb{Z})$ such that $f(X) = Y$. Define $X = Y$, then since $Y \in P(\mathbb{N})$, $Y \in P(\mathbb{Z})$. Also, to see that $g(Y) = Y$, we need to prove that $Y \cap \mathbb{N} = Y$. This is equivalent (by a proposition we have seen previously) to the fact that $Y \subseteq \mathbb{N}$. This follows since $Y \subseteq \mathbb{N}$. □

Also note that $Im(g) = P(\mathbb{N})$, (since we just proved that g is surjective) and it is not injective since for example $g(\{-1,1\}) = \{1\} = g(\{1\})$.

(3) The function $h : (0,\infty) \to (0,\infty)$ defined by $h(x) = \frac{1}{x}$ is surjective.

Proof. Let $y \in (0,\infty)$, we want to prove that there is $x \in (0,\infty)$ such that $h(x) = y$. Namely, we want that $\frac{1}{x} = y$. Then define $x = \frac{1}{y}$. Since $0 < y$, also $0 < x$ and therefore $x \in (0,\infty)$ and we have that $h(x) = \frac{1}{\frac{1}{y}} = y$ as wanted. □

(4) $G : P(\mathbb{N}) \times P(\mathbb{N}) \to P(\mathbb{N} \times \mathbb{N})$ defined by $G(\langle X, Y\rangle) = X \times Y$ is not onto.

Proof. For example $\{\langle 1,1\rangle, \langle 2,2\rangle\} \in Range(G) \setminus Im(G)$. Suppose toward a contradiction that $G(\langle X, Y\rangle) = \{\langle 1,1\rangle, \langle 2,2\rangle\}$. Then by definition of G, $X \times Y = \{\langle 1,1\rangle, \langle 2,2\rangle\}$. By set equality, this means that $\langle 1,1\rangle$, $\langle 2,2\rangle \in X \times Y$. which by the definition of Cartesian product implies that $1, 2 \in X$ and $1, 2 \in Y$. But then $\langle 1,2\rangle \in X \times Y$ but $\langle 1,2\rangle \notin \{\langle 1,1\rangle, \langle 2,2\rangle\}$, contradiction. □

Proposition 3.16. *Let $f : A \to B$ and $g : B \to C$ be any functions.*

(1) If f, g are injective then so is $g \circ f$.
(2) If f, g are surjective then so is $g \circ f$

Definition 3.17. A function $f : A \to B$ is invertible if there is a function $g : B \to A$ such that:

$$g \circ f = id_A \quad \text{and } f \circ g = id_B$$

Example 3.18. (1) $f : \{a, b, c\} \to \{1, 2, 3\}$ defined by

$$f(x) = \begin{cases} 1 & x = a \\ 2 & x = b \\ 3 & x = c \end{cases}$$

is invertible as witnessed by the function $g : \{1, 2, 3\} \to \{a, b, c\}$,

$$g(x) = \begin{cases} a & x = 1 \\ b & x = 2 \\ c & x = 3 \end{cases}$$

(2) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x + 1$ is invertible since the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x - 1$ satisfy that $g \circ f = f \circ g = Id_{\mathbb{R}}$.
(3) The function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) = n + 1$ is **not** invertible. The function $g(n) = n-1$ is not a function from $\mathbb{N}$ to $\mathbb{N}$ as $g(0) = -1$. To formal way to prove it is to use the next theorem (and the fact the g is not onto). If we restrict the range of f to $\mathbb{N}_+$ then g above from $\mathbb{N}_+$ to $\mathbb{N}$ witnesses that f is invertible.
(4) There is no $f : \{a, b, c\} \to \{1, 2, 3, 4\}$ which is invertible.
(5) $f : P(\mathbb{N}) \to P(\mathbb{N})$ defined by $f(X) = \mathbb{N} \setminus X$ is invertible as $f \circ f = Id_{P(\mathbb{N})}$.

Theorem 3.19. *If g_1, g_2 are two inverse functions of f then $g_1 = g_2$. We denote the inverse function of f by f^{-1}.*

Proof. Suppose the g_1, g_2 are two inverse function of f, then

$$g_1 \circ f = id_A \quad \text{and } f \circ g_1 = id_B$$

$$g_2 \circ f = id_A \quad \text{and } f \circ g_2 = id_B$$

It follows that

$$g_1 = g_1 \circ Id_B = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = Id_A \circ g_2 = g_2$$

□

Theorem 3.20. *A function $f : A \to B$ is invertible if and only if it is one to one and onto.*

Proof. Suppose that f is invertible and let $f^{-1} : B \to A$ be the inverse function. Let us prove that f is one to one and onto:

- one to one: Let $a_1, a_2 \in A$, suppose that $f(a_1) = f(a_2)$, we want to prove that $a_1 = a_2$. Then $f^{-1}(f(a_1)) = f^{-1}(f(a_2))$ and since $f^{-1} \circ f = Id_A$ we get that
$$a_1 = f^{-1}(f(a_1)) = f^{-1}(f(a_2)) = a_2$$
- onto: Let $b \in B$, we want to prove that there is $a \in A$ such that $f(a) = b$. Let $a = f^{-1}(b) \in A$. Then $f(a) = f(f^{-1}(b))$ and since $f \circ f^{-1} = Id_B$, we have that $f(a) = f(f^{-1}(b)) = b$ as wanted.

For the other direction, suppose that f is one to one and onto B. We want to prove that f is invertible, namely that there is a function $g : B \to A$ such that $f \circ g = Id_B$ and $g \circ f = Id_A$. Here is the definition of g: For any element of b, there is (since f is onto B) a unique (since f is one to one) element $a_b \in A$ such that $f(a_b) = b$. Define $g(b) = a_b$. Let us prove that g is inverse to f:

- $\underline{g \circ f = Id_A}$: Let $a \in A$, then denote $f(a) = b \in B$. By definition $g(b) = a_b$ is the unique element in A such that $f(a_b) = b$ and since $f(a) = b$ it follows that $a = a_b$. Hence $g(f(a)) = g(b) = a_b = a$. It follows that $g \circ f = Id_A$.
- $\underline{f \circ g = Id_B}$: Let $b \in B$, by definition, $g(b) = a_b$ and a_b has the property that it is (the unique which is) mapped to b, namely $f(a_b) = b$. Hence $f(g(b)) = f(a_b) = b$. Again it follows that $f \circ g = Id_B$.

□

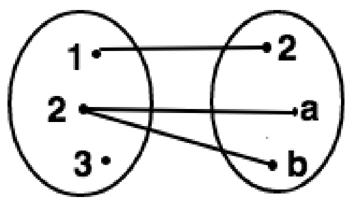
3.2. **General relations.** Toward a formal definition of a function, we would like to describe that certain objects relate to other objects. To turn relations into a formal mathematical object, we need to define them as sets. First, how would we code that an object a relates to an object b? we can use the ordered pair $\langle a, b\rangle$. A single relation describes many such connections, hence it is a set of ordered pairs:

Definition 3.21. A relation from the set A to the set B is set $R \subseteq A \times B$.

Example 3.22. (1) $R = \{\langle 1, 2\rangle, \langle 1, 3\rangle\}$ is a relation from $\{1, 2\}$ tp $\{1, 2, 3\}$ since
$$R \subseteq \{1, 2\} \times \{1, 2, 3\}$$
. R is also a relation from $\mathbb{R}$ to $\mathbb{N}$.

(2) $\{\langle 1, \sqrt{2}\rangle, \langle 2, 4\rangle\}$ is not a relation from $\mathbb{N}$ to $\mathbb{N}$.

(3)
$$id_{\mathbb{N}} = \{\langle n,n\rangle \mid n \in \mathbb{N}\}$$
$$\leq_{\mathbb{N}} = \{\langle n,m\rangle \in \mathbb{N}^2 \mid \exists k \in \mathbb{N}.n+k = m\}, \ \ <_{\mathbb{N}} = \{\langle n,m\rangle \in \mathbb{N}^2 \mid \exists k \in \mathbb{N}_+.n+k = m\}$$
are three relations from $\mathbb{N}$ to $\mathbb{N}$. Note that
$$\leq = < \cup id_{\mathbb{N}}$$
(4) $A = \{\langle x,y\rangle \in \mathbb{R}^2 \mid x - y \in \mathbb{Q}\}$ for example $\langle 3+\sqrt{2}, \sqrt{2}\rangle \in A$, $\langle 1, \pi\rangle \notin A$.
(5) $R = \{\langle X,Y\rangle \in P(\mathbb{N}) \times P(\mathbb{Z}) \mid X \subseteq Y\}$. R is a relation from $P(\mathbb{N})$ to $P(\mathbb{Z})$.
(6) It is sometimes convinient to imagine a relation as two potato's representing the sets A and B, and then and arrows from A to B. For example, if $R = \{\langle 1,2\rangle, \langle 2,a\rangle, \langle 2,b\rangle\}$ From $\{1,2,3\}$, to $\{2,a,b\}$:



(7) $S = \{\langle x,y\rangle \in \mathbb{Z}^2 \mid x \text{ divides } y\}$, Then S is a relation from $\mathbb{Z}$ to $\mathbb{Z}$.
(8) In general, for every set A we denote the identity relation on the set A by $id_A = \{\langle a,a\rangle \mid a \in A\}$.
(9) A funciton is also a relation. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$. This function establishes connections between the real number x and the real number x^2, So the formal definition of the function as a set is $f = \{\langle x, x^2\rangle \mid x \in \mathbb{R}\}$.

Remark 3.23. In most cases a relation (i.e. a set of pairs) has a "meaning", which is some notion we already familiar with, just not in terms of sets of pairs. In the previous examples, $\leq_{\mathbb{N}}$ is just a formal representation for the usual $\leq$ where we only consider natural numbers. The relation D is just the divisibility relation on between integers, and id_A is just the equality relation where we only consider elements of the set A. **However**, a general relation R, is just an abstract object. It does not necessarily have a meaning as in the previous examples. Examples $(1), (2), (6)$ do not arise from a natural notion. We can always artificially force a meaning to it, but this would be of no use.

Definition 3.24. Let R be a relation from A to B. Define:

Definition 3.25. (1) $\mathrm{dom}(R) = \{a \in A \mid \exists b \in B, \ \langle a,b\rangle \in R\}$.

(2) $\mathrm{im}(R) = \{b \in B \mid \exists a \in A,\ \langle a, b\rangle \in R\}$.
(3) $R^{-1} = \{\langle b, a\rangle \mid \langle a, b\rangle \in R\}$.
(4) $Id_A = \{\langle a, a\rangle \mid a \in A\}$.

Important: When handling general relations, do not try to find a "meaning" for it. Instead, you should simply think of a set of pairs. When handling a specific relation, it is important to understand the idea behind it (by finding examples pairs of elements which belongs to the relation).

Problem 5. *Let R be a relation from A to B, S be a relation from B to C. Define*

$$S \circ R = \{\langle a, c\rangle \in A \times C \mid \exists b \in B,\ \langle a, b\rangle \in R \wedge \langle b, c\rangle \in S\}$$

Prove that:

(1) $R \circ Id_A = R$.
(2) $Id_B \circ R = R$.
(3) $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$.
(4) If T is a relation from C to D them $(T \circ S) \circ R = T \circ (S \circ R)$.

3.3. **abstract functions.** The formal way to define a function is as relations:

Definition 3.26. Let A, B be two sets. A *function* from A to B is a relation from A to B such that:

(1) f is total *Total* on A: $\forall a \in A.\exists b \in B.\langle a, b\rangle \in f$. (i.e. $\mathrm{dom}(f) = A$)
(2) f is *univalent* or a *partial function on* A: $\forall a \in A.\forall b_1, b_2 \in B.\langle a, b_1\rangle \in f \wedge \langle a, b_2\rangle \in f \Rightarrow b_1 = b_2$.

Notation 3.27. If f is a function from A to B we denote it by $f : A \to B$. Also if $f : A \to B$ is a function, we denote $f(a) = b$ if and only if $\langle a, b\rangle \in f$. So $f(a)$ is the unique object in the set B that the function f attaches to the element a.

Example 3.28. (1) Let $f = \{\langle 1, a\rangle, \langle 3, b\rangle, \langle 2, a\rangle\}$. To see that f is a function from $\{1, 2, 3\}$ to $\{a, b, c\}$, we need to prove that for every $x \in \{1, 2, 3\}$ the is a unique $y \in \{a, b, c\}$ such that $\langle x, y\rangle \in f$ (and then we can denote $f(x) = y$). Since there are only 3 elements in f we can go one-by-one over the elements of f and check that this is indeed the case manually. Now that we are sure that f is a function, we can write $f : \{1, 2, 3\} \to \{a, b\}$ and

$$f(1) = a,\ f(2) = a,\ f(3) = b.$$

(2) The identity relation on a set A, is a function $id_A : A \to A$ satisfying $id_A(a) = a$ for every $a \in A$.
(3) Consider $S = \{\langle X, x\rangle \in P(\mathbb{N}) \times \mathbb{N} \mid x \in X\}$. This is **not** a function from $P(\mathbb{N})$ to $\mathbb{N}$ since it is not total. For example[5], $\emptyset \in P(\mathbb{N})$, and there is no x such that $\langle \emptyset, x\rangle \in S$, otherwise we would have $x \in \emptyset$.

[5]To prove that e function is not total/univalent, we should provide a counter example.

Let us try and remove $\emptyset$ to see if we get a function. Is S a function from $P(\mathbb{N}) \setminus \{\emptyset\}$ to $\mathbb{N}$? This is still not a function since it is not univalent. For example, $\langle\{1,2,3\},1\rangle, \langle\{1,2,3\},2\rangle \in S$. Also it is not $Total$

(4) Let A, B be any sets. For every $b \in B$ the *constant function with value* b is the the relation f_b from A to B

$$f_b = \{\langle x, b\rangle \mid x \in A\} = A \times \{b\}.$$

Claim: f_b is a function from A to B.

Proof. We need to prove that f_b is total on A and univalent.
Total: We need to prove that for every $x \in A$ there is $y \in B$ such that $\langle x, y\rangle \in f_b$. Let $x \in A$. Define $y = b$, then by the definition of f_b, $\langle x, b\rangle \in f_b$.
Univalent: We need to prove that for every $a \in A$ and for every $b_1, b_2 \in B$, if $\langle a, b_1\rangle, \langle a, b_2\rangle \in f_b$ then $b_1 = b_2$. Let $a \in A$, $b_1, b_2 \in B$ and suppose that $\langle a, b_1\rangle, \langle a, b_2\rangle \in f_b$. We want to prove that $b_1 = b_2$. By the definition of f_b, since we have that $b_1 = b = b_2$. $\square$

Hence $f_b : A \to B$ is a function satisfying $\forall a \in A. f_b(a) = b$.

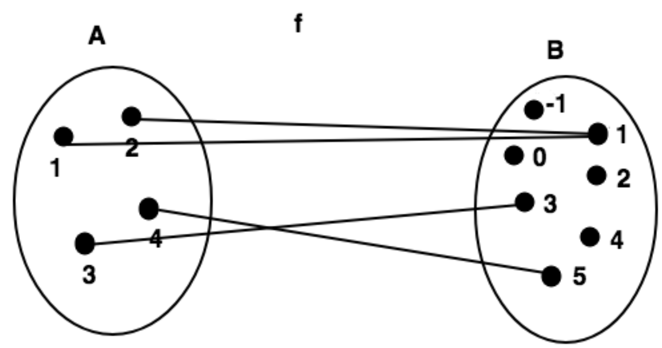
(5) $\pi_1 : A \times B \to A$ $\pi_1 = \{\langle\langle a, b\rangle, c\rangle \in (A \times B) \times A \mid a = c\}$ Is called the *projection to the left coordinate*, it satisfies that $\pi(\langle a, b\rangle) = a$. Similarly, the *projection to the right coordinate* is denoted $\pi_2 : (A \times B) \to B$ and it satisfies $\pi_2(\langle a, b\rangle) = b$.

(6) To summation operation on the rational number (or on the natural numbers/integers/reals) is a function $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. We are used to write $3 + 5 = 8$ instead of $+(\langle 3, 5\rangle) = 8$.

(7) Let $g : P(A) \times P(B) \to P(A)$ defined by $g = \{\langle\langle X, Y\rangle, Z\rangle \in (P(A) \times P(B)) \times P(A) \mid Z = X \cap Y\}$ we have that $g(X, Y) = X \cap Y$

(8) Given a set of pairs R in $A \times B$ we can represent R as a collection of arrows from he set A to the set B. This is very convenient when considering functions. For example, to verify the R is a function from A to B we should simply verify(not prove!) that there is **exactly one** arrow attached to **every** element of A. For example, consider

$$f : \{1,2,3,4\} \to \{-1,0,1,2,3,4,5\} \ f = \{\langle 1,1\rangle, \langle 2,1\rangle, \langle 3,3\rangle\langle 4,5\rangle\}$$



Definition 3.29. A sequence of elements in the set A is a function $f : \mathbb{N} \to A$. In calculus we sometime denote $a_n = f(n)$ and $(a_n)_{n=0}^{\infty} = f$.

Example 3.30. The sequence $1, \frac{1}{2}, \frac{1}{3}, \ldots$ is formally the function $f : \mathbb{N} \to \mathbb{Q}$, $f = \{\langle n, \frac{1}{n+1}\rangle \mid n \in \mathbb{N}\}$ satisfying $f(n) = \frac{1}{n+1}$.

Definition 3.31. Let $f : A \to B$ be a function. The *domain* of f is simply A, we denote $\mathrm{dom}(f) = A$. The *range* of f is B and we denote $\mathrm{rng}(f) = B$. The *image* of f is the set $\mathrm{Im}(f) = \{f(a) \mid a \in A\}$.

Definition 3.32. Let A, B be two sets. We denote *the set of all functions* from A to B by

$$^{A}B = \{f \in P(A \times B) \mid f \text{ is a function from } A \text{ to } B\}$$

Example 3.33. Let F_2 be the relation from $^{\mathbb{R}}\mathbb{R}$ to $\mathbb{R}$ defined by

$$F_2 = \{\langle f, r\rangle \in {}^{\mathbb{R}}\mathbb{R} \times \mathbb{R} \mid \langle 2, r\rangle \in f\}.$$

Prove that F is a function.

Proof. Total: We nee to probe that for every $f \in \mathbb{R}\mathbb{R}$ (here the domain of F_2 is itself a set of functions!) there is $r \in \mathbb{R}$ such that $\langle f, r\rangle \in F$. Let $f \in \mathbb{R}\mathbb{R}$. we need to find $r \in \mathbb{R}$ such that $\langle 2, r\rangle \in f$. Since f is a function from $\mathbb{R}$ to $\mathbb{R}$, it is in particular a total relation on $\mathbb{R}$, and since $2 \in \mathbb{R}$, there exists $r \in \mathbb{R}$ such that $\langle 2, r\rangle \in f$, hence $\langle f, r\rangle \in F_2$.
Univalent: We want to prove that for any $f \in {}^{\mathbb{R}}\mathbb{R}$ and any $r_1, r_2 \in \mathbb{R}$, if $\langle f, r_1\rangle, \langle f, r_2\rangle \in F_2$ then $r_1 = r_2$. Supposet that $\langle f, r_1\rangle, \langle f, r_2\rangle \in F_2$, then by definition $\langle 2, r_1\rangle, \langle 2, r_2\rangle \in f$. Since f is a function, it is in particular univalent and therefore $r_1 = r_2$. $\square$

Note that we have $F_2(f) = f(2)$ for every function $f \in {}^{\mathbb{R}}\mathbb{R}$.

In order to discard the need to formulate functions as sets of pair we simply need to understand when two functions are equal[6].

Theorem 3.34. *Let f, g be any function. Then the following are equivalent:*

1. *$f = g$ (equality of sets of pairs!).*
2. *$\mathrm{dom}(f) = \mathrm{dom}(g)$ and for every $x \in \mathrm{dom}(f)$, $f(x) = g(x)$.*

Proof. $\underline{\Rightarrow}$: Suppose that $f = g$, then clearly $\mathrm{dom}(f) = \mathrm{dom}(g)$. Let $x \in \mathrm{dom}(f)$, and denote by $f(x) = y$. Then $\langle x, y\rangle \in f$ and since $f = g$ $\langle x, y\rangle \in g$ hence $g(x) = y = f(x)$.

$\underline{\Leftarrow}$: For the other direction, suppose that $\mathrm{dom}(f) = \mathrm{dom}(g) =: A$ and that for every $x \in A$, $f(x) = g(x)$. We want to prove that $f = g$ (equality of sets)

- $\underline{\subseteq}$: Let $\langle x, y\rangle \in f$. Then $x \in A$ and $f(x) = y$. Thus $x \in \mathrm{dom}(g)$ and $f(x) = y = g(x)$ which implies that $\langle x, y\rangle \in g$.
- $\underline{\supseteq}$: The other direction is symmetric.

$\square$

Problem 6. *Let $f : A \to B$ be a function.*

1. *Prove that if $X \subseteq A$, then $f \cap X \times B$ is a function and equals $f \restriction X$.*
2. *Show that if $f : A \to B$, $g : B \to C$ are functions then $g \circ f$ (the composition of the relations) is a function from A to C and that for every $a \in A$, $g \circ f(a) = g(f(a))$.*
3. *Prove that if f is one-to-one and onto B then f^{-1} (the inverse relation) is a function and moreover that $f^{-1} \circ f = Id_A$ and $f \circ f^{-1} = Id_B$.*

3.4. **Relations on a single set.** The first kind of relations we are interested in are relations R from a set A to itself.

Definition 3.35. A relation R from A to A (i.e. $R \subseteq A^2$) is called a relation *on the set A*.

For example, $\leq_{\mathbb{N}}$ is a relation of $\mathbb{N}$, id_A is a relation on A and the divisibility relation S is a relation in $\mathbb{Z}$.

Example 3.36. Let us denote by $\subseteq_A = \{\langle X, Y\rangle \in P(A)^2 \mid X \subseteq Y\}$. Then $\subseteq_A$ is a relation on $P(A)$.

Instead of writing for example $\langle 2, 3\rangle \in \leq_{\mathbb{N}}$ or $\langle \{1\}, \{39, 1, 14\}\rangle \in \subseteq_{\mathbb{Z}}$, we would like to keep the usual notation that $2 \leq_{\mathbb{N}} 3$ and $\{1\} \subseteq_{\mathbb{Z}} \{39, 1, 14\}$. Hence we have the following notation:

Notation 3.37. Given a general relation R on a set A, we define $aRb \equiv \langle a, b\rangle \in R$.

In order to develop some theory and prove interesting theorems about relations, we will need to add more structure/properties to the relation. The most important kind of relations on a single set are *equivalence relations* and *orders*.

[6] As we did with tuples.

3.5. **Equivalence relations.** As we have seen previously, sets are equal if and only if they have the same elements. This is a quit rigid equality. There are mathematical theories where it is convenient to identify between two objects although they are not equal as sets, we say that they are *equivalent*. For example, to define a rational numbers $\frac{n}{m}$ from the integers, it is natural to identify it with the pair $\langle n, m\rangle$. However, note that while $\frac{1}{2} = \frac{2}{4}$, the pairs $\langle 1,2\rangle, \langle 2,4\rangle$ are distinct. What we usually do, is to set some criterion to determine when two objects are equivalent. Formally, this would mean that we have some relation R on a set A, and two members $a, b \in A$ will be equivalent if aRb. In our example of rationals, we would need to find a criterion which makes $\langle 1,2\rangle, \langle 2,4\rangle$ equivalent for examples, and not only them, but also $\langle 4,2\rangle, \langle 8,2\rangle$ and $\langle -1,9\rangle, \langle 2,-18\rangle$ and so on.

Example 3.38. To find the right criteria for the rations, we need to express the equality $\frac{a}{b} = \frac{c}{d}$ in terms of integers, so let simply cross-multiply the equation and get $ad = bc$. Going back to the beginning, we define a relation R on the **set of pairs** $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$. Note that this is not a relation on $\mathbb{Z}$, rather then on pairs, and we exclude 0 by only considering pairs of the form $\langle a, b\rangle$ where $b \neq 0$. Now we set the criterion that $\langle a, b\rangle R\langle c, d\rangle$ (namely, the pairs $\langle a, b\rangle$ and $\langle c, d\rangle$ are equivalent) if and only if $ad = bc$. Formally, we define the relation R as follows:

$$R = \Big\{\langle\langle a, b\rangle, \langle c, d\rangle\rangle \in (\mathbb{Z} \times \mathbb{Z} \setminus \{0\})^2 \mid ad = bc\Big\}$$

Since equivalence relations imitate equality, there are some necessary properties which must be posed on a general relation in order for it to be an equivalence relation:

Definition 3.39 (Properties of relations and equivalence relation)**.** Let R be a relation on a set A. We say that:

(1) R is *reflexive* (on A) if: $\forall a \in A.aRa$.
(2) R is *symmetric* if: $\forall a, b \in A.aRb \Rightarrow bRa$.
(3) R is *transitive* if: $\forall a, b, c \in A.(aRb) \wedge (bRc) \Rightarrow aRc$.
(4) R is an *equivalence relation* if it is reflexive, symmetric and transitive.

Example 3.40. (1) Let us give some non mathematical relations on the "set" of all humans to illustrate these properties:
   (a) The *brotherhood relation*: two humans x, y are brothers if and only if they have the same biological parents.[7]
   <u>The brotherhood relation is reflexive</u>: Indeed, **every** human x is a brother of himself, as by this definition x has the same two biological parents as himself.
   <u>The brotherhood relation is symmetric</u>: If x is a brother of y then clearly y is a brother of x because they both have the same biological parents.

[7]This is simply a convenient choice of definition, one can consider other definitions for brotherhood.

The brotherhood relation is transitive: Suppose that x is a brother of y and y is a brother of z. Then x as the same two biological parents as y and y has the same two biological parents as z. Then x has the same two biological parents as z, hence x and z are brothers

We conclude that the brotherhood relation is an equivalence relation.

(b) The *descendent relation*: for two humans (dead or alive) we say that x is a descendent of y (or that y is an ancestor of x) is x is the son of a son of a son ... of a son of y. It is a matter of definition if this relation is reflexive, namely, is x is a descendent of himself. It is clearly transitive.This is not symmetric, since for example, Jeffery Jordan is a descendent (the son of) Michael Jordan, but Michael Jordan is not the a descendent of Jeffery Jordan.[8]

(2) Let $A = \{1, 2, 3, 4, 5, 6\}$ then

$$E = \{\underbrace{\langle 1,1\rangle, \langle 2,2\rangle, \langle 3,3\rangle, \langle 4,4\rangle, \langle 5,5\rangle, \langle 6,6\rangle}_{id_A}, \langle 1,5\rangle, \langle 5,1\rangle, \langle 2,3\rangle, \langle 3,2\rangle, \langle 3,6\rangle, \langle 6,3\rangle, \langle 2,6\rangle, \langle 6,2\rangle\}$$

is an equivalence relation on A.

(3) Among the most important equivalence relations is the congruence relation. Recall that for a natural number $n > 0$ and two integers z_1, z_2 we say that $z_1 \equiv z_2 \mod n$ if $z_1 \mod n = z_2 \mod n$. In order to avoid the use modulo in the definition congruency, we can formulate it as follows:

$$E_n = \{\langle z_1, z_2\rangle \in \mathbb{Z}^2 \mid z_1 - z_2 \text{ is divisible by } n\}$$

Let us prove that E_n is an equivalence relation.

Reflexive: we want to prove that for every $z \in \mathbb{Z}$, zE_nz. Let $z \in \mathbb{Z}$, we want to prove that $z - z = 0$ is divisible by n, but this is true sine every number divides 0(recall the formal definition of divisibility and prom this easy fact!).

Symmetric: We want to prove that for every $z_1, z_2 \in \mathbb{Z}$, if $z_1E_nz_2$ then $z_2E_nz_1$. Let $z_1, z_2 \in \mathbb{Z}$ and suppose (this is an implication!) that $z_1E_nz_2$, we want to prove that $z_2E_nz_1$. [9] By definition of E_n, we conclude that n divides $z_1 - z_2$ and therefore there is $k \in \mathbb{Z}$ such that $z_1 - z_2 = k \cdot n$. Hence $z_2 - z_1 = (-k) \cdot n$ and also $-k \in \mathbb{Z}$. It follows again by the definition of E_n that $z_2E_nz_1$.

Transitive: Suppose that $z_1E_nz_2$ and $z_2E_nz_3$, we want to prove that $z_1E_nz_3$. By definition of E_n, this means that n divides $z_1 - z_2$ and

[8] Note that in order to prove that a relation is not reflexive/symmetric/transitive we should always give a **specific** counter example, since these properties are universal properties and therefore their negation is an existential property.

[9] Usually, we will start directly with "suppose that $z_1E_nz_2$, we want to prove that $z_2E_nz_1$".

also $z_2 - z_3$. By definition f divisibility, there are $k_1, k_2 \in \mathbb{Z}$ such that $z_1 - z_2 = k_1 n$ and $z_2 - z_3 = k_2 n$. Summing the two equations, we get:

$$z_1 - z_3 = (z_1 - z_2) + (z_2 - z_3) = k_1 n + k_2 n = (k_1 + k_2)n)$$

Since $k_1 + k_2 \in \mathbb{Z}$, it follows that $z_1 - z_3$ is divisible by n. By the definition of E_n, it follow that $z_1 E_n z_3$.

We conclude that E_n is an equivalence relation.

(4) $S = \{\langle n, m\rangle \in \mathbb{Z}^2 \mid \exists k \in \mathbb{Z} n + k^2 = m\}$ is reflexive, not symmetric, since for example $0S1$ (as $0 + 1^2 = 1$) but $1 \not S 0$ (prove that!). It is not transitive since for example $1 + 1^2 = 2$ and $2 + 1^2 = 3$ however $3 - 1 = 2$ is not a square of a natural (or even rational) number.

(5) The following relation will serve to construct the integers from the natural numbers. On $\mathbb{N}^2$ we define the following relation

$$\sim_Z = \Big\{\langle\langle n, m\rangle, \langle k, l\rangle\rangle \in (\mathbb{N} \times \mathbb{N})^2 \mid n + l = m + k\Big\}$$

Problem 7. *Prove that $\sim_Z$ is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.*

(6) Let us prove that the relation

$$\sim_Q = \Big\{\langle\langle a, b\rangle, \langle c, d\rangle\rangle \in (\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))^2 \mid ad = bc\Big\}$$

we use to construct the rational numbers is indeed an equivalence relation on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$:

Reflexive: Let $\langle a, b\rangle \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$, [10] we want to prove that $\langle a, b\rangle \sim_Q \langle a, b\rangle$. This follows, since $ab = ab$ and by the definition of $\sim_Q$.

Symmetric: Suppose that $\langle a, b\rangle \sim_Q \langle c, d\rangle$, we want to prove that $\langle c, d\rangle \sim_Q \langle a, b\rangle$. By our assumption we see that $ad = bc$, and since we can switch the order of number multiplication we get that $da = cb$ and therefore $\langle c, d\rangle \sim_Q \langle a, b\rangle$.

Transitive: Suppose that $\langle a, b\rangle \sim_Q \langle c, d\rangle$, $\langle c, d\rangle \sim_Q \langle e, f\rangle$. We want to prove that $\langle a, c\rangle \sim_Q \langle e, f\rangle$. By the assumption we have that $ad = bc$ and $cf = de$. Note that $adf = bcf = bde$ and since[11] $d \neq 0$, we can eliminate it from the equation to see that $af = be$. By definition of $\sim_Q$, it follows that $\langle a, b\rangle \sim_Q \langle e, f\rangle$.

It follows that $\sim_Q$ is an equivalence relation.

(7) For any set A, the identity relation id_A and $A \times A$ are always equivalence relations on the set A.

(8) Here are two examples of equivalence relations on $\mathbb{R}^3$:

$$H_1 = \{\langle\langle a, b, c\rangle, \langle a', b', c'\rangle\rangle \in \mathbb{R}^3 \mid a = a'\}$$

$$H_2 = \{\langle\langle a, b, c\rangle, \langle a', b', c'\rangle\rangle \in \mathbb{R}^3 \mid a + b + c = a' + b' + c'\}.$$

[10] We want to prove that $\forall a \in A. a \sim_Q a$. In our case $A = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ is a set of pairs (!) hence we want to prove that $\forall \langle a, b\rangle \in \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}). \langle a, b\rangle \sim_Q \langle a, b\rangle$.

[11] Indeed $\langle c, d\rangle \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\}$, $c \in \mathbb{Z}$ and $d \in \mathbb{Z} \setminus \{0\}$. Therefore $d \neq 0$.

The equivalence criterion that the relation H_1 sets is to identify between triples with the same first coordinate. The equivalence that H_2 sets is to identify triples with the same sum.

(9) Here is an equivalence relations on the set $P(\mathbb{N}) \setminus \{\emptyset\}$:
$$T_1 = \{\langle X, Y\rangle \in (P(\mathbb{N}) \setminus \{\emptyset\})^2 \mid \min(X) = \min(Y)\}$$
T_1 identifies sets with the same minimal elements. Here is an equivalence relation on the set $P(\mathbb{N})$:
$$T_2 = \{\langle X, Y\rangle \in (P(\mathbb{N}) \setminus \{\emptyset\})^2 \mid X \cap \mathbb{N}_{even} = X \cap \mathbb{N}_{odd}\}$$
T_2 identifies sets which includes exactly the same even numbers.

Back to our example of the rational numbers, what is the object $\frac{1}{2}$? is it $\langle 1, 2\rangle$ or is it $\langle 2, 4\rangle$? the definition of $\frac{1}{2}$ is just the set of those pairs $\{\langle 1, 2\rangle, \langle 2, 4\rangle, \rangle 3, 6\rangle, \langle -1, -2\rangle....\}$. The point is that we "glue" together all the conditions which are equivalent to $\langle 1, 2\rangle$. Formally, we call this an *equivalence class*:

Definition 3.41. Let E be an equivalence relation on a set A. The *equivalence class* of an element $a \in A$ is the set of all conditions $b \in A$ such that a is E-equivalent to b. Formally, we denote the equivalence class of a by
$$[a]_E = \{b \in A \mid aEb\}$$
An E-equivalent class is just $[a]_E$ for some $a \in A$.

Example 3.42. We use the same notations from the previous example.

(1) In the brotherhood relation we have for example the following equivalence classes:
$$[\text{Orville Wright}]_{brotherhood} = \{\text{Orville Wright, Wilbur Wright}\}$$
$$[\text{Steph Curry}]_{brotherhood} = \{\text{Steph Curry, Seth Curry, Sydel Curry}\}$$
$$[\text{Kim Kardashian}]_{brotherhood} = \{\text{Kim Kard., Kourtney Kard., Khloé Kard., Rob Kard.}\}$$

(2) For $A = \{1, 2, 3, 4, 5, 6\}$ and E from example (2), We have that:
$$[1]_E = \{1, 5\}$$
$$[2]_E = \{2, 3, 6\}$$
$$[3]_E = \{2, 3, 6\}$$
$$[4]_E = \{4\}$$
$$[5]_E = \{1, 5\}$$
$$[6]_E = \{2, 3, 6\}$$
This is not a coincidence that $[1]_E = [5]_E$ and that $[2]_E = [3]_E = [6]_E$, can you guess way?

(3) The equivalence classes of E_n are
$$[0]_{E_n} = \{0, n, -n, 2n, -2n, 3n, ...\} = \{zn \mid z \in \mathbb{Z}\}$$
$$[1]_{E_n} = \{1, n-1, -n+1, 2n-1, -2n+1, ...\} = \{zn + 1 \mid z \in \mathbb{Z}\}$$
A general equivalence class is just:
$$[i]_{E_n} = \{zn + i \mid z \in \mathbb{Z}\}$$
and $i \equiv j \mod n$ if and only if $[i]_{E_n} = [j]_{E_n}$.

(4) Using equivalence classes and the equivalence relation $\sim_Q$ we can now formally define the rational number $\frac{n}{m} = [\langle n, m\rangle]_{\sim_Q}$. For example, the number $\frac{1}{2}$ is just $[\langle 1, 2\rangle]_{\sim_Q}$. We will see later that $[\langle 1, 2\rangle]_{\sim_Q} = [\langle 2, 4\rangle]_{\sim_Q}$ for example, where the last equality is an actual set equality!

(5) As for $\sim_Z$, we think of a pair $\langle n, m\rangle \in \mathbb{N}^2$ and representing $n-m$. So we identify between $n \in \mathbb{N}$ with $[\langle n, 0\rangle]_{\sim_Z}$ and define $-n = [\langle 0, n\rangle]_{\sim_Z}$.

(6) The equivalence class of a general triple $\langle a, b, c\rangle \in \mathbb{R}^3$ has the form:
$$[\langle a, b, c\rangle]_{H_1} = \{\langle a, x, y\rangle \mid x, y \in \mathbb{R}\}$$
and
$$[\langle a, b, c\rangle]_{H_2} = \{\langle x, y, (a + b + c - x - y)\rangle \mid x, y \in \mathbb{R}\}$$

(7) We have fore example
$$[\{4, 7, 3, 22\}]_{T_1} = \{X \in P(\mathbb{N}) \mid 3 = \min(X)\}$$
and
$$[\{4, 7, 3, 22\}]_{T_2} = \{X \in P(\mathbb{N}) \mid X \cap \mathbb{N} = \{2, 22\}\}$$

Proposition 3.43. *Let E be an equivalence relation on A. Then for every $a, b \in A$:*

(1) Either $[a]_E = [b]_E$.

(2) Or $[a]_E \cap [b]_E = \emptyset$

Moreover, $[a]_E = [b]_E$ if and only if aEb.

Proof. Let $a, b \in A$. We formally need to prove a $\vee$-statement. Let us split into cases:

(1) Suppose $[a]_E \cap [b]_E = \emptyset$, the (2) holds and we are done.

(2) Suppose $[a]_E \cap [b]_E \neq \emptyset$. We want to prove that $[a]_E = [b]_E$, which is sets equality. Let us prove a double inclusion:

  (a) $\underline{[a]_E \subseteq [b]_E}$: Let $x \in [a]_E$. We want to prove that $x \in [b]_E$. Let $c \in [a]_E \cap [b]_E$, which exists by the assumption in this case. By definition of equivalence relation, xEa, cEa and cEb.
  - By symmetry, since cEa, then aEc.
  - By transitivity, since xEa and aEc, then xEc.
  - Again by trasitivity since xEc and cEb, xEb.

  By the definition of equivalence class it follows that $x \in [b]_E$.

  (b) $\underline{[b]_E \subseteq [a]_E}$: Follows from the symmetry between a and b.

This concludes the proof that $[a]_E = [b]_E$ or $[a]_R \cap [b]_E = \emptyset$. For the moreover part, we nee to prove a double implication:

(1) $\Longrightarrow$: Suppose that $[a]_E = [b]_E$, we need to prove that aEb. Since E is reflexive, aEa and therefore $a \in [a]_E$. By the equality of the set $[a]_E = [b]_E$ we conclude that $a \in [b]_E$ and by the definition of equivalence class we conclude that aEb.
(2) $\Longleftarrow$: Suppose that aEb, we need to prove that $[a]_E = [b]_E$. Again since E is reflexive we have that $a \in [a]_E$ and by the definition of equivalence class we have that $a \in [b]_E$. Thus $a \in [a]_E \cap [b]_E$, which means that $[a]_E \cap [b]_E \neq \emptyset$. By the first part, this must means that $[a]_E = [b]_E$.

$\square$

Corollary 3.44. *The following are equivalent:*

(1) $a \not E b$.
(2) $[a]_E \neq [b]_E$.
(3) $[a]_E \cap [b]_E = \emptyset$.

Proof. exercise. $\square$

Definition 3.45. Let E be an equivalence relation on A. The *quotient set* of A by E (a.k.a "A modulo E") is the set of **all** equivalence classes.[12]. We denote it by[13]

$$A/E = \{[a]_E \mid a \in A\}$$

Example 3.46. (1) The "set" Humans/brotherhood consist of all possible equivalence classes, each equivalence class is the set of siblings from a given family. We can label each equivalence class according to the family name and think of the quotient

$$\text{Humans/brotherhood} = \{\text{"The Kardeshians"}, \text{"The Curry's"}, \text{"The Wright's"}, ...\}$$

(2) $A/E = \{\{1,5\},\{2,3,6\},\{4\}\}$.
(3) We have that

$$\mathbb{Z}/E_n = \{\{zn + i \mid z \in \mathbb{Z}\} \mid i = 0,1,2,...,n-1\}$$

Since each equivalence class in E_n is associated with a residue modulo n, we think of $\mathbb{Z}/E_n$ as the sets of residues modulo n.
(4) The integers are defined by $\mathbb{Z} = \mathbb{N}^2/\sim_Z$
(5) The rational numbers are defined as

$$\mathbb{Q} = (\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})/\sim_Z$$

(6)

$$\mathbb{R}^3/H_1 = \{\{\langle a,x,y\rangle \mid x,y \in \mathbb{R}\} \mid a \in \mathbb{R}\}$$

[12] Needless to say, without repetitions.
[13] Do not confused A/E with set difference $A \setminus E$.

Here every equivalence class can be identified with a single real number a.

$$\mathbb{R}^3/H_2 = \{\{\langle x, y, (s-x-y)\rangle \mid x, y \in \mathbb{R}\} \mid s \in \mathbb{R}\}$$

Also here the equivalence classes can be identifies with a single real number s which represents the sum $a+b+c$.

(7)

$$(P(\mathbb{N}) \setminus \{\emptyset\})/T_1 = \{\{X \in P(\mathbb{N}) \setminus \{\emptyset\} \mid \min(X) = n\} \mid n \in \mathbb{N}\}$$

And each equivalence class can be identified with a natural number.

$$P(\mathbb{N})/T_2 = \{\{X \in P(\mathbb{N}) \mid X \cap \mathbb{N}_{even} = Y\} \mid Y \in P(\mathbb{N}_{even})\}$$

And each equivalence class can be identified with a set of even numbers.

Definition 3.47 (Partition)**.** Let A be any set. A partition of the set A is any set $\Pi \subseteq P(A)$ such that:

(1) $\emptyset \notin \Pi$.
(2) $\cup\Pi = A$.
(3) If $X, Y \in \Pi$, $X \neq Y$, then $X \cap Y = \emptyset$.

Example 3.48. (1) $\{\{1,5\},\{2,3,6\},\{4\}\}$ is a partition of $\{1,2,3,4,5,6\}$.
(2) $\{\mathbb{N}_{even}, \mathbb{N}_{odd}\}$ is a partition of $\mathbb{N}$.

Corollary 3.49. *If E is an equivalent relation on A then A/E is a partition of A.*

Proof. Follows directly from Proposition 3.43. □

Theorem 3.50. *Let Π be a partition on A. Let R_Π be the relation on A defined by*

$$xR_\Pi y \iff \exists B \in \Pi,\ x, y \in B$$

Then:

(1) R_Π is an equivalence relation on A.
(2) $A/R_\Pi = \Pi$.

Proof. (1) Let us prove that R_Π is an equivalence relation:

$\underline{R_\Pi \text{ is reflexive}}$: Let $a \in A$, since $\cup\Pi = A$, there is $X \in \Pi$ such that $a \in X$ and therefore by definition of R_Π, $\langle a, a\rangle \in R_\Pi$.

$\underline{R_\Pi \text{ is symmetric}}$: Suppose that $\langle a, b\rangle \in R_\Pi$, then there is $X \in \Pi$ such that $a, b \in X$. Hence $b, a \in X$, and therefore $\langle b, a\rangle \in R_\Pi$.

$\underline{R_\Pi \text{ is transitive}}$: Suppose that $\langle a, b\rangle \in R_\Pi$ and $\langle b, c\rangle \in R_\Pi$, then there are $X, Y \in \Pi$ such that $a, b \in X$ and $b, c \in Y$. Since $b \in X \cap Y$, we conclude that $X \cap Y \neq \emptyset$ and since Π is a partition, $X = Y$. hence $a, c \in X$ and therefore $\langle a, c\rangle \in R_\Pi$.

(2) To see that $A/R_\Pi = \Pi$ we prove a double inclusion:

$\underline{\subseteq}$: Let $[a]_{R_\Pi} \in A/R_\Pi$. Then there is $X \in \Pi$ such that $a \in X$. We claim that $[a]_{R_\Pi} = X$ and from this it follows that $[a]_{R_\Pi} \in \Pi$. Again we prove it by double inclusion:

- $\subseteq$: Let $b \in [a]_{R_\Pi}$, then $aR_\Pi b$ and therefore there is $Y \in \Pi$ such that $a, b \in Y$. Since $a \in X \cap Y$ we conclude that $X = Y$ and therefore $b \in X$.
- $\supseteq$: If $b \in X$ then $a, b \in X \in \Pi$ and therefore $aR_\Pi b$ which implies that $b \in [a]_{R_\Pi}$.

$\subseteq$: Let $X \in \Pi$, we want to prove that $X \in A/R_\Pi$. Since $X \neq \emptyset$, pick any $a \in X$, we claim that $X = [a]_{R_\Pi} \in A/R_\Pi$. The prof is similar to the previous part.

$\square$

Problem 8. *If R is an equivalence relation on A, then $R = R_{A/R}$.*

Definition 3.51. A relation R does not depend on the choice of representatives of E if whenever aEa' and bEb' then $aRb \Rightarrow a'Rb'$.

Example 3.52. (1) $[\langle n, m\rangle]_{\sim_Z} + [\langle n', m'\rangle]_{\sim_Z} = [\langle n+n', m+m'\rangle]_{\sim_Z}$ Does not depend on the choice of representatives.

Proof. If $\langle n_1, m_1\rangle \sim_Z \langle n_2, m_2\rangle$ and $\langle n_1', m_1'\rangle \sim_Z \langle n_2', m_2'\rangle$, then $n_1 + m_2 = n_2 + m_1$ and $n_1' + m_2' = n_2' + m_1'$. We would like to prove that

$$\langle n_1 + n_1', m_1 + m_1'\rangle \sim_Z \langle n_2 + n_2', m_2 + m_2'\rangle$$

. To see this,

$$n_1+n_1'+m_2+m_2' = n_1+m_2+n_1'+m_2' = n_2+m_1+n'2+m_1' = m_1+m_1'+n_2+n_2'$$

as wanted. $\square$

(2) $[n]_{E_m} \cdot [n']_{E_m} = [n \cdot n']_{E_m}$ does not depend on the choice of representative.

Proof. Suppose that $nE_m n_0$ and $n'E_m n_0'$ we want to prove that $nn'E_m n_0 n_0'$. Note that $m | n - n_0$ and $m | n' - n_0'$. Hence

$$nn' - n_0 n_0' = nn' - n'n_0 + n'n_0 - n_0 n_0' = n'(n - n_0) + n_0(n' - n_0').$$

This is a sum of two numbers which are divisible by m and therefore $nn' - n_0 n_0'$ is divisible by m. $\square$

(3) $F([\langle a, b, c\rangle]_{H_1}) = a$ Does not depend on the choice of representatives. Clearly if $\langle a, b, c\rangle H_1 \langle a', b', c'\rangle$, then $a = a'$ and therefore $F([\langle a, b, c\rangle]_{H_1}) = F([\langle a', b', c'\rangle]_{H_1})$.

3.6. **Ordered sets.**

Definition 3.53. We say that a relation R on A is:

1. Weakly anti-symmetric if: for all $a, b \in A$, if aRb and bRa then $a = b$.
2. Strongly anti-symmetric if: for all $a, b \in A$, if $a\ R\ b$ then $b \not R\ a$.
3. Weak partial order if R is reflexive, transitive, and weakly anti-symmetric.
4. Strong partial order if R is transitive and strongly anti-symmetric.

Definition 3.54. A partial order R (either weak or strong) is called total/linear if every $a, b \in A$ are R-comparable, namely, if

$$aRb \vee bRa \vee a = b$$

Problem 9. *A relation R on A is called anti-reflexive if $\forall a \in A,\ a \not R a$. Prove that the following are equivalent:*

(1) R is strongly anti symmetric.
(2) R is anti reflexive and weakly anti symmetric.

Example 3.55. (1) the regular order $<$ of real numbers is a strong linear order on $\mathbb{R}$ and $\leq$ is a weak linear order on $\mathbb{R}$.

(2) $\subsetneq$ is a strong non-linear order on $P(A)$ for (almost) every set A and $\subseteq$ is a weak non-linear order on $P(A)$ for (almost) every set A.

Proof. Let us some weak anti-symmetry of $\subseteq$: Let $X, Y \in P(A)$, if $X \subseteq Y$ and $Y \subseteq X$ then $X = Y$ by double inclusion. If A has at least two elements $a, b \in A$ then $\{a\}, \{b\}$ are incomparable in $\subseteq$-relation and therefore $\subseteq$ is not linear on $P(A)$. $\square$

(3) The lexicographic order. Suppose that $<_A$ is a partial order on A and $<_B$ is a partial order on B. Define the lexicographic order on $A \times B$ by

$$\langle a, b\rangle <_{Lex} \langle a', b'\rangle \text{ if and only if } a <_A a' \vee (a = a' \wedge b <_B b')$$

We leave transitivity to the reader. Let us prove that it is strongly anti-symmetric. Assume that $\langle a, b\rangle <_{Lex} \langle a', b'\rangle$. We want to prove that $\neg(\langle a', b'\rangle <_{Lex} \langle a, b\rangle)$, namely, that $\neg(a' <_A a)$ and $\neg(a' = a \wedge b' <_B b)$. Let us split into cases:

(a) If $a <_A a'$, then by anti-symmetry of $<_A$, $a \neq a'$ and $\neg(a' <_A a)$. Hence we are done.
(b) If $a = a'$ and $b <_B b'$, then $\neg a' <_A a$ since $<_A$ is anti-reflexive. Also since $<_B$ is anti-reflexive, $\neg(b' <_B b)$ and agin we are done.

(4) The domination order on ${}^{\mathbb{N}}\mathbb{N}$ is defined by

$$f \leq g \iff \forall n \in \mathbb{N}, f(n) \leq g(n)$$

This is a weak order on ${}^{\mathbb{N}}\mathbb{N}$. The eventual domination order is defined by

$$f \leq^* g \iff \exists N \forall n \geq N,\ f(n) \leq g(n)$$

This is not an order on ${}^{\mathbb{N}}\mathbb{N}$ since there are $f \neq g$ such that $f \leq^* g$ and $g \leq^* f$ (find an example!). However:

Problem 10. *Let*

$$E = \{\langle f, g\rangle \in ({}^{\mathbb{N}}\mathbb{N})^2 \mid \exists N \forall n \geq N, f(n) = g(n)\}$$

(a) Prove that E is an equivalence relation.
(b) Prove that the relation $[f]_E \leq^* [g]_E$ iff $f \leq^* g$ does not depend on the choice of representatives and partially orders ${}^{\mathbb{N}}\mathbb{N}/E$.

(5) Define $<_{Lex}$ in ${}^{\mathbb{N}}\mathbb{N}$ by $f <_{Lex} g$ iff $f \neq g \wedge f(n_{f,g}) < g(n_{f,g})$ where $n_{f,g} = \min\{m \in \mathbb{N} \mid f(m) \neq g(m)\}$.

Exercise 6. *Prove that $<_{Lex}$ is an order on ${}^{\mathbb{N}}\mathbb{N}$.*

Solution 7. *Let us prove it is transitive. Suppose that $f <_{Lex} g$ and $g <_{Lex} h$. Let us split into cases:*

(a) If $n_{f,g} = n_{g,h} = n^*$, then for every $n < n^*$, $f(n) = g(n) = h(n)$ hence $n_{f,h} \geq n^*$. Also $f(n^*) < g(n^*) < h(n^*)$ and so $f(n^*) < h(n^*)$. Thus $n^* = n_{f,h}$ and $f(n_{f,h}) < h(n_{f,h})$ as wanted.

(b) If $n_{f,g} < n_{g,h}$, we have for every $n < n_{f,g}$ $f(n) = g(n) = h(n)$, hence $n_{f,h} \geq n_{f,g}$ and also $f(n_{f,g}) < g(n_{f,g}) = h(n_{f,g})$. Hence $n_{f,h} = n_{f,g}$ and $f(n_{f,h}) < h(n_{f,h})$ as desired.

(c) The case $n_{f,g} > n_{g,h}$ is similar.

Let us prove that $<_{Lex}$ is anti-symmetric. Suppose that $f <_{Lex} g$, then $f(n_{f,g}) < g(n_{f,g})$. Hence $\neg(g(n_{f,g}) < f(n_{f,g}))$ so $\neg(g <_{Lex} f)$.

Definition 3.56. A pair $\langle A, \leq_A \rangle$ is called an ordered set or a poset (partially ordered set).

Definition 3.57. Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B$ be two ordered sets. A function $f : AB$ is called order-preserving if:

$$\forall a_1, a_2 \in A,\ a_1 \leq_A a_2 \Leftrightarrow f(a_1) \leq_B f(a_2)$$

f is called an *isomorphism* if it is an order-preserving bijection. f is an *embedding* if it is order-preserving and injective.

Example 3.58. (1) Consider the regular order $<$ on $\mathbb{N}$ and $<_{Lex}$ on $\mathbb{N} \times \mathbb{N}$. The function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ defined by $f(n) = \langle 5, n \rangle$ is an embedding. It is clearly injective, to see it is order-preserving, let $n_1, n_2 \in \mathbb{N}$, then if is a straightforward verification that

$$n_1 < n_2 \iff f(n_1) = \langle 5, n_1 \rangle <_{Lex} \langle 5, n_2 \rangle$$

(2) $f : \mathbb{N} \to P(\mathbb{N})$ defined by $f(n) = \{0, ..., n\}$ is an embedding of $\mathbb{N}$ with $\leq_{\mathbb{N}}$ into $P(\mathbb{N})$ with $\subseteq$.

(3) $f : \mathbb{N} \to \mathbb{N}_{even}$ defined by $f(n) = 2n$ is an isomorphism of $\mathbb{N}$ with the regular order into $\mathbb{N}_{even}$ with the regular order.

Definition 3.59. We say that $\langle A, \leq_A \rangle \simeq \langle B, \leq_B \rangle$ is there exists an isomorphism $f : A \to B$. We say that $\langle A, \leq_A \rangle \precsim \langle B, \leq_B \rangle$ if there is an embedding $f : A \to B$.

Exercise 7. *Prove that $\mathbb{Q}$ is not isomorphic to $\mathbb{Z}$.*

Proof. Suppose otherwise and let $f : \mathbb{Q} \to \mathbb{Z}$ be an isomorphism. Consider $f(0) = z$. Since f is an isomorphism, there is $q \in \mathbb{Q}$ such that $f(q) = z + 1$. Since $z < z + 1$ and f is order-preserving, $0 < q$. Hence $0 < \frac{q}{2} < q$. Then $z < f(\frac{q}{2}) < z+1$ is an integer strictly between z and $z+1$, contradiction. $\square$

4. Number systems

4.1. **Natural numbers.** We want a definition which is purely set-theoretic.

Definition 4.1. Define $0 = \emptyset$. For any set A define $A + 1 = A \cup \{A\}$.

Why not define the natural numbers by induction? Since induction itself requires the natural numbers and we end up with a circular definition. We need to take a different approach

Definition 4.2. A set X is inductive if:

(1) $0 \in X$.
(2) $\forall x \in X . x + 1 \in X$.

Definition 4.3. A natural number is a set x such that for every inductive set B, $x \in B$.

Clearly, 0 is a natural number and also $0+1$, $(0+1)+1$, $((0+1)+1)+1$.

Exercise 8. *The intersection of inductive sets is an inductive set.*

Proposition 4.4. *The set of natural numbers (if it exists) is an inductive set and is included in every inductive set.*

Proof. The second part is immediate from the definition of natural numbers. To see that the set of natural numbers is inductive, we mentioned that 0 is a natural number hence condition (1) of "inductive set" is satisfied. Suppose that x is a natural number and let us prove that $x + 1$ is a natural number. Let B be any inductive set, then $x \in B$ by definition of a natural number. Since B is inductive, $x + 1 \in B$ and therefore $x + 1$ is a natural number. $\square$

We cannot prove the existence of an inductive set based on the axioms so far.

Axiom (Ax7. Infinity)**.** There exists an inductive set.

Corollary 4.5. *Ax7 holds if and only the set of natural numbers exists.*

Proof. We have already mentioned that if the set of natural numbers exists then it is an inductive set. For the other direction, we can use any inductive set and the axiom of comprehension to prove the existence of the set of natural numbers. $\square$

Definition 4.6. Denote by $\mathbb{N} = \omega$ the set of all natural numbers.

Corollary 4.7 (Induction principle for ω)**.** *If $T \subseteq \omega$ is inductive then $T = \omega$.*

The reason this is called the induction principle is that when we prove something by induction what we actually prove is that the set of natural numbers for which a certain statement is true is an indutive set and by the previous corollary this set must be all of ω.

Example 4.8. Prove that every natural number n is either 0 or there is $m \in \omega$ such that $n = m + 1$.

Proof. The set $\{0\} \cup \{n \in \mathbb{N} \mid \exists m \in \mathbb{N},\ m+1 = n\}$ is inductive and therefore equals $\mathbb{N}$. □

4.1.1. *The recursion theorem and the arithmetic operations.* recursion is a definition technique, but what does it define? sequences:

Definition 4.9. A *sequence* of elements of a set A is a function $f : \mathbb{N} \to A$ enumerated by the natural numbers.

Example 4.10. In calculus we denote a sequence by $(a_n)_{n=0}^{\infty}$, for example $a_n = \frac{1}{n+1}$. Formally this is just a function $a : \mathbb{N} \to \mathbb{R}$ defined by $a(n) = \frac{1}{n+1}$, and we identify between $a(n)$ and a_n.

Definition 4.11. A *recursive* definition of a sequence $h : \omega \to A$ has two parts:

(1) The initial value of the sequence: A definition for $h(0) \in A$.
(2) The recursive condition: A Function F which is used to compute the next element in the sequence $h(n+1)$ from the previous elements. Namely $F(h(n)) = h(n+1)$.

Remark 4.12. A more general form of recursion allows F to use finitely many values from A.

Example 4.13. Define $h(0) = 1$ and $h(n+1) = h(n) + 2n + 1$. Then $F : \mathbb{N} \to \mathbb{N}$ can be taken as $F(x) = x + 2n + 1$. It is not hard to prove by induction that $h(n) = (n+1)^2$

Example 4.14. Define $n!$ as follows: $0! = 1$ and $(n+1)! = (n+1) \cdot n!$. This means that $F(\langle a_1, ..., a_n \rangle) = (n+1) \cdot a_1 \cdot ... \cdot a_n$ is as wanted.

Definition 4.15 (Arithmetic operations)**.** (1) We define $n + m$ for every n by recursion on m:
$$n + 0 = n \text{ and } n + (m+1) = (n+m) + 1.$$
(2) $n \cdot m$ by recursion on m:
$$n \cdot 0 = \text{ and } n \cdot (m+1) = (n \cdot m) + n.$$
(3) n^m is defined by recursion on m:
$$n^0 = 1 \text{ and } n^{(m+1)} = (n^m) \cdot n.$$

Let us proof for example that addition is commutative:

Theorem 4.16. *For every n, m, $n + m = m + n$*

Proof. By induction on n we prove that for every m, $n + m = m + n$. For $n = 0$, we prove by induction on m that $0 + m = m + 0$. For $m = 0$, we have that $0 + 0 = 0 + 0$. Suppose this is true for m and let us prove for $m + 1$:
$$0 + (m+1) = (0+m) + 1 = (m+0) + 1 = m + 1 = (m+1) + 0$$

Suppose this is true for n and let us prove it for $n+1$. Again by induction on m we prove that $(n+1)+m = m+(n+1)$. For $m = 0$ we prove as before. Suppose this holds for m and let us prove for $m+1$

$$(n+1)+(m+1) = ((n+1)+m)+1 = (n+(m+1))+1 = ((m+1)+n)+1 = (m+1)+(n+1)$$

□

The proof of the theorem below is as above and summarizes some of the most important properties of the arithmetic operations

Theorem 4.17. *(1) Associativity:* $(n+m)+k = m+(n+k)$ *and* $(n \cdot m) \cdot k = m \cdot (n \cdot k)$.
(2) Commutativity: $n+m = m+n$ *and* $n \cdot m = m \cdot n$.
(3) Distributivity: $n \cdot (m+k) = n \cdot m + n \cdot k$.
(4) No zero divisors: $n \cdot m = 0 \Rightarrow n = 0 \vee m = 0$.

Proof. Let us prove associativity for addition. For $k = 0$ $(n+m)+0 = n+m = n+(m+0)$. Suppose this was true for k and let us prove for $k+1$

$$(n+m)+(k+1) = ((n+m)+k)+1 = (n+(m+k))+1 = n+((m+k)+1) = n+(m+(k+1))$$

Let us prove (3) by induction on k. $n \cdot (m+0) = n \cdot m$ and $n \cdot m + n \cdot 0 = n \cdot m + 0 = n \cdot m$. Suppose this was true for k and let up prove for $k+1$:

$$n \cdot (m+(k+1)) = n \cdot ((m+k)+1) = n \cdot (m+k)+n = (n \cdot m+n \cdot k)+n = n \cdot m+(n \cdot k+n) = n \cdot m+n \cdot (k+1)$$

Let us prove (1) for the multiplication by induction on k we prove the equality for every n, m. For $k = 0$ we have that $(n \cdot m) \cdot 0 = 0$ by definition of the multiplication. Also $n \cdot (m \cdot 0) = n \cdot 0 = 0$ by definition.

Suppose this was true for some k and let us prove for $k+1$

$$(n \cdot m) \cdot (k+1) = (n \cdot m) \cdot k+(n \cdot m) = n \cdot (m \cdot k)+(n \cdot m) = n \cdot (m \cdot k+m) = n \cdot (k+1))$$

Let us prove (4). Suppose that $n, m \neq 0$. Then $n = k+1$ and $m = r+1$ by previous results (See example). Then $n \cdot m = (k+1) \cdot (r+1) = (k+1) \cdot r + (k+1) = ((k+1) \cdot r + k) + 1$ which is a successor and therefore non zero. □

Problem 11. *Prove the power identities:*
(1) $n^{m+k} = n^m \cdot n^k$.
(2) $(n \cdot m)^k = n^k \cdot m^k$
(3) $(n^m)^k = n^{m \cdot k}$.

Solution 8. *By induction on k,* $n^{m+0} = n^m$ *and* $n^m \cdot n^0 = n^m \cdot 1 = n^m$. *Suppose this was true for k and let us prove for $k+1$, then*

$$n^{m+(k+1)} = n^{(m+k)+1} = n^{m+k} \cdot n = (n^m \cdot n^k) \cdot n = n^m \cdot (n^k \cdot n) = n^m \cdot (n^{k+1})$$

Theorem 4.18 (The Recursion theorem)**.** *Let $F : A \to A$ and $a \in A$. Then there is a unique function $f : \omega \to A$ such that:*
(1) $f(0) = a$.
(2) For every $n \in \omega$, $f(n+1) = F(f(n))$.

Proof. Clearly there is at most one such function f (prove it by induction!). To see the existence, consider the set T of all partial function g such that:

(1) $0 \in \mathrm{dom}(g)$ and $g(0) = a$.
(2) For every $m \in \omega$, if $m+1 \in \mathrm{dom}(g)$ then $m \in \mathrm{dom}(g)$ and moreover $g(m+1) = F(g(m))$.

We claim that $f = \cup T$ is as wanted. To see that f is total we show that $\mathrm{dom}(f) = \cup_{g \in T} \mathrm{dom}(g)$ is an inductive set. We have the function $g_0 = \{\langle 0, a\rangle\} \in T$ so $0 \in \mathrm{dom}(g_0) \subseteq \mathrm{dom}(f)$. Suppose that $n \in \mathrm{dom}(f)$ and $g \in T$ such that $n \in \mathrm{dom}(g)$. If $n+1 \in \mathrm{dom}(g)$ we are done. otherwise, define $g' = g \cup \{\langle n+1, F(g(n))\rangle\}$, it is not hard to check that also $g' \in T$ and also $n+1 \in \mathrm{dom}(g') \subseteq \mathrm{dom}(f)$, as desired.

To see that f is univalent, we prove that $\{n \mid \exists! m,\ \langle n, m\rangle \in f\}$ is inductive. For $n = 0$, if $\langle 0, m\rangle \in f$ there there is $g \in T$ such that $g(0) = m$. it follows that $m = a$ and hence there is exactly one such m. Suppose this is true for n and let us prove that $n+1$ is in the set. Assume that $\langle n+1, m_1\rangle \in g_1$ and $\langle n+1, m_2\rangle \in g_2$ for $g_1, g_2 \in T$. Then $n \in \mathrm{dom}(g_1)$, $n \in \mathrm{dom}(g_2)$ and by the induction hypothesis $g_1(n) = g_2(n)$. It follows that $m_1 = g_1(n+1) = F(g_1(n)) = F(g_2(n)) = g_2(n+1) = m_2$.

To conclude that f has the desired property prove that $f \restriction \mathrm{dom}(g) = g$ for every $g \in T$.

$\square$

Definition 4.19 (The order of the natural numbers)**.** We define $n < m$ if and only if $n \in m$.

Theorem 4.20. *(1) $<$ is a strong linear order on $\mathbb{N}$.*
(2) $<$ is well-ordered, namely, for every $\emptyset \neq X \subseteq \mathbb{N}$, there is $n \in X$ such that $n \leq m$ for every $m \in X$.
(3) For every $n, m, k \in \mathbb{N}$,
(a) $n < m$ iff $n + k < m + k$.
(b) If $k \neq 0$ then $n < m$ iff $n \cdot k < m \cdot k$.

Proof. $(1), (2)$ will be part of a more general theorem later. Let us prove $(3a)$ and leave $(3b)$ as an exercise. We prove it by induction on k. For $k = 0$ the equivalence is clear. Suppose this is true for k and let us prove for $k+1$. If $n < m$ we want to prove that $n + (k+1) < m + (k+1)$. Towards a contradiction, if $n+(k+1) \geq m+(k+1)$, then $m+k \in (n+k)+1 = n+k \cup \{n+k\}$. Then either $m+k = n+k$ which contradicts the induction hypothesis, or $m+k \in n+k$ and therefore $m+k < n+k$ which also contradicts the induction hypothesis. If $n+(k+1) < m+(k+1)$, then $n+k < m+k$, just otherwise, $n+k \geq m+k$ and therefore $(m+k)+1 \subseteq (n+k)+1$ which implies that $m+(k+1) \leq n+(k+1)$, contradiction. $\square$

Corollary 4.21 (The cancellation law)**.** *Suppose that $n + m = n' + m$ then $n = n'$.*

Proof. Since $<$ is a linear order we split into cases:

(1) If $n < n'$ then $n + m < n' + m$ contradicting the assumption.
(2) The case $n' < n$ is similar.
(3) $n = n'$ is the only possible case.

$\square$

4.2. **Defining the integers and rationals.** We have the set $\mathbb{N}$ together with $+, \cdot$ defined and has the usual properties. Recall that we have defined the following relation on $\mathbb{N}^2$: $\langle n, m\rangle \sim_Z \langle k, l\rangle$ iff $n + l = k + m$. We will suppress the Z as there is only one relation in this subsection. We already claimed previously that this is an equivalence relation. Let us prove it since this requires some gentile properties of addition:

Proposition 4.22. *$\sim$ is an equivalent relation.*

Proof. reflexivity and symmetry are easy to verify. and only requires that $n + m = m + n$. As for transitivity, suppose that $n + l = k + m$ and $k + r = s + l$ we wand to prove that $n + r = s + m$ Indeed, $n + r + l = k+m+r = s+l+m = s+m+l$. By the cancellation law, $n+r = s+m$. $\square$

We think of $[< n, m >]_\sim = n - m$.

Definition 4.23. $\mathbb{Z} = \mathbb{N}^2 / \sim$.

We identify $\mathbb{N}$ inside $\mathbb{Z}$ by $n \mapsto [\langle n, 0\rangle]_\sim$. Also we denote by $-n = [\langle 0, n\rangle]_\sim$ and more generally $-[\langle n, m\rangle]_\sim = [\langle m, n\rangle]_\sim$.

Definition 4.24. We define $[\langle n, m\rangle]_\sim + [\langle n', m'\rangle]_\sim = [\langle n + n', m + m'\rangle]_\sim$.

We already proved this operation does not depend on the choice of representatives. Also we define $z_1 - z_2$ as $z_1 + (-z_2)$.

Exercise 9. *Define properly multiplication (think of $(n-m)\cdot(n'-m')$ and prove it does not depend on the choice of representatives.*

Definition 4.25. We define the order by $[\langle n, m\rangle] < [\langle n', m'\rangle]$ iff $n + m' < n' + m$

Proposition 4.26. *The usual properties of addition/multiplication and the order on the integers hold. In particular we have commutativity of addition and multiplication, and the following cancellations role: for every a, b and every $c \neq 0$,*

$$a \cdot c = b \cdot c \Rightarrow a = b$$

Remark 4.27. In fact $\mathbb{Z}$ with addition and multiplication is an integral domain: a ring with no zero divisors.

We can repeat the same construction we had from $\mathbb{N}$ to $\mathbb{Z}$ in order to construct $\mathbb{Q}$ from $\mathbb{Z}$ replacing addition with multiplication:

Definition 4.28. Define an equivalence relation $\sim$ on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ by

$$\langle z_1, z_2\rangle \sim \langle t_1, t_2\rangle \text{ iff } z_1 t_2 = z_2 t_1$$

Again, this is an equivalence relation due to the properties of multiplication on $\mathbb{Z}$.

Definition 4.29. $\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z} \setminus \{0\})/\sim$

Definition 4.30. $[\langle z_1, z_2\rangle]_\sim + [\langle t_1, t_2\rangle]_\sim = [\langle z_1t_2 + t_1z_2, z_2t_2\rangle]_\sim$
$[\langle z_1, z_2\rangle]_\sim \cdot [\langle t_1, t_2\rangle]_\sim = [\langle z_1t_1, z_2t_2\rangle]_\sim$
if $s, t \neq 0$ define $([\langle s, t\rangle]_\sim)^{-1} = [\langle t, s\rangle]_\sim$.

We think of $[\langle t, s\rangle]_\sim$ as $\frac{t}{s}$ and identify $z \mapsto [\langle z, 1\rangle]_\sim$.

Problem 12. *Prove that for every $[\langle n, m\rangle]_\sim \in \mathbb{Q}$ there is $n', m' \in \mathbb{Z}$ such that $m' > 0$ and $[\langle n, m\rangle]_\sim = [\langle n', m'\rangle]_\sim$.*

Definition 4.31. Suppose that $n, n' \in \mathbb{Z}$ and $m, m' \in \mathbb{N}_+$. Define $[\langle n, m\rangle]_\sim < [\langle n', m'\rangle]_\sim$ iff $nm' < mn'$.

Again one should check that all the regular properties of the operations and order

Proposition 4.32. *$\mathbb{Q}$ as no least and last element and it is dense in itself.*

Proof. To see there is no last element (a similar proof shows that there is no least element), let $[\langle n, m\rangle]_\sim \in \mathbb{Q}$ be any rational number, and assume without loss of generality that $m > 0$. Then $nm < (n+1)m$ by properties of $<$ on $\mathbb{Z}$. By definition of $<$ on $\mathbb{Q}$ it follows that $[\langle n, m\rangle]_\sim < [\langle n+1, m\rangle]_\sim$.

To see that $\mathbb{Q}$ is dense in itself, let $[\langle n_1, m_1\rangle]_\sim, [\langle n_2, m_2\rangle]_\sim \in \mathbb{Q}$ be such that $q_1 := [\langle n_1, m_1\rangle]_\sim < [\langle n_2, m_2\rangle]_\sim =: q_2$ and $m_1, m_2 > 0$ (We calculate the average $\frac{\frac{n_1}{m_1} + \frac{n_2}{m_2}}{2} = \frac{n_1m_2 + n_2m_1}{2m_1m_2}$). Define $q_3 := [\langle n_1m_2 + n_2m_1, 2m_1m_2\rangle]_\sim$ and let us prove that $q_1 < q_3 < q_2$. Indeed, $q_1 < q_3$ since

$$n_1(2m_1m_2) = n_1m_1m_2 + n_1m_1m_2 \underset{\text{Since } n_1m_2 < n_2m_1}{<} n_1m_1m_2 + m_1n_2m_1$$
$$= m_1(n_1m_2 + n_2m_1)$$

Also to see that $q_3 < q_2$,

$$m_2(n_1m_2+n_2m_1) = m_2n_1m_2+m_2n_2m_1 \underset{\text{Since } n_1m_2 < n_2m_1}{<} m_2n_2m_1+n_2m_1m_2 = 2n_2m_1m_2$$

□

Theorem 4.33. *$\mathbb{Q}$ is countable i.e. there is a bijection between $\mathbb{N}$ and $\mathbb{Q}$.*

We will prove this theorem later.

Theorem 4.34 (Cantor)**.** *If $\langle A, \leq_A \rangle$ is a countable ordered set with no least and last element which is dense in itself the $\langle A, \leq_A \rangle \simeq \langle \mathbb{Q}, \leq \rangle$.*

Proof. Suppose that $\mathbb{Q} = \{q_n \mid n \in \mathbb{N}\}$ is an enumeration of $\mathbb{Q}$ an $A = \{a_n \mid n \in \mathbb{N}\}$ is an enumeration of A. We construct the isomorphism $f : \mathbb{Q} \to A$ by induction. We start with q_0 and define $f(q_0) = a_0$. Let us do q_1 for clarity reasons. We split into cases

(1) If $q_0 < q_1$, then pick a_m for the minimal m such that $a_0 < a_m$. Note that such an m exists since A has no last element. Define $f(q_1) = a_m$.
(2) If $q_1 < q_0$ we choose a_m for the minimal m such that $a_m < a_0$ and define $f(q_1) = a_m$.

Now before taking care of q_2, we make sure we took care of a_1, if $a_1 = a_m$ we are done. Otherwise, we take a_1 and split into cases:

(1) If $a_1 < a_m, a_0$, then we choose q_k for the minimal k such that $q_k < q_0, q_1$ and define $f^{-1}(a_1) = q_k$ or equivalently, we define $f(q_k) = a_1$.
(2) If $a_1 > a_m, a_0$ we act similarly.
(3) If $\min\{a_0, a_m\} < a_1 < \max\{a_0, a_m\}$ (namely $a_m < a_1 < a_0$ or $a_0 < a_1 < a_m$), then we choose q_k, for the minimal k such that $\min\{q_0, q_1\} < q_k < \max\{q_0, q_1\}$ and define $f^-1(a_1) = q_k$.

In general we assume that at the n^{th} step f is defined on $\{q_{1_i}, ..., q_{i_N}\}$ such that $\{0, ..., n\} \subseteq \{i_1, ..., i_N\}$ and $q_{i_1} < ... < q_{i_N}$. Moreover, we assume that $\{a_0, ..., a_n\} \subseteq \{f(q_{i_1}), ..., f(q_{i_N})\}$ and $f(q_{i_1}) < .... < f(q_{i_N})$. If $n + 1 \in \{i_1, ..., i_N\}$ we do nothing at the $\mathbb{Q}$ side. Otherwise, we split into cases:

(1) If $q_{n+1} < q_{i_1}$, we choose m such that $a_m < f(q_{i_1})$ (which exists since there is no least element in A) and define $f(q_{i_{n+1}}) = a_m$.
(2) If $q_{n+1} > q_{i_N}$, we act similarly.
(3) Otherwise, there is a unique $1 \leq r < N$ such that $q_{i_r} < q_{n+1} < q_{i_{r+1}}$, we choose $f(q_{i_r}) < a_m < f(q_{i_{r+1}})$ which exists since A is dense in itself and define $f(q_{n+1}) = a_m$.

If $a_{n+1} \in \{f(q_{i_1}), ..., f(q_{i_N}), a_m\}$, then we do nothing on the A side. Otherwise, we again split into the same cases according to the interval that a_{n+1} fall in and choose q_k in the corresponding interval in the $\mathbb{Q}$-side, then we define $f^{-1}(a_{n+1}) = q_k$. The function f which we define is clearly order-preserving, as we made sure to choose the images and preimages in the correct interval. It is one-to-one since we always choose different elements and it is onto as for each n, at the n^{th} stage we ensured that $\{q_0, ...q_n\} \subseteq \text{dom}(f)$ and $\{a_0, ..., a_n\} \subseteq \text{im}(f)$. $\square$

4.3. **The real numbers.** The previous construction applied to $\mathbb{Q}$ will in fact result in $\mathbb{Q}$ and will not add anything new. How do we construct the reals? what is the missing property of $\mathbb{Q}$ that defers it from the reals? This is not an algebraic property but rather a topological/order-theoretic property:

Definition 4.35. Let $\langle A, <_A \rangle$ be a linearly ordered set. A set $X \subseteq A$ is called:

(1) Bounded from above (below) is there is $a \in A$ such that for every $x \in X$, $x \leq_A a$ ($x \geq_A a$). a is called an upper (lower) bound
(2) Bounded if it is both bounded from above and below.

Definition 4.36. A least upper bound (last lower bound) of a set $X \subseteq A$ is an element $a \in X$ which is an upper (lower) bound and for any upper (lower) bound $b \in A$, $a \leq_A b$ ($a \geq_A b$).

Example 4.37. 2.5 is an upper bound for the open interval $(0, 2) \subseteq \mathbb{R}$. However 2 is the least upper bound for that set.

Example 4.38. For $\mathbb{Q}$, consider $X = \{q \in \mathbb{Q} \mid q^2 < 2\}$. Then X is bounded by 2 for example but there is no least upper bound of X in $\mathbb{Q}$. To see this, suppose otherwise, and let q^* be such a least upper bound. Since $q^* = \frac{m}{n} \in \mathbb{Q}$, it is impossible that $(q^*)^2 = 2$. If $(q^*)^2 < 2$, consider $q' = \frac{2q^*+2}{q^*+2}$. Then one can check that:

(1) $q^* < q'$.
(2) $(q')^2 < 2$.

Contradicting that q^* is an upper bound.

If $q^* > 2$, then again we define the same $q' = \frac{2q^*+2}{q^*+2}$. This time we have:

(1) $q^* > q'$.
(2) $(q')^2 > 2$

From (2) it follows that whenever $q^2 < 2$ then $q^2 < (q')^2$ and therefore $q < q'$. Hence q' is an upper bound for X contradicting that q^* is the least upper bound.

Definition 4.39. A linearly ordered set A is called complete if every bounded non-empty set has a least upper bound.

The completeness property of the reals is what enables taking limits in calculus and is in fact equivalent to many known theorems from calculus (for example that every Cauchy sequence converges).

There is a general method to "complete" an order i.e.adding those missing points. We will introduce the construction only for $\mathbb{Q}$ and refer the reader to the literature for the general construction. This idea is due to Dedikind and was used by him in a very similar way in his famous prime ideal decomposition theorem (Dedekind-Kummer theorem).

Definition 4.40. A set $B \subseteq \mathbb{Q}$ is called a Dedekind cut if:

(1) $B \neq \emptyset$.
(2) B is bounded from above i.e. there is $q \in \mathbb{Q}$ such that for every $b \in B$, $b < q$.
(3) B is downward closed i.e. if whenever $b \in B$ and $q \in \mathbb{Q}$, is such that $q < b$, then $q \in B$.
(4) B has no last element i.e. for every $q \in B$ there is $p \in B$ such that $q < p$.

Definition 4.41.

$$\mathbb{R} := \{X \in P(\mathbb{Q}) \mid X \text{ is a Dedekind cut}\}$$

Definition 4.42. The order of $\mathbb{R}$ is $r < s$ iff $r \subseteq s$.

$<$ is a linear ordering of $\mathbb{R}$.

Proof. The fact that it is a strong order is easy. Let us check that the order is linear. Let $r_1, r_2 \in \mathbb{R}$. Suppose that $r_1 \neq r_2$ and let us split into cases:

(1) If there is $q \in r_1 \setminus r_2$, then for every $p \in r_2$, we must have that $p < q$, just otherwise, $q \leq p$ and then $q \in r_2$ since $p \in r_2$ and r_2 is downward closed, which is a contradiction. Hence $p \in r_1$ since r_1 is downward closed and $q \in r_1$. We conclude that $r_2 \subsetneq r_1$.
(2) The case where there is $q \in r_2 \setminus r_2$ is symmetric.

□

There is a standard way to identify $\mathbb{Q}$ inside $\mathbb{R}$, by $q \mapsto \mathbb{Q}_<[q] := \{p \in \mathbb{R} \mid p < q\}$.

Problem 13. *This function is an embedding of $\mathbb{Q}$ in $\mathbb{R}$.*

Example 4.43. The set $X = \{q \in \mathbb{Q} \mid q < 0 \vee q^2 < 2\}$ is a Dedekind cut and there is no $q \in \mathbb{Q}$ such that $X = \mathbb{Q}_<[q]$

Theorem 4.44. *$\mathbb{Q}$ is dense in $\mathbb{R}$*

Proof. If $X_1 < X_2$ are any cuts, fix any $q \in X_2 \setminus X_1$ then $X_1 \leq q < X_2$. Since X_2 has no maximal element, there is $q' \in X_2$ such that $q < q'$, then clearly, $X_1 < q' < X_2$. □

Theorem 4.45. *$\mathbb{R}$ is complete*

Proof. Let $\mathcal{F} \subseteq \mathbb{R}$ be a non empty bounded set of reals, then $\bigcup \mathcal{F} \in \mathbb{R}$ is a Dedekind cut which is the supremum of $\mathcal{F}$. □

Theorem 4.46. *$\mathbb{R}$ is the unique (up to isomorphism) ordered $\langle A, R\rangle$ set such that:*

(1) $\langle A, R\rangle$ has no first and last element.
(2) $\langle A, R\rangle$ contains a countable dense subset. (separability)
(3) $\langle A, R\rangle$ is complete.

Lemma 4.47. *For every $r \in \mathbb{R}$, $r = \sup \mathbb{Q} \cap (-\infty, r)$.*

Proof. Clearly r is an upper bound for $\mathbb{Q} \cap (-\infty, r)$ and therefore $\sup \mathbb{Q} \cap (-\infty, r) \leq r$. If $r' < r$, then since $\mathbb{Q}$ is dense in $\mathbb{R}$, there is $q \in \mathbb{Q}$ such that $r' < q < r$ and therefore r' is not a bound for $\mathbb{Q} \cap (-\infty, r)$. □

Proof of Theorem. Let $\langle S, \leq_S \rangle$ be an ordered set with a countable dense subset A with no least and last element. By Cantor's Theorem $A \simeq \mathbb{Q}$ and let $f : \mathbb{Q} \to A$ be the witnessing isomorphism. Define $F : \mathbb{R} \to S$ by $F(r) = \sup f''\mathbb{Q} \cap (-\infty, r)$. This is a well-defined function since S is complete. We leave to the reader to check that this is indeed an isomorphism. □

Definition 4.48. We define the operations on $\mathbb{R}$ as follows:

$$A + B = \{a + b \mid a \in A, b \in B\}$$

Proposition 4.49. *If* $r_1, r_2 \in \mathbb{R}$ *then* $r_1 + r_2 \in \mathbb{R}$

Proof. Let us check the three properties of a Dedekind cut:

(1) Since $r_1, r_2 \neq \emptyset$, there are $q_i \in r_i$, $i = 1, 2$ and thus $q_1 + q_2 \in r_1 + r_2$ which implies that $r_1 + r_2 \neq \emptyset$.
(2) Let q_i be an upper bound for r_i for $i = 1, 2$. Then for every $p_1 + p_2 \in r_1 + r_2$, where $p_i \in r_i$, then $p_1 \leq q_1$ an $p_2 \leq q_2$. Hence, we have that
$$p_1 + p_2 \leq q_1 + q_2.$$
It follows that $q_1 + q_2$ is an upper bound.
(3) Let us prove that $r_1 + r_2$ is downward closed. Let $q < q_1 + q_2 \in r_1 + r_2$. Then $q - q_1 < q_2 \in r_2$ and therefore $q - q_1 \in r_2$ as r_2 is downward closed. It follows that $q = q_1 + (r - q_1) \in r_1 + r_2$ as wanted.
(4) Let us prove that $r_1 + r_2$ has no last element. Let $q_1 + q_2 \in r_1 + r_2$. Then there is $p_1 \in r_1$ and $p_2 \in r_2$ such that $q_1 < p_1$ and q_2. It follows that $q_1 + q_2 < p_1 + p_2 \in r_1 + r_2$ as wanted.

□

Let us give another example.

Proposition 4.50. *For every real number* r, $r + 0 = r$

Proof. Let us prove double inclusion. Let $q + p \in r + 0$. Then $p < 0$ and therefore $q + p < q$. Since r is downward closed, $q + p \in r$. Let $q \in r$. Since r has no last element, there is $p \in r$ such that $q < p$. It follows that $q - p < 0$ and therefore $q = p + (q - p) \in r + 0$. □

Definition 4.51. For $x \in \mathbb{R}$ we define:
$$-x = \{q \in \mathbb{Q} \mid \exists s > q, \ -s \notin x\}.$$

Problem 14. *Prove that* $-x \in \mathbb{R}$.

Proposition 4.52. $x + (-x) = 0$.

Proof. Let us prove a double inclusion.

(1) Let $q + p \in x + (-x)$, then there is $s > p$ such that $-s \notin x$. In particular, since $q \in x$, $q < -s$. We conclude that $q + p < q + s < -s + s = 0$ (the last equality is equality of rationals).
(2) Let $p < 0$, then $-\frac{p}{2} > 0$. Let $t \in x$ be any element. Find $n \in \mathbb{N}_+$ be the least such that $t + n \cdot (-\frac{p}{2}) \notin x$ (such an n exists since $x \neq \mathbb{Q}$) Let $q = t + (n - 1) \cdot (-\frac{p}{2})$. So $q \in x$ and if we let $s = \frac{p}{2} - q$, then $-s = q - \frac{p}{2} \notin x$. Moreover, $p - q < \frac{p}{2} - q = s$. By definition $p - q \in -x$. We conclude that
$$p = q + (p - q) \in x + (-x).$$

□

Definition 4.53. $|x| = x \cup -x$.

Definition 4.54. Define $x \cdot y$ as follows:

(1) If $x, y \geq 0$, define $x \cdot y = 0 \cup \{p \cdot q \mid p \in x, q \in y \text{ and } p, q \geq 0\}$.
(2) If $x, y < 0$, define $x \cdot y = |x| \cdot |y|$.
(3) If $x < 0 \leq y$ or $y < 0 \leq x$ then $x \cdot y = -(|x| \cdot |y|)$.

Theorem 4.55. *Let $\sqrt{2} = \{q \in \mathbb{Q} \mid q < 0 \vee q^2 < 2\}$. Then $(\sqrt{2})^2 = 2$*

Proof.
$$\sqrt{2} \cdot \sqrt{2} = 0 \cup \{p \cdot q \mid p, q \in \sqrt{2} \text{ and } p, q \geq 0\}$$
Let $p, q \geq 0$ be such that $p, q \in \sqrt{2}$. Without loss of generality suppose that $p \leq q$. Hence $p \cdot q \leq p \cdot p < 2$ hence $p \cdot q \in 2$. For the other direction, let $p < 2$. If $p \leq 0$ then clearly $p \in \sqrt{2} \cdot \sqrt{2}$. So suppose that $p > 0$

Lemma 4.56. *There is N such that for every $k \geq N$, there is $m \in \mathbb{N}$ such that $kp < m^2 < 2k$*

Proof of Lemma. Find $N_0 \in \mathbb{N}$ so that $N_0 p \geq 1$ (N_0 can be any number which is at least the denominator of $|p|$) For every $k \geq N_0$, we find n_k such that $n_k^2 \leq kp < (n_k + 1)^2$. Note that the sequence n_k is weakly monotone with k and goes to infinity with k. Note that:
$$\frac{(n_k + 1)^2}{kp} - 1 \leq \frac{2n_k + 1}{n_k^2} \to_{k \to \infty} 0$$
Hence there is $N \geq N_0$ such that for every $k \geq N$
$$\frac{(n_k + 1)^2}{kp} - 1 < \frac{2 - p}{p}$$
For every such k we have that $kp < (n_k + 1)^2 < 2k$ as wanted.

□

To conclude the theorem we find any n such that $n^2 \geq N$ and then $n^2 p < m^2 < n^2 2$ for some m. Hence $p < \frac{m}{n}^2 < 2$. Note that $\frac{m}{n} \in \sqrt{2}$ and therefore $p < \frac{m}{n}^2 \in \sqrt{2}\sqrt{2}$. By downward closure, $p \in \sqrt{2}\sqrt{2}$. □

A decimal representation of a real number r is an integer number n and a sequence $(a_k)_{k=1}^{\infty}$ such that $a_k \in \{0, ..., 9\}$ and $n.a_1 a_2 ... a_N := n + \sum_{k=1}^{N} \frac{a_k}{10^k} \to_{N \to \infty} r$. We denote that by $n.a_1 a_2 a_3 ....$ Note that a real number can have two representations:

Example 4.57. 1.99999.... = 2.00000... Since the constant function $2 = 2 + \frac{0}{10} + \frac{0}{100}...$ converges to 2 but also, by the standard formula for the sum of a geometric series,
$$1 + \sum_{k=1}^{N} \frac{9}{10^k} = 1 + \frac{9}{10}\left(\frac{1 - \frac{1}{10^N}}{1 - \frac{1}{10}}\right) \to_{N \to \infty} 1 + \frac{9}{10} \frac{1}{1 - \frac{1}{10}} = 1 + 1 = 2$$

The next theorem shows that avoiding representations ending with infinitely many zeros results in a unique representation.

Theorem 4.58. *Let r be a real number. Then there is a unique integer n and sequence $(a_k)_{k=1}^{\infty}$ such that $r = n.a_1 a_2 a_3 ...$*

Proof. Existence: n is defined to be the maximal integer such that $n < r$. in particular $n < r \leq n+1$ and so $0 < r-n \leq 1$. There is an $a_1 \in \{0, ..., 9\}$ such that $\frac{a_1}{10} < r-n \leq \frac{a_1+1}{10}$ (equivalent to $n+\frac{a_1}{10} < r$) hence $0 < r-(n.a_1) \leq \frac{1}{10}$. Suppose we have defined $a_1, ..., a_k \in \{0, ..., 9\}$ such that $0 < r-(n.a_1....a_k) \leq \frac{1}{10^k}$. Let $a_{k+1} \in \{0, ..., k\}$ be such that $\frac{a_k}{10^{k+1}} < r-(n.a_1....a_k) \leq \frac{a_k+1}{10^{k+1}}$. It follows that $0 < r-(n.a_1....a_k a_{k+1}) <\leq \frac{1}{10^{k+1}}$. Since $\frac{1}{10^k} \to_{k\to\infty} 0$, it follows that $n.a_1a_2a_3... = r$. Note that the sequence a_k cannot be eventually zero since this would mean that for some N $r - n.a_1a_2...a_N = 0$ contradicting the choice of a_N so that this difference is greater than 0. For uniqueness, suppose that $a_0.a_1a_2... = r = b_0.b_1b_2....$ Suppose a contradiction that there is k such that $a_k \neq b_k$ and let k be minimal. Without loss of generality assume that $a_k < b_k$ Let us split into cases:

(1) If $k = 0$, we find any $M > 0$ such that $b_M > 0$. Such an M exists by our assumption that the sequence is not eventually 0. Note that for every $m \geq M$,
$$b_0.b_1b_2....b_m \geq b_0.b_1b_2....1 > b_0 \geq a_0.a_1a_2...a_m$$
Since this strong inequality holds for a tail of m's, the limits cannot be the same, contradiction.

(2) If $k > 0$, we find any $M > 0$ such that $b_M > 0$. Note that for every $m \geq M$,
$$b_0.b_1b_2....b_m \geq b_0.b_1b_2....b_M > b_0.b_1...b_k \geq a_0.a_1...(a_k+1) \geq a_0.a_1a_2...a_m$$
contradicting the limit equality.

□

Theorem 4.59. $\mathbb{R}$ *is not countable*

Cantor's original proof. Suppose otherwise, then $\mathbb{R} = \{r_n \mid n \in \mathbb{N}\}$. Let us define a sequence:
$$a_0 < a_1 < a_2...a_n < ... < b_n < b_{n-1}... < b_2 < b_1 < b_0$$
as follows: $a_0 = r_0$ and $b_0 = r_k$ for the minimal k such that $r_k > a_0$. Suppose that $a_n < b_n$ were defined and let $a_{n+1} = r_k$ for the minimal k such that $a_n < r_k < b_n$. Let $b_{n+1} = r_k$ for the minimal k such that $a_{n+1} < r_k < b_n$. By the completeness of $\mathbb{R}$ there is $a = \sup_{n<\omega} a_n$. Note that for every n, $a_n < a < b_n$. There is k^* such that $a = r_{k^*}$ and there if $l > k^*$ such that for some n, $b_n = r_l$. This means that at stage $n-1$, we had $a_{n-1} < b_{n-1}$ and we chose $b_n = r_k$ for the minimal k such that $a_{n-1} < r_k < b_{n-1}$ and this minimal k was l. However, $a_{n-1} < a = r_{k^*} < b_{n-1}$ also satisfies this property and $k^* < l$, contradiction. □

4.4. **Two questions about the real numbers.**

Question 1. *Can the real numbers be well-ordered?*

Definition 4.60. An ordered set $\langle A, R\rangle$ is called c.c.c (countable chain condition) if whenever I is a set of disjoint open intervals in A, then I is at most countable.

Problem 15. $\mathbb{R}$ *is c.c.c.*

Question 2. *Suslin hypothesis: If we replace separability by c.c.c do we still obtain a characterization of* $\mathbb{R}$

5. Equinumerability

Definition 5.1. Let A, B be any sets. We say that:

(1) $A \approx B$ "A and B are equinumerable" if there is a bijection $f : A \to B$.
(2) $A \prec B$ "A is at most the size of B" if there is an injective function $f : A \to B$.
(3) $A \not\approx B$ if $\neg(A \approx B)$, namely if there is no bijection $f : A \to B$.
(4) $A \prec B$ if $A \preceq B$ and $A \not\approx B$.

Example 5.2. (1) $\{1, 2, 3\} \approx \{2, 7, 19\}$ as witnessed by the bijection

$$f(x) = \begin{cases} 2 & x = 1 \\ 7 & x = 2 \\ 19 & x = 3 \end{cases}$$

(2) $\mathbb{N} \approx \mathbb{N}_{even}$ as witnessed by the function $f : \mathbb{N} \to \mathbb{N}_{even}$, $f(n) = 2n$.
(3) $A \preceq P(A)$ for every set A as witnessed by the function $f : A \to P(A)$, $f(a) = \{a\}$.
(4) $(0, 1) \simeq (1, 3)$ as given by $f : (0, 1) \to (1, 3)$, $f(x) = 2x + 1$.
(5) $\{X \in P(\mathbb{N}) \mid 0 \in X\} \approx P(\mathbb{N})$ by $f : P(\mathbb{N}) \to \{X \in P(\mathbb{N}) \mid 0 \in X\}$, $f(X) = \{0\} \cup \{x + 1 \mid x \in X\}$.
(6) $\mathbb{N} \times \mathbb{N} \preceq P(\mathbb{N})$ witnessed by $f : \mathbb{N} \times \mathbb{N} \to P(\mathbb{N})$, $f(\langle n, m\rangle) = \{n, n + m\}$.
(7) $A \subseteq B \to A \preceq B$ as witnessed by the function $f : A \to B$, $f(a) = a$.
(8) Clearly $A \approx B$ implies $A \preceq B$.

Claim 5.2.1. *for any sets* A, B, C*:*

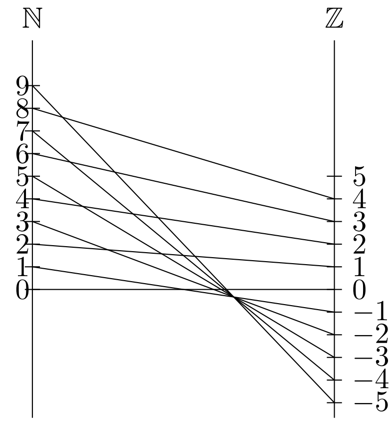
(1) $A \approx A$.
(2) $A \approx B \to B \approx A$.
(3) $A \approx B \wedge B \approx C \to A \approx C$ *and* $A \preceq B \preceq C \to A \preceq C$.

Are there two infinite sets which are not equinumerable?

Proposition 5.3. $\mathbb{N} \simeq \mathbb{Z}$

Proof. Define $f : \mathbb{N} \to \mathbb{Z}$ by

$$f(n) = \begin{cases} \frac{n}{2} & n \in \mathbb{N}_{even} \\ -\frac{n+1}{2} & n \in \mathbb{N}_{odd} \end{cases}$$



□

$\mathbb{Z}$ is like "two copies" of $\mathbb{N}$. What about infinitely many copies of $\mathbb{N}$? $\mathbb{N} \times \mathbb{N}$.

Proposition 5.4. $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$

Proof. Define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $f(\langle n, m\rangle) = 2^n(2m+1) - 1$. □

We will have an easier proof later.

Proposition 5.5. *Let A, A', B, B' be sets such that $A \approx A'$ and $B \approx B'$. Then:*

(1) $P(A) \approx P(A')$.
(2) $A \times B \approx A' \times B'$.
(3) ${}^{B}A \approx {}^{B'}A'$.
(4) If A, B are disjoint and A', B' are disjoint then $A \uplus B \approx A' \uplus B'$.

The above proposition is true upon replacing $\approx$ by $\preceq$ everywhere.

Proof. Let us prove for example (1). Let $f : A \to A'$ be a bijection. One should check that $F : P(A) \to P(A')$ defined by $F(X) = f''X$ is a bijection. □

Example 5.6. $\mathbb{N} \simeq \mathbb{Z} \times \mathbb{Z}$.

What about $\mathbb{Q}$? clearly $\mathbb{N} \preceq \mathbb{Q}$

Claim 5.6.1. (AC) *Suppose that $A \neq \emptyset$. Then $A \preceq B$ iff there is $f : B \to A$ onto.*

Proof. Suppose that $g : A \to B$ is one-to-one. Let us $a^* \in A$ be some elements. Define $f : B \to A$ by

$$f(b) = \begin{cases} a^* & b \notin Im(g) \\ g^{-1}(b) & b \in Im(g) \end{cases}$$

This is well defined since g is invertible on its image. For the other direction, suppose that $f : B \to A$ is onto. Let us define $g : A \to B$ one-to-one. For every $a \in A$, since f is onto, there is some (choose!) $b_a \in f^{-1}[\{a\}]$. Define $g(a) = b_a$. Then g is one to one since if $a \neq a'$ then $b_a \in f^{-1}[\{a\}]$ and $b_{a'} \in f^{-1}[\{a'\}]$ which are disjoint sets and therefore $b_a \neq b_{a'}$. Hence g is one-to-one. $\square$

Example 5.7. $\mathbb{Q} \preceq \mathbb{Z} \times \mathbb{Z} \approx \mathbb{N}$. The function $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ defined by

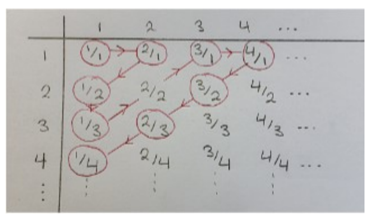
$$f(\langle z_1, z_2 \rangle) = \begin{cases} \frac{z_1}{z_2} & z_2 \neq 0 \\ 0 & else \end{cases}$$

is onto

So we are in the situation where $\mathbb{N} \preceq \mathbb{Q}$ and $\mathbb{Q} \preceq \mathbb{N}$. Does it mean that $\mathbb{N} \approx \mathbb{Q}$? Yes! but this requires a highly non-trivial theorem which we will prove later. Instead, let us give direct proof:

Theorem 5.8. $\mathbb{N} \approx \mathbb{Q}$

Proof. We are about to construct a function $f : \mathbb{N}_+ \to \mathbb{Q}_+ = \{q \in \mathbb{Q} \mid q > 0\}$ one-to-one and onto, by recursion on $\mathbb{N}_+$. To do so, we think of the $\mathbb{Q}_+$ as elements in the matrix $\mathbb{N}_+ \times \mathbb{N}_+$



We go by induction on the diagonal rows (namely pair $\langle k_1, k_2 \rangle$ such that $k_1 + k_2 = n$ starting at $n - 2$). We define $f(1) = 1/1 = 1$. Suppose we reached the n^{th} row. In row $n + 1$, we keep defining f on new (finitely many) values only for those pairs which represent a rational number which haven't appeared before (to ensure the function is one-to-one). The resulting function f is a bijection from $\mathbb{N}_+$ to $\mathbb{Q}_+$. Let us now define a function $g : \mathbb{N} \to \mathbb{Q}$ by

$$g(n) = \begin{cases} 0 & n = 0 \\ f(\frac{n}{2}) & n \in \mathbb{N}_{even} \setminus \{0\} \\ -f(\frac{n+1}{2}) & n \in \mathbb{N}_{odd} \end{cases}$$

$\square$

So far we failed to find two infinite sets which are not equinumerable.

Theorem 5.9. *(AC) If A is infinite then* $\mathbb{N} \prec A$.

Proof. We construct the function $f : \mathbb{N} \to A$ by recursion, there is always a possibility to continue the definition of f and pick a new element since otherwise, A was finite. $\square$

Definition 5.10. A set A is countable if $A \approx \mathbb{N}$. A is uncountable if $\mathbb{N} \prec A$.

Theorem 5.11. *The following sets are countable:* $\mathbb{Z}, \mathbb{N}_{even}, \mathbb{Q}, \mathbb{N}\times\mathbb{N}, \mathbb{N}^n (n \geq 1)$

Proof. It remains to show that $\mathbb{N}^n$ is countable. We prove that by induction on n. For $n = 1$, this is clear. Suppose that $\mathbb{N}^n \approx \mathbb{N}$, then

$$\mathbb{N}^{n+1} \approx \mathbb{N}^n \times \mathbb{N} \approx \mathbb{N} \times \mathbb{N} \approx \mathbb{N}.$$

$\square$

Theorem 5.12 (Cantor's Diagonalization Theorem)**.** $\mathbb{N} \prec {}^{\mathbb{N}}\{0,1\}$

Proof. It is not hard to prove that $\mathbb{N} \preceq {}^{\mathbb{N}}\{0,1\}$. So it remains to prove that $\mathbb{N} \not\approx {}^{\mathbb{N}}\{0,1\}$. Assume toward a contradiction that $F : \mathbb{N} \to {}^{\mathbb{N}}\{0,1\}$ was onto. Let us show how to produce a function $g : \mathbb{N} \to \{0,1\}$ (i.e. an element in the range of F) such that for every n, $F(n) \neq g$ (i.e. g is not in the image of F). This will produce a contradiction to the assumption that F is onto.

For each n, $F(n) : \mathbb{N} \to \{0,1\}$ so we write it as a binari sequence

$$f_n := F(n) = \langle F(n)(0), F(n)(1), F(n)(2), ...\rangle$$

So the list of functions $F(0), F(1), F(2)$ can be written in a matrix:

$$\begin{array}{ccccccc}
\underline{f_0(0)} & f_0(1) & f_0(2) & f_0(3) & \dots & f_0(n) & \dots \\
f_1(0) & \underline{f_1(1)} & f_1(2) & f_1(3) & \dots & f_1(n) & \dots \\
f_2(0) & f_2(1) & \underline{f_2(2)} & f_2(3) & \dots & f_2(n) & \dots \\
f_3(0) & f_3(1) & f_3(2) & \underline{f_3(3)} & \dots & f_3(n) & \dots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \dots & \ddots \\
f_n(0) & f_n(1) & f_n(2) & f_n(3) & \dots & \underline{f_n(n)} & \dots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots
\end{array}$$

Note that each value in this matrix is 0 or 1. We would like to define a function $g : \mathbb{N} \to \{0,1\}$, namely a binary sequence $\langle g(0), g(1), g(2), ...\rangle$ such that g defers from each row at some n. so we change the values from 0 to 1, Start by setting $g(0) = 0$ if $f_0(0) = 1$ or $g(0) = 1$ if $f_0(0) = 0$ ("flip the bit") algebraically we can write that as $1 - f_0(0)$. Moving to f_1, we flip the value $f_1(1)$ and define $g(1) = 1 - f_1(1)$. In general, we flip the values on the diagonal and define $g(n) = 1 - f_n(n)$. To that g is as wanted, suppose toward a contradiction that $g = f_n$ for some n, then by function equality we get that $1 - f_n(n) = g(n) = f_n(n)$ hence $f_n(n) = \frac{1}{2}$, contradiction. $\square$

Corollary 5.13. *For every set* A*,* $A \prec {}^{A}\{0,1\}$.

Proof. If $A = \emptyset$ this is straightforward. So assume $A \neq \emptyset$. Toward a contradiction, suppose that $F : A \to {}^{A}\{0,1\}$ is onto and denote by $f_a = F(a)$. Define $g : A \to \{0,1\}$ by

$$g(a) = 1 - f_a(a)$$

The continuation is as before. $\square$

Theorem 5.14. $P(A) \approx {}^{A}\{0,1\}$

Proof. For a subset $B \subseteq A$ we define the indicator function $\chi^A_B : A \to \{0,1\}$ by

$$\chi^A_B(a) = \begin{cases} 1 & a \in B \\ 0 & a \notin B \end{cases}$$

The function $\chi^A : P(A) \to {}^A\{0,1\}$ defined by $\chi^A(B) = \chi^A_B$ is a bijection (prove that!). □

Theorem 5.15 (Cantor's Theorem)**.** $A \prec P(A)$

Proof. $a \mapsto \{a\}$ is an injection from A to $P(A)$ hence $A \preceq P(A)$. Suppose toward a contradiction that $A \approx P(A)$, then by the previous theorem $A \approx {}^A\{0,1\}$, contradiction. □

Corollary 5.16. $\mathbb{N} \prec P(\mathbb{N}) \prec P(P(\mathbb{N})) \prec ...$

Theorem 5.17 (Cantor-Schröeder-Bernstein)**.** *Let A, B be sets and supose that $A \preceq B \wedge B \preceq A$ then $A \approx B$.*

Proof. Let $f : A \to B$ and $g : B \to A$ be injective functions. And let $k = g \circ f : A \to A$ be the injective composition of those functions. Note that $g : B \to Im(g)$ is invertible and let $g^{-1} : Im(g) \to B$ be the inverse map. Define the following sequence of sets:

$$A_0 = A \setminus Im(g),\ A_{n+1} = k'' A_n$$

Let $D = \cup_{n\in\mathbb{N}} A_n$. Now we are ready to define the function $h : A \to B$ which is going to be a bijection:

$$h(x) = \begin{cases} g^{-1}(x) & x \notin D \\ g^{-1}(k(x)) & x \in D \end{cases}$$

Let us prove that h is well defined (i.e. that we can apply g^-1 in the definition of h) Indeed, if $x \in D$ then $k(X) = g(f(x)) \in Im(g)$ and if $x \notin D$, then in particular $x \notin A_0 = A \setminus Im(g)$. Hence $x \in Im(g)$.

Claim 5.17.1. *$x \in D$ if and only if $h(x) \in g^{-1''}D$.*

Proof. If $x \notin D$, then $h(x) = g^{-1}(x)$, and since g^{-1} is one-to-one, $g^{-1}(x) \notin g^{-1''}D$. If $x \in D$, then $x \in A_n$ for some n and therefore $k(x) \in K'' A_n = A_{n+1} \subseteq D$. It follows that $h(x) = g^{-1}(k(x)) \in g^{-1''}D$. □

h is one-to-one: Suppose that $y = h(x_1) = h(x_2)$. If $y \notin g^{-1''}D$, then by the claim $x_1, x_2 \notin D$ and therefore $g^{-1}(x_1) = h(x_1) = h(x_2) = g^{-1}(x_2)$. Since g^{-1} is one-to-one, $x_1 = x_2$. If $y \in g^{-1''}D$, then $x_1, x_2 \in D$ and therefore $g^{-1}(k(x_1)) = h(x_1) = h(x_2) = g^{-1}(k(x_2))$. Since both g^{-1}, k are one -to-one we have that $x_1 = x_2$.

h is onto: Let $b \in B$ and consider $g(b) \in Im(g)$. If $g(b) \notin D$, then $h(g(b)) = g^{-1}(g(b)) = b$. If $g(b) \in D$, then there is n such that $g(b) \in A_n$. Note that $n > 0$ since $A_0 = A \setminus Im(g)$. Hence $g(b) \in k'' A_{n-1}$ and there if $a \in A_{n-1} \subseteq D$ such that $k(a) = g(b)$. It follows that $h(a) = g^{-1}(k(a)) = g^{-1}(g(b)) = b$. □

Example 5.18. Prove that ${}^{\mathbb{N}}N \approx P(\mathbb{N})$

Proof. On one hand we have $P(\mathbb{N}) \approx {}^{\mathbb{N}}\{0,1\} \preceq {}^{\mathbb{N}}\mathbb{N}$ (the last equality is due to inclusion) on the other hand we have ${}^{\mathbb{N}}\mathbb{N} \subseteq P(\mathbb{N} \times \mathbb{N}) \approx P(\mathbb{N})$. So by Cantor-Schroeder-Berstein $P(\mathbb{N}) \approx {}^{\mathbb{N}}\mathbb{N}$. $\square$

Theorem 5.19. $\mathbb{R} \approx {}^{N}\{0,1\}$

Proof. On one hand we have that every $x \in \mathbb{R}$ is a Dedekind cut so $x \in P(\mathbb{Q})$ and therefore

$$\mathbb{R} \preceq P(\mathbb{Q}) \approx P(\mathbb{N}) \approx {}^{N}\{0,1\}$$

For the other direction, we will define a function $F : {}^{\mathbb{N}}\{1,2\} \to \mathbb{R}$ defined by

$$F(f) = 0.f(0)f(1)f(2)...$$

is one-to-one as every decimal representation is not eventually 0. Also it is clear that $\{0,1\} \approx \{1,2\}$ hence

$${}^{\mathbb{N}}\{0,1\} = {}^{\mathbb{N}}\{1,2\} \preceq \mathbb{R}$$

By Cantor- Schroeder-Berstein, $\mathbb{R} \approx {}^{\mathbb{N}}\{0,1\}$ $\square$

In particular $\mathbb{R}$ is uncountable.

Problem 16. *Prove that* ${}^{\mathbb{N}}\{0,1\} \times {}^{\mathbb{N}}\{0,1\} \approx {}^{\mathbb{N}}\{0,1\}$ *[Hint: consider the interweaving function that take two binary sequences* $\langle a_0, a_1, ...\rangle, \langle b_0, b_1, ...\rangle$ *and outputs* $\langle a_0, b_0, a_1, b_1, a_2, b_2, ...\rangle$*]*

About this result, Cantor said: “My eyes can see it but I cannot believe it”.

Theorem 5.20. *for every* $n \geq 1$, $\mathbb{R}^n \approx \mathbb{R}$.

Proof. It suffices to prove that $\mathbb{R} \times \mathbb{R} \approx \mathbb{R}$ and then the same inductive argument as with the case of the natural numbers will work. Indeed,

$$\mathbb{R} \times \mathbb{R} \approx {}^{\mathbb{N}}\{0,1\} \times {}^{\mathbb{N}}\{0,1\} \approx {}^{\mathbb{N}}\{0,1\} \approx \mathbb{R}$$

$\square$

Theorem 5.21. *For every* $\alpha < \beta$ *reals* $[\alpha, \beta] \approx (\alpha, \beta) \approx (\alpha, \infty) \approx \mathbb{R}$

Proof. First we note that $tn : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ is one-to-one and onto hence $(-\frac{\pi}{2}, \frac{\pi}{2}) \approx \mathbb{R}$. Since $(-\frac{\pi}{2}, \frac{\pi}{2}) \subseteq [-\frac{\pi}{2}, \frac{\pi}{2}] \subseteq (\frac{\pi}{2}, \infty) \subseteq \mathbb{R}$ we also have that all those sets are equinumerable. Now it is not hard to find bijections of the from $f(x) = ax + b$ which moves (α, β) to $(-\frac{\pi}{2}, \frac{\pi}{2})$ and $[\alpha, \beta]$ to $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and (α, ∞) to $(-\frac{\pi}{2}, \infty)$. $\square$

Definition 5.22. The continuum hypothesis (CH): Every set $A \subseteq \mathbb{R}$ is either finite, countable, or is equinumerable to the reals.

Theorem 5.23 (Godel and Cohen)**.** *The continuum hypothesis cannot be proven nor refuted from* ZFC*.*

Theorem 5.24. *(AC) The countable union of at most countable sets is at most countable*

Proof. Let A_n be a sequence of sets such that for each n, A_n is at most countable. Let us define B_n as follows, $B_0 = A_0$ and $B_{n+1} = A_{n+1} \setminus (\cup_{k=0}^{n} A_k)$. Since $B_n \subseteq A_n$, our assumption that A_n is at most countable implies that there is $f_n : B_n \to \mathbb{N}$ which is one-to-one. Note that if $n \neq m$ then $B_n \cap B_m = \emptyset$ and also that $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} B_n$. Define $g : \bigcup_{n\in\mathbb{N}} A_n \to \mathbb{N} \times \mathbb{N}$ by $g(n) = \langle m_n, f_{m_n}(n)\rangle$, where $m_n \in \mathbb{N}$ is the unique index such that $n \in B_{m_n}$. Then g is one-to-one and therefore $\bigcup_{n\in\mathbb{N}} An \preceq \mathbb{N} \times \mathbb{N} \preceq \mathbb{N}$. □

Corollary 5.25. *The following sets are countable: $\{X \in P(\mathbb{N}) \mid X \text{ is finite }\}$, the set of finite sequence of natural numbers, the set of all algebraic numbers.*

Proof. (1) Clearly $A_1 := \{X \in P(\mathbb{N}) \mid X \text{ is finite }\}$ is infinite and therefore $\mathbb{N} \preceq A_1$. To see that it is at most uncountable, note that $A_1 = \cup_{n\in\mathbb{N}} P(\{0, ..., n\})$ which is a countable union of finite (so at most countable) sets and therefore A_1 is at most countable.

(2) We are asked to prove that the set $\cup_{n\in\mathbb{N}_+} \mathbb{N}^n$ is countable. It is clearly infinite and is already given to us as a countable union of countable sets which is therefore at most countable.

(3) An algebraic number is a real number r which is a root of a non-zero polynomial with integer coefficients. Let $\mathbb{Z}[x]$ denote the set of all polynomials with integer coefficients. Then each non-zero polynomial has some degree $n \in \mathbb{N}$ and has the form $p(x) = z_n x^n + z_{n-1}x^{n-1} + ...z_1 x + z_0$. Let $\mathbb{Z}_n[X]$ be the set of all polynomials of degree at most n. Then clearly, $\mathbb{Z}_n[X] \approx \mathbb{Z}^{n+1}$ and therefore $\mathbb{Z}_n[X]$ is countable. Note that $\mathbb{Z}[X] = \cup_{n\in\mathbb{N}} \mathbb{Z}_n[X]$ and therefore is a countable union of countable sets (hence countable). Now the set of algebraic numbers is just $\cup_{p(x)\in\mathbb{Z}[X]} roots(p(x))$ where $roots(p(x)) = \{r \in \mathbb{R} \mid p(r) = 0\}$. Recall that every polynomial has only finitely many roots and therefore the set of algebraic numbers is a countable union of finite sets and therefore at most countable.

□

Corollary 5.26. *The following sets are uncountable: $\{X \in P(\mathbb{N}) \mid X \approx \mathbb{N}\}$, $\mathbb{R} \setminus \mathbb{Q}$, $\{r \in \mathbb{R} \mid r \text{ is transendental}\}$,*

Proof. Lets just prove one of them. If for example $\mathbb{R}\setminus\mathbb{Q}$ was countable, then $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} \setminus \mathbb{Q})$ would have been a countable union of countable sets and therefore countable. Contradiction. □

6. Cardinal numbers

With finite sets we have a natural number assigned to each finite set A according to the number of elements in A which we dented by $|A|$. This number determines completely when to sets are equinumerable:

Proposition 6.1. *Let A, B be two finite sets. Then $A \approx B$ if and only if $|A| = |B|$.*

We would like to extend this also to infinite sets and assign a quantity/number, which we call *a cardinal*, to each set which will determine the equinumerability relation. A first attempt would be to define that a cardinal is just an $\approx$-equivalence class. Indeed, we saw the following: For every sets A, B, C

(1) $A \approx A$.
(2) $A \approx B \Rightarrow B \approx A$.
(3) $A \approx B \wedge B \approx C \Rightarrow A \approx C$.

Does it mean that $\approx$ is an equivalence relation?

The problem is that there is no set on which this relation is defined and therefore there is no formal object which is the $\approx$-equivalence class.

To overcome this difficulty, we will need to choose somehow a representative $\kappa_{\mathcal{C}} \in \mathcal{C}$ for every $\approx$-equivalence class $\mathcal{C}$ and we would like to write for every $X \in \mathcal{C}$, $|X| = \kappa_{\mathcal{C}}$. For example, since $n = \{0, ..., n-1\}$ has n elements, we can choose $\kappa_{\mathcal{C}} = n$ as the representative of the class $\mathcal{C} = \{A \mid A \approx n\}$. Then we need to prove the following:

Definition 6.2. For every finite set A there is a unique n such that $A \approx n$.

Another equivalence class is the class of countable sets:

Definition 6.3. Denote by $\aleph_0 = \mathbb{N}$. We define $|A| = \aleph_0$ if and only if A is countable.

For now, let us assume that we have made such a canonical choice (this will be formally defined later) so if κ is a cardinal and $A \approx \kappa$ we may write $|A| = \kappa$.

Definition 6.4. Let κ, λ be cardinals. we define:

(1) $\kappa + \lambda = |A \uplus B|$ where A, B are **disjoint** sets such that $|A| = \kappa$ and $|B| = \lambda$.
(2) $\kappa \cdot \lambda = |A \times B|$ where $|A| = \kappa$ and $|B| = \lambda$.
(3) $\kappa^\lambda = |{}^B A|$ where $|A| = \kappa$ and $|B| = \lambda$.

We need to check that these operations does not depend on the choice of A, B.

Exercise 10. *If $A \approx A'$ and $B \approx B'$ then:*

(1) Given that A, B are disjoint and A', B' are disjoint, $A \uplus B \approx A' \uplus B'$.
(2) $A \times B \approx A' \times B'$.
(3) ${}^B A \approx {}^{B'} A'$

Proof. Let us prove (1) for example. Fix a bijections $f : A \to A'$ and $g : B \to B'$. Define $h : A \uplus B \to A' \uplus B'$ by

$$h(x) = \begin{cases} f(x) & x \in A \\ g(x) & x \in B \end{cases}$$

One should check that h is indeed a bijection. $\square$

Theorem 6.5 (Basic properties)**.** *Let κ, λ, σ be any cardinals (finite or infinite) then*

(1) $\kappa + \lambda = \lambda + \kappa$, $\kappa \cdot \lambda = \lambda \cdot \kappa$ *(commutativity)*
(2) $(\kappa + \lambda) + \sigma = \kappa + (\lambda + \sigma)$, $\kappa \cdot (\lambda \cdot \sigma) = (\kappa \cdot \lambda) \cdot \sigma$.*(Associativity)*
(3) $\kappa \cdot (\lambda + \sigma) = \kappa \cdot \lambda + \kappa \cdot \sigma$.*(Distributively)*
(4) $\kappa + 0 = \kappa,\ \kappa \cdot 0 = 0,\ \kappa \cdot 1 = \kappa,\ \kappa^1 = \kappa,\ 1^\kappa = 1,\ 0^0 = 1$, *for* $\kappa > 0$, $0^\kappa = 0$. *(Neutral elements)*
(5) *For every* n $\underbrace{\kappa + \kappa + \kappa + \kappa + ... + \kappa}_{n\ times} = n \cdot \kappa$, $\underbrace{\kappa \cdot \kappa \cdot \kappa \cdot \kappa \cdot ... \cdot \kappa}_{n\ times} = \kappa^n$.

Proof. Let us prove for example (3), Let A, B, C be such that $|A| = \kappa$, $|B| = \lambda$, $|C| = \sigma$ such that $B \cap C = \emptyset$. We need to prove that $A \times (B \uplus C) \approx (A \times B) \uplus (A \times C)$. Note that indeed $A \times B \cap A \times C = A \times (B \cap C) = A \times \emptyset = \emptyset$ and that $A \times (B \cup C) = (A \times B) \cup (A \times C)$. Since we have set equiality we have in particular equinumerability.

Let us prove for example in (2) that $\kappa \times \lambda = \lambda \times \kappa$. We need to prove that $A \times B \approx B \times A$. Clearly the function $f : A \times B \to B \times A$ defined by $f(\langle a, b\rangle) = \langle b, a\rangle$ a bijection between these sets.

Let us prove in (4) that $0^0 = 1$. We need to prove that ${}^{\emptyset}\emptyset \approx 1$. Indeed, one should check formally that $\emptyset : \emptyset \to \emptyset$ is the unique function in that set hence ${}^{\emptyset}\emptyset = \{\emptyset\}$ (which actually equals $1 = \{0\} = \{\emptyset\}$).

□

Remark 6.6. It should be proven (and the proof is omitted here) that for natural numbers this is the usual definition of addition, multiplication and power.

Proposition 6.7. (1) $\aleph_0 + \aleph_0 = \aleph_0$.
(2) $\aleph_0 + n = \aleph_0$.
(3) $\aleph_0 \cdot \aleph_0 = \aleph_0$.

Proof. For (1), $\aleph_0 = |\mathbb{N}_{eben}|$ and also $\aleph_0 = |\mathbb{N}_{odd}|$ which are disjoint sets. Hence

$$\aleph_0 + \aleph_0 = |\mathbb{N}_{even} \cup \mathbb{N}_{odd}| = |\mathbb{N}| = \aleph_0$$

For (2), let $n \in \mathbb{N}$. Then $\mathbb{N} \setminus \{0, ..., n-1\}$ is countable (witnessed by the function $m \mapsto m + n$). Hence $\aleph_0 + n = |(\mathbb{N} \setminus \{0, ..., n-1\}) \cup \{0, ..., n-1\}| = |\mathbb{N}| = \aleph_0$.

For (3), we saw that $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$ which implies that $\aleph_0 \cdot \aleph_0 = \aleph_0$ □

Corollary 6.8. $|\mathbb{R}| = 2^{\aleph_0}$

Proof. Indeed, we prove that $\mathbb{R} \approx {}^{\mathbb{N}}\{0, 1\}$ and by definition of exponent, $|\mathbb{R}| = 2^{\aleph_0}$. □

Proposition 6.9. (1) $2^{\aleph_0} + 2^{\aleph_0} = 2^{\aleph_0}$.
(2) $2^{\aleph_0} + \aleph_0 = 2^{\aleph_0}$.
(3) $2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0}$.

Proof. For (1) $\mathbb{R} \approx [0, \infty) \approx (-\infty, 0)$ and therefore

$$2^{\aleph_0} + 2^{\aleph_0} = |(-\infty, 0) \cup [0, \infty)| = |\mathbb{R}| = 2^{\aleph_0}.$$

For (2), Note that $(0,1) \subseteq \mathbb{R} \setminus \mathbb{N} \subseteq \mathbb{R}$. So by Cantor-Schroeder-Bernstein, $\mathbb{R} \setminus \mathbb{N} \approx \mathbb{R}$. Hence

$$2^{\aleph_0} + \aleph_0 = |(\mathbb{R} \setminus \mathbb{N}) \cup \mathbb{N}| = |\mathbb{R}| = 2^{\aleph_0}.$$

For (3), we have seen that $\mathbb{R} \times \mathbb{R} \approx \mathbb{R}$ □

Definition 6.10. We define $\kappa \leq \lambda$ if $A \preceq B$ where $|A| = \kappa$ and $|B| = \lambda$

Problem 17. *Prove that $\kappa \leq \lambda$ does not depend on the choice of representatives.*

By cantor-Schroeder-Bernstein theorem we have that:

Corollary 6.11. *$\kappa \leq \lambda$ and $\lambda \leq \kappa$ then $\kappa = \lambda$.*

Zermelo'z theorem says that every two cardinalities are comperable:

Theorem 6.12 ((AC))**.** *For every two cardinals κ, λ, either $\kappa \leq \lambda$ or $\lambda \leq \kappa$.*

Theorem 6.13 (Monotonicity)**.** *If $\kappa \leq \lambda$ and $\sigma \leq \tau$ then*

(a) $\kappa + \sigma \leq \lambda + \tau$.
(b) $\kappa \cdot \sigma \leq \lambda \cdot \tau$.
(c) $\kappa^\sigma \leq \lambda^\tau$ (except for the case $0^0 = 1 > 0 = 0^\kappa$ for every $\kappa > 0$).

Proof. Let us prove fr example (c). The assumption $\kappa \leq \lambda$ and $\sigma \leq \tau$ translates to sets A, B, C, D of cardinality $\kappa, \lambda, \sigma, \tau$ respectively such that $A \preceq B$ and $C \preceq D$. We need to prove that ${}^A C \preceq {}^B D$. Let us split into cases:

(1) If $B = \emptyset$, then also $A = \emptyset$ (since $A \preceq B$).
   (a) If $D = \emptyset$ then also $C = \emptyset$ and then ${}^C A = \{\emptyset\} = {}^D B$.
   (b) $D \neq \emptyset$, in this case, the assumptions of the theorem implies that $\tau \neq \emptyset$ and therefore ${}^C A = \emptyset = {}^D B$.

(2) Suppose that $B \neq \emptyset$ and let $b^* \in B$ be any element and let $F : A \to B$, $G : C \to D$ be injections. We need to define an injection $\Phi : {}^C A \to {}^D B$. For a function $f : C \to A$, we define $\Phi(f) : D \to B$ by

$$\Phi(f)(d) = \begin{cases} b^* & d \notin Im(G) \\ F(f(G^{-1}(b))) & d \in Im(G) \end{cases}$$

Let us prove that Φ is one-to-one. Let $f, g : C \to A$, and assume that $f \neq g$. We need to prove that $\Phi(f) \neq \Phi(g)$. By function inequality, there is $c \in C$ such that $f(c) \neq g(c)$. Consider $d = G(c) \in Im(G)$. Then by definition

$$\Phi(f)(d) = F(f(G^{-1}(d))) = F(f(G^{-1}(G(c)))) = F(f(c)) \underset{F \text{ is 1-1}}{\neq} F(g(c)) = ... = \Phi(g)(d)$$

□

Corollary 6.14. $2^{\aleph_0} \cdot \aleph_0 = 2^{\aleph_0}$

Proof. $2^{\aleph_0} = 2^{\aleph_0} \cdot 1 \leq 2^{\aleph_0} \cdot \aleph_0 \leq 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0}$. ☐

Definition 6.15. We denote by $\kappa < \lambda$ if $\kappa \leq \lambda$ and $\kappa \neq \lambda$

Corollary 6.16. *For every cardinal κ, $\kappa < 2^\kappa$.*

Proof. Let A be of cardinality κ. By Cantor's theorem $A \prec {}^A\{0,1\}$. By definition $2^\kappa = |{}^A\{0,1\}|$ hence $\kappa < 2^\kappa$. ☐

Rules of exponent:

Theorem 6.17. *(1)* $(\kappa^\lambda)^\sigma = \kappa^{\sigma \cdot \lambda}$.
(2) $\kappa^{\lambda+\sigma} = \kappa^\lambda \cdot \kappa^\sigma$.
(3) $(\kappa \cdot \lambda)^\sigma = \kappa^\sigma \cdot \lambda^\sigma$

Proof. Let us prove for example (1). Let A, B, C be of cardinalities κ, λ, σ respectively. we need to prove that

$$^C(^BA) \approx {}^{C\times B}A$$

Define $\Phi : {}^{C\times B}A \to {}^C(^BA)$ as follows. For every $f : C \times B \to A$, let $\Phi(f) : C \to {}^BA$ be the that takes $c \in C$ and outputs $\Phi(f)(c) : B \to A$, which in turn is defined by

$$\Big(\Phi(f)(c)\Big)(b) = f(\langle c, b\rangle)$$

Let us check that the function $\Psi : {}^C(^BA) \to {}^{C\times B}A$ defined by $\Psi(g)(\langle c, b\rangle) = (g(c))(b)$ is inverse to Φ. We shall only prove that $\Psi \circ \Phi = Id_{{}^{C\times B}A}$ and leave to the reader the second composition. Let $f : C \times B \to A$ we need to prove that $\Psi(\Phi(f)) = f$ so let $\langle c, b\rangle \in C \times B$, then

$$\Psi(\Phi(f))(\langle b, c\rangle) = (\Phi(f)(c))(b) = f(\langle c, b\rangle)$$

as wanted. ☐

Corollary 6.18. $(\aleph^0)^{\aleph_0} = 2^{\aleph_0}$ *and* $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}$

Proof. $2^{\aleph_0} \leq (\aleph_0)^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} \leq 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}$ ☐

7. The Axiom of Choice

Every time we perform the following:

"$X \neq \emptyset$ let $x \in X$"

we are making a choice. This line can appear only finitely many times in a formal proof and therefore we are allowed to choose finitely many times. However, we encounter a problem if we would like to choose infinitely many times.

Definition 7.1. Let $\mathcal{F}$ be set of non-empty sets. A choice function for $\mathcal{F}$ is a function $f : \mathcal{F} \to \cup\mathcal{F}$ such that for each $A \in \mathcal{F}$, $f(A) \in A$.

Example 7.2. Let $\mathcal{F} = P(\mathbb{N}) \setminus \{\emptyset\}$, then $f(X) = \min(X)$ is a choice function for $\mathcal{F}$.

Problem 18. *Find a choice function for* $P(\mathbb{Q}) \setminus \{\emptyset\}$

Example 7.3. Let $\mathcal{F} = P(\mathbb{R}) \setminus \{\emptyset\}$. Then it is provable that there is no explicit choice function for $\mathcal{F}$.

Axiom (Ax9. Choice)**.** For every set $\mathcal{F}$ such that $\emptyset \notin \mathcal{F}$, there exists a choice function.

We denote the axiom of choice by AC. Here are some basic theorems which use the axiom of choice:

(1) If $g : A \to B$ is onto then there is $f : B \to A$ such that $g \circ f = Id_B$.
(2) If A is infinite then $\mathbb{N} \precsim A$.
(3) $A \precsim B$ iff there is a function $f : B \to A$ which is onto.
(4) The countable union of countable sets is countable.

Other non set-theoretic examples:

(1) Every field has an algebraically closed closure.
(2) Every ideal is contained in a maximal ideal.
(3) There exists a set which is not Lebesgue measurable.
(4) Tychonoff's theorem: a product of compact topological spaces is compact.
(5) Hahn-Banach theorem.
(6) Completeness theorem for first order logic.
(7) the compactness theorem for first order logic.
(8) $\mathbb{R}$ can be well ordered.

List of axioms of the system ZF (Zermelo -Frenkel):

Ax0. Existence
Ax1. Extensionality.
Ax2. Foundation (will be introduced later)
Ax3. Comprehention.
Ax4. Pairing
Ax5. Union.
Ax6. Replacement.
Ax7. Infinity.
Ax8. Powerset.

The axioms of the system ZFC(Zermelo-Frenkel-Choice) are $ZF + AC$.

Theorem 7.4. *The following are equivalent:*

(1) AC.
(2) Every set can be well ordered (The well order theorem).
(3) Zermelo's Theorem.
(4) Zorn's lemma.

We will introduce and prove the equivalent statements above in the next few sections.

Corollary 7.5 (AC)**.** *There is a system of representatives for all possible cardinalities.*

Corollary 7.6 (AC)**.** *For any set A, $A \times A \approx A$.*

How do we avoid choice:

Theorem 7.7. *Suppose that $\mathcal{A}$ is a set of open pairwise disjoint intervals in $\mathbb{R}$, then $|\mathcal{A}| \leq \aleph_0$.*

Proof. We are not allowed to use AC. Let us pick (one choice!) a bijection between $\mathbb{Q}$ and $\mathbb{N}$ and assume that $\mathbb{Q} = \{q_n \mid n \in \mathbb{N}\}$. Let us define $f : \mathcal{A} \to \mathbb{Q}$ by setting $f(X) = q_n$ for the minimal n such that $q_n \in X \cap \mathbb{Q}$. Then f exists by the axiom of comprehension (and others in ZF). since the intevals are pairwise disjoint, it is not hard to check that f is one-to-one and therefore $|\mathcal{A}| \leq \aleph_0$. $\square$

8. WELL ORDERS AND ORDINALS

Recall that a (strong) order on a set A is a relation R which is transitive, reflexive, and strongly-anti-symmetric. R is total if every any two members $a, b \in A$ are R-comparable, namely: $a = b \vee aRb \vee bRa$.

Definition 8.1. An total order R on A is called a *well-order* if:

$$\forall X \subseteq A.X \neq \emptyset \Rightarrow \exists \min{}_R(X)$$

where $\min_R(X)$ is a (unique) element in $x \in X$ such that $\forall y \in X.x \neq y \Rightarrow xRy$.

Example 8.2.
- Every total order on a finite set is a well-order.
- $\mathbb{N}$ with the regular order is a well order.
- $\mathbb{N} \times \mathbb{N}$ with the lexicographic order is a well order.
- Consider the following order of ${}^{\mathbb{N}}\mathbb{N}$ given by fRg iff $f(n^*) < g(n^*)$ where $n^* = \min\{n \mid f(n) \neq g(n)\}$. Then R is a total ordering of ${}^{\mathbb{N}}\mathbb{N}$ which is not a well-order.

Theorem 8.3. *(AC) Every set can be well-ordered.*

The proof for this will be given later. For now let us prove the other direction:

The well order Theorem implies the axiom of choice: Let $\mathcal{F}$ be any family of non-empty sets. Let $A = \bigcup \mathcal{F}$. By the well order theorem, there is a well ordering $\prec$ on A. Define a choice function $f : \mathcal{F} \to A$ by $f(X) = \min_\prec X$. Note that since $\emptyset \notin \mathcal{F}$, f is well-defined. $\square$

Definition 8.4. Let $\langle A, R \rangle$ be an ordered set, define $A_R[x] = \{y \in A \mid yRx\}$.

Lemma 8.5. *If $\langle A, R \rangle$ is a well order then for any $x \in A$, $\langle A, R \rangle \not\simeq \langle A_R[x], R \rangle$.*

Proof. Suppose that $f : R \to A_R[x]$ witnesses otherwise, let $B = \{y \mid f(y)Ry\}$. B is not empty since $f(x) \in A_R[x]$ and therefore $f(x)Rx$. Let $x^* = \min_R(B)$, then $f(x^*)Rx^*$ and since f is order preserving $f(f(x^*))Rf(x^*)$, hence $f(x^*) \in B$, contradictiong the minimality of x^*. $\square$

Problem 19. *Find a counter-example for the previous lemma in case that $\langle A, R\rangle$ is not well ordered.*

Lemma 8.6. *Suppose $\langle A, R\rangle, \langle B, S\rangle$ are well-orders and $\langle A, R\rangle \simeq \langle B, S\rangle$. Then the isomorphism between them is unique.*

Proof. Suppose that g_1, g_2 are two isomorphisms and toward contradiction assume that $g_1 \neq g_2$. Let $x_* = \min\{x \in A \mid g_1(x) \neq g_2(x)\}$. Then $g_1(x^*) \neq g_2(x^*)$. Without loss of generality, suppose taht $b := g_1(x^*) S g_2(x^*)$ and let yRx^* be such that $g_2(y) = b$, then $g_1(y) S g_1(x^*) = b = g_2(y)$, thus $g_1(y) \neq g_2(y)$ and therefore $y \in \{x \mid g_1(x) \neq g_2(x)\}$ contradiction the minimality of x^*. □

Definition 8.7. Let $\langle A, R\rangle$ be a well-ordering A set $X \subseteq A$ is called an initial segment if $\forall y \in X \forall z \in A.zRy \to z \in X$.

Lemma 8.8. *Let $\langle A, R\rangle$ be a well-ordering and $X \subseteq A$. Then X is an initial segment iff $X = A$ or $\exists x \in A.A_R[x] = X$.*

Proof. Exercise. [Hint: define $x = \min A \setminus X$] □

Theorem 8.9 (The trichotomy theorem of well-ordering)**.** *Let $\langle A, R\rangle$,$\langle B, S\rangle$ be well-ordering. Then exactly one of the following holds:*

(1) $\langle A, R\rangle \simeq \langle B, S\rangle$.
(2) there is $x \in A$ such that $\langle A_R[x], R\rangle \simeq \langle B, S\rangle$.
(3) there is $y \in B$ such that $\langle A, R\rangle \simeq \langle B_S[y], S\rangle$.

Proof. Let

$$f = \{\langle a, b\rangle \in A \times B \mid \langle A_R[a], R\rangle \simeq \langle B_S[b], S\rangle\}$$

First we claim that $dom(f), Im(f)$ are initial segments. To see this, is suffices to prove that they are downward closed. For example, if $a'Ra$ and $a \in \mathrm{dom}(f)$ then there is b such that $\langle A_R[a], R\rangle \simeq \langle B_S[b], S\rangle$. Let $g : A_R[a] \to B_S[b]$ be an isomorphism witnessing this. Note that $A_R[a']$ is an initial segment of $A_R[a]$ and therefore $g \restriction A_R[a']$ is defined, order preserving and $1-1$. Let $b' = g(a')$, it is not hard to verify that $Im(g) = B_S[b']$ and therefore $g \restriction A_R[a']$ witnesses the fact that $\langle A_R[a'], R\rangle \simeq \langle B_S[b'], S\rangle$ which implies that $a' \in dom(f)$. Similarily, $Im(f)$ is an initial segment. Also f must be (univalent and) injective since otherwise, we would have had $a_1 R a_2$ such that $b = f(a_1) = f(a_2)$ and in particular $\langle A_R[a_1], R\rangle \simeq \langle B_S[b], S\rangle \simeq \langle A_R[a_2], R\rangle$ which contradicts the lemma that a well ordering is not isomorphic to its proper initial segments.

Finally, we claim that it is impossible that both $dom(f), Im(f)$ are proper initial segment, sense otherwise, $dom(f) = A_R[x]$ and $Im(f) = B_S[y]$ and we let $x' = \min A \setminus A_R[x]$ and $y' = \min B \setminus B_S[y]$, then we can extend f to be defined on $A_R[x']$ by sending $f(x) = y$ witnessing that $x' \in dom(f)$, contradiction. □

Corollary 8.10. *The Well-ordering theorem implies Zermelo's theorem.*

Proof. Let A, B be two set. Find any well orderings R, S on A, B respectively. By the trichotomy theorem Either (1) holds in which case we have produced bijection witnessing $A \approx B$, or (2), in which case there is an injective function $f : B \to A$ which witnesses that $B \prec A$, or (3) which similarly implies $A \preceq B$. □

8.1. **ordinals.** The basic theory is due to Von Neuman.

Definition 8.11. A set x is called trastivie if

$$\forall y \in x \forall z \in y . z \in x$$

Or equivalently,

$$\forall y \in x . y \subseteq x$$

Example 8.12. $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$

Exercise 11. *If $\mathcal{F}$ is a set of transitive sets then $\bigcup \mathcal{F}, \bigcap \mathcal{F}$ are both transitive sets.*

Transitive sets are sets for which the $\in$-relation is transitive.

Definition 8.13. A set α is called an ordinal if α is a transitive set and

$$\in_\alpha := \{\langle x, y\rangle \in \alpha^2 \mid x \in y\}$$

is a well order on α.

Remark 8.14. The axiom of foundation and the axiom of choice will later tell us that an infinite decreasing $\in$- sequence does not exist and therefore it will suffice to require that $\in_\alpha$ is a total order.

Axiom (Ax. 2 Foundation)**.** For every set $A \neq \emptyset$ there is $x \in A$ such that $x \cap A = \emptyset$.

Proposition 8.15 (AC)**.** *The following are equivalent:*

1. *The axiom of foundation.*
2. *There is no infinite deceasing sequence in the $\in$-relation.*

Proof. (1) implies (2) is clear and does not require the axiom of choice. For the other direction, let us prove that $\neg(1)$ implies $\neg(2)$. Let A be a set witnessing the failure of the axiom of foundation. Let f be a choice function for $P(A) \setminus \{\emptyset\}$. We would like to construct a decreasing sequence in the $\in$-relation. We shall define a sequence $a_n \in A$ recursively. Since $A \neq \emptyset$ let $a_0 = f(A)$ be an element. Suppose we have defined $a_n \in a_{n-1} \in ... \in a_1 \in a_0$ and let us define a_{n+1}. By our assumption on A, there is no element $x \in A$ such that $A \cap x = \emptyset$, and therefore $a_n \cap A \neq \emptyset$. Let $a_{n+1} = f(A \cap a_n)$, then $a_{n+1} \in a_n \cap A$. □

Example 8.16. $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ the set $\{\emptyset, \{\emptyset\}\ \{\{\emptyset\}\}\}$ is an example of a transitive set which is not an ordinal (since $\emptyset$ and $\{\{\emptyset\}\}$ are not $\in$-comperable). If $x = \{x\}$ then x is not an ordinal since we will have $x \in x$ and therefore $\in$ is not anti reflexive. For the same reason, for every ordinal α, $\alpha \notin \alpha$.

Theorem 8.17. *(1) If α is an ordinal and $x \in \alpha$ then x is an ordinal and $x = \alpha_{\in}[x]$.*

(2) $\alpha \subseteq \beta$ iff $\alpha \in \beta \vee \alpha = \beta$.

(3) If α, β are ordinals such that $\alpha \simeq \beta$ then $\alpha = \beta$.

(4) For every two ordinal α, β, $\alpha \in \beta \vee \beta \in \alpha \vee \alpha = \beta$.

(5) If C is a set of ordinals then there is $\min_{\in}(C)$.

(6) If C is a set of ordinals then $\cup C$ is an ordinal and had the property of supremum, namely, it is an upper bound of C: $\forall \alpha \in C.\alpha \subseteq \cup C$ and if β is an upper bound for C then $\cup C \subseteq \beta$.

Proof. (1) exercise. (2), from right to left is easy. From left to right, suppose that $\alpha \subseteq \beta$ and $\alpha \neq \beta$, let $\gamma = \min(\beta \setminus \alpha)$, we claim that $\gamma = \alpha$. If $x \in \gamma$, then $x \in \beta$ and bu minimality of γ, $x \in \alpha$. If $x \in \alpha$, then $x \in \beta$ by inclusion. x, γ are comparable in $\in$, but $\gamma = x$ and $\gamma \in x$ is ruled out since $\gamma \in \beta \setminus \alpha$, so $x \in \gamma$. By double inclusion $\alpha = \gamma$. For (3), suppose that there is $x \in \alpha$ such that $f(x) \neq x$ and let x be the minimal such x. Then x is an ordinal and $x = f[x] \subseteq \beta$, but then $x \in \beta$ and $x \in f(x)$ so there is $y \in \alpha$ such that $f(y) = x \neq y$ but then $y \in x$ since f is order-preserving which contradicts the minimality of x. (4) follows from (1), (3) and the trichotomy theorem □

Corollary 8.18. *$\neg\exists z.\forall x.x$ is an ordinal $\Rightarrow x \in z$*

Proof. Otherwise, let $On = \{\alpha \in z \mid \alpha \text{ is an ordinal}\}$ (which exists by comprehansion), then On is a transitive set (by (1) of the previous theorem) and $\in$ well orders On (by (3) and (5)) and therefore On is itself an ordinal, so $On \in On$. However, no ordinal can be a member of itself, contradiction. □

We denote the class of all ordinals by On.

Remark 8.19. As we have just proved, there is no formal object which is On in the mathematical universe, thus, there is no formal distinction between "$x \in On$" and "x is an ordinal", or "$A \subset On$" and "$\forall x \in A$, x is an ordinal".

Axiom (Ax6. Replacment)**.** The axiom of replacement states that for every set A and every formula $\phi(x, y)$ such that $\forall a \in A \exists! y \phi(a, y)$, the set $\{y \mid \exists a \in A,\ \phi(a, y)\}$ exists.

Theorem 8.20. *For any well-ordered set $\langle A, R\rangle$ there is a unique ordinal α such that $\langle A, R\rangle \simeq \langle \alpha, \in \rangle$. We call this α the order-type of $\langle A, R\rangle$ and denote it by* $\text{otp}(A, R)$*.*

Proof. Uniqueness follows from before. To prove existence, let $B = \{a \in A \mid \exists x \in On.\langle A_R[a], R\rangle \simeq \langle x, \in \rangle\}$. Note that for every $a \in B$, there is a unique ordinal x which witness $a \in B$. So we may apply replacement to B and form the set $C = \{x \in On \mid \exists a \in B.\langle A_R[a], R\rangle \simeq \langle x, \in \rangle\}$. We claim that C is an ordinal. First, since C is a set of ordinal, the $\in$ relation on C is a well order. To see that C is transitive, note that if $y \in x \in C$ and $\langle A_R[a], R\rangle \simeq \langle x, \in \rangle$ then there is $b \in A_R[a]$ such that $\langle A_R[b], R\rangle \simeq \langle y, \in \rangle$. Hence $b \in B$ and

$y \in C$. It follows that C is an ordinal. A similar argument proves that B is an initial segment of A and if $B = A_R[c]$ for some c then $c \in B$ by definition so $B = A$. □

Remark 8.21. Without the axiom of replacement, one cannot prove theorem 4.8 as there is a model of $ZFC - \{Ax6\}$ for which theorem 4.8 fails.

Notation 8.22. $\alpha < \beta$ iff $\alpha \in \beta$ and $\alpha \leq \beta$ iff $\alpha < \beta \vee \alpha = \beta$ iff $\alpha \subseteq \beta$.

Theorem 8.23. *(1) If α is an ordinal then $\emptyset \leq \alpha$.*
(2) If α is an ordinal then $\alpha + 1 := \alpha \cup \{\alpha\}$ is an ordinal and is the successor of α in the sense that it is the minimal ordinal greater than α.
(3) If A is a set of ordinals without a greatest element then $\sup A := \cup A$ is an ordinal strictly greater then all the ordinals in A.

Proof. Exercise. □

Definition 8.24. A successor ordinal is an ordinal of the form $\alpha + 1$, otherwise it is called limit.

Theorem 8.25 (Hertog's Theorem)**.** *For every set A there is an ordinal α such that $\alpha \not\preceq A$ i.e. there is no injection from α into A.*

Proof. Suppose otherwise, that there is a set A such that for every ordinal α there is an injection of α into A. In particular, for every ordinal $\beta \geq \alpha$, $\beta \sim \alpha$. Let $S = \{R \in P(A \times A) \mid X \in P(A),\ R \text{ well-orders } X\}$. S exists by the power set axiom and comprehension. Define for each $R \in S$, $F(R) = \text{otp}(\alpha, R)$. Then by replacement the following is a set exists $E = \{F(R) \mid R \in S\}$. By our assumption, for every ordinal β, there is an injection $f : \beta \to A$ and therefore we can translate the order $(\beta, \in)$ to a well order R on a subset $X \subseteq A$ such that $\text{otp}(\alpha, R) = \beta$. In other words we conclude that $\beta \in E$ and therefore $E = On$. This is a contradiction to the fact that On was already proven not to be a set. □

Corollary 8.26. *There is an uncountable ordinal*

Proof. otherwise every ordinal can be injected into $\mathbb{N}$ contradicting Hartog's theorem. □

Definition 8.27. Let ω_1 be the least uncountable cardinal.

Proposition 8.28. $\omega_1 = \{\alpha \in On \mid \alpha \textit{ is countable}\}$

Proof. If α is countable then it is impossible that $\omega_1 \leq \alpha$ since this would mean that $\omega_1 \subseteq \alpha$, contradiction α being countable. Hence $\alpha \in \omega_1$. If α is uncountable then $\omega_1 \leq \alpha$ since ω_1 is minimal. Hence $\alpha \notin \omega_1$, as wanted. □

Corollary 8.29. *Zermelo's theorem implies the well order theorem.*

Proof. Let A be any set. Then there is α such that $\alpha \not\preceq A$. By Zermelo's theorem this must imply that $A \preceq \alpha$ and therefore there is an injection $f : A \to \alpha$. Now we can define a well ordering on A as follows: $a \preceq b$ iff $f(a) < f(b)$. Hence A can be well ordered. □

8.2. **Transfinite recursion and induction.** We will formulate the induction and recursion theorem in a way that can be applied to what we call classes. Formally, a class does not exist as a mathematical object (as we have seen for V and for On). Given a formula $\pi(x)$ with a free variable x (we allow other free variables, indeed, the class we are defining might depend on parameters) we think of the class C_ϕ as the "collection" (whatever that means) $C_\phi = \{x \mid \phi(x)\}$. So whenever C_ϕ appears in a mathematical statement, it should be clear how to replace C_ϕ by ϕ, for example:

(1) $\forall x \in C_\phi.x$ satisfy... just mean $\forall x.\phi(x) \Rightarrow x$ satisfy...
(2) $C_\phi \subseteq On$ means

$$\forall x.\phi(x) \Rightarrow x \text{ is an ordinal.}$$

Note that if C_ϕ is a class and A is a set then $C_\phi \cap A = \{x \in A \mid \phi(x)\}$ which is a set that exists by comprehansion.

The next theorems are formulated for classes and take their usual meaning when the class is in fact a set:

Theorem 8.30. *Let $0 \neq C$ be a class of ordinal (formally, let ϕ be a formula such that $(\exists x.\phi(x)) \wedge (\forall x.\phi(x) \Rightarrow x$ is an ordinal$)$). Then there is $y = \min(C)$ (formally, $\exists y.\phi(y) \wedge \forall x.\phi(x) \to x \geq y$).*

Proof. Let $\alpha \in C$ be any ordinal, then $D = \alpha + 1 \cap C$ is a non-empty set of ordinals, and therefore $y = \min(D)$ exists. Let us prove that $y = \min(C)$. let $x \in C$, then either $x > \alpha$ in which chase $x > \alpha \geq y$ of $x \leq y$ but then $x \in D$ and therefore $x \geq y$. □

Formally, what we have above is a theorem scheme, one for every formula ϕ. This theorem enables us to prove the induction theorem over all the ordinal!:

Theorem 8.31 (The induction theorem)**.** *Let C be a class of ordinals such that for every ordinal α, if $\alpha \subseteq C$ then $\alpha \in C$, then $C = On$.*

Proof. Suppose otherwise, let $\beta = \min(On \setminus C)$. Then for every $\alpha < \beta$, $\alpha \in C$, but then $\beta \in C$ by our assumption. Contradiction. □

Corollary 8.32. *Let C be a class of ordinals such that*

(1) $\alpha \in C \Rightarrow \alpha + 1 \in C$.
(2) For every limit ordinal δ, if $\forall \beta < \delta$, $\beta \in C$ then $\in C$.

Then $C = On$.

Theorem 8.33 (The recursion theorem)**.** *Suppose that $F(x, y)$ is a formula such that $\forall x \exists! y.F(x, y)$. Then one can write down a formula $G(v, w)$ such that*

$$\forall \alpha \in On.\exists! w.G(\alpha, w) \wedge \forall \alpha \in On \exists x.\exists y.(x = G \restriction \alpha \wedge F(x, y) \wedge G(\alpha, y)$$

Before proving the theorem, let us explain the formulation of the theorem. The formula $F(x, y)$ is thought of as the formula $f(x) = y$ for some "function" $f : V \to V$ which accommodates some recursive information. Then the theorem says that there is a function $g : On \to V$ (which is given by the formula $G(v, w)$) such that for every $\alpha \in On$, $g(\alpha) = f(g \restriction \alpha)$).

To see how this relates to the usual way we define functions recursively, recall that in a recursive definition of a function, we assume that $\forall\beta < \alpha$, $g(\beta)$ has already been defined (in other words, $g \restriction \alpha$ has been defined) and given this unknown definition we define $g(\alpha)$. The purpose of the function f is to take that unknown $x = g \restriction \alpha$, which can be have any possible values, and the output $g(x)$ is what we would have wanted for the value of $g(\alpha)$ to be. The recursion theorem simply tells you that given a function f (which is defined on any possible sequence x) the function g which satisfies $g(\alpha) = f(g \restriction \alpha)$ exists. Since we are talking about classes, this is all formulated with formulas instead of functions.

Remark 8.34. In many situations we use the induction and recursion theorem simultaneously when we define a function g and assume that $g\restriction\alpha$ has already been defined and satisfies some properties, then we define $g(\alpha)$ and prove it satisfies some properties.

Example 8.35. Ordinal arithmetic: for a fixed α, we define:

- $\alpha + \beta$ by recursion on β
  (1) $\alpha + 0 = \alpha$.
  (2) $\alpha + (\beta + 1) = (\alpha + \beta) + 1$.
  (3) For a limit ordinal δ, we define $\alpha + \delta = \sup_{\beta<\delta} \alpha + \beta$.
- $\alpha \cdot \beta$ by recursion on β
  (1) $\alpha \cdot 0 = \alpha$.
  (2) $\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha$.
  (3) For a limit ordinal δ, we define $\alpha \cdot \delta = \sup_{\beta<\delta} \alpha \cdot \beta$.
- α^β by recursion on β
  (1) $\alpha^0 = 1$.
  (2) $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$.
  (3) For a limit ordinal δ, we define $\alpha^\delta = \sup_{\beta<\delta} \alpha^\beta$.

$1 + \omega = \sup_{n<\omega} 1 + n = \omega < \omega_1$

$2 \cdot \omega = \sup_{n<\omega} 2 \cdot n = \omega < \omega + \omega = \omega + 2$
$2^\omega = \sup_{n<\omega} 2^n = \omega$ (so 2^ω as ordinals and as cardinal is not the same!)
$\omega + \omega^2 = \omega^2$
$(\omega + 1)^2 = (\omega + 1) \cdot (\omega + 1) = (\omega + 1) \cdot \omega + \omega + 1 = \omega^2 + \omega + 1$.

Proposition 8.36. *(1) If $\alpha < \beta$ then for every γ, $\gamma + \alpha < \gamma + \beta$.*
(2) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$

Proof. For (1), We prove by trasfinite induction of β that for every $\alpha < \beta$ and every γ, $\gamma + \alpha < \gamma + \beta$. For $\beta = 0$, the claim is vacuously true (since there is no $\alpha < 0$). Suppose that the claim holds for β and let us prove it for $\beta + 1$. Let $\alpha < \beta + 1$ and γ be any ordinal. Let us split into cases:

- If $\alpha < \beta$, then by the induction hypothesis and the definition of "+" in the successor case,

$$\gamma + \alpha < \gamma + \beta < (\gamma + \beta) + 1 = \gamma + (\beta + 1)$$

- If $\alpha = \beta$, then as in the first case we get $\gamma + \beta < \gamma + (\beta + 1)$.

For limit β, let $\alpha < \beta$, then $\alpha + 1 < \beta$. By the induction hypothesis applies to $\alpha + 1$ and the definition of "+" is the limit case,

$$\gamma + \alpha < \gamma + (\alpha + 1) \leq \sup_{\delta < \beta} \alpha + \delta = \alpha + \beta$$

For (2), agan we prove it by induction on γ, for every α, β.

- For $\gamma = 0$ we have that:

$$(\alpha + \beta) + 0 = \alpha + \beta = \alpha + (\beta + 0)$$

- At successor step $\gamma + 1$, we have that

$$(\alpha+\beta)+(\gamma+1) = ((\alpha+\beta)+\gamma)+1 = (\alpha+(\beta+\gamma))+1 = \alpha+((\beta+\gamma)+1) = \alpha+(\beta+(\gamma+1))$$

- At limit steps γ, suppose that for every $\delta < \gamma$ we have that $(\alpha+\beta)+\delta = \alpha + (\beta + \delta)$, then

$$(\alpha + \beta) + \gamma = \sup_{\delta < \gamma} (\alpha + \beta) + \delta = \sup_{\delta < \beta} \alpha + (\beta + \delta) =^* \alpha + (\beta + \gamma)$$

  To see why $*$ holds, we will use (1) and the definition of supremum. Indeed, if $\delta < \gamma$ then from (1) we get that $\beta + \delta < \beta + \gamma$ and therefore (again from (1)), $\alpha + (\beta + \delta) < \alpha + (\beta + \gamma)$. Hence $\sup_{\delta < \gamma} \alpha + (\beta + \delta) \leq \alpha + (\beta + \gamma)$. Note that $\beta + \gamma = \sup_{\delta < \gamma} \beta + \delta$ by definition and therefore (since $\beta + \delta$ is strictly incresing with δ) we conclude that $\beta + \gamma$ is a limit ordinal and that $\sup\{\alpha + \rho \mid \rho < \beta + \gamma\}$. It follows that

$$\alpha + (\beta + \gamma) = \sup_{\rho < \beta + \gamma} \alpha + \rho$$

  Hence we need to check that

$$\sup_{\delta < \gamma} \alpha + (\beta + \delta) = \sup_{\rho < \beta + \gamma} \alpha + \rho$$

  We have that $\{\beta + \delta \mid \delta < \gamma\} \subseteq \beta + \gamma$ so " $\leq$ " is clear (the sup is taken over more elements). For the other direction, let $\rho < \beta + \gamma$ then there is $\delta < \gamma$ such that $\beta + \delta > \rho$ and by (1) we have that $\alpha + (\beta + \delta) > \alpha + \rho$ so " $\geq$ " follows.

$\square$

The next theorem concludes the equivalence between AC, the well-order theorem and Zermelo's theorem:

Theorem 8.37. *The axiom of choice implies the well-order theorem.*

Proof. Let A be a set and let f be a choice function for $P(A) \setminus \{\emptyset\}$. Fix any $x \notin A$ (which exists since A cannot be the set of all sets). Define by recursion a function g from On to $A \cup \{x\}$ as follows:

$$H(\alpha) = \begin{cases} f(A \setminus \{H(\beta) \mid \beta < \alpha\}) & \{H(\beta) \mid \beta < \alpha\} \subsetneq A \\ x & \text{otherwise} \end{cases}$$

Note that there must be α such that $H(\alpha) = x$, just otherwise, for each α, $H(\alpha) \in A \setminus \{H(\beta) \mid \beta < \alpha\}$ and therefore $H(\alpha) \neq H(\beta)$ for every $\alpha \neq \beta$. So for every α, $H \restriction \alpha$ is an injection of α into A contradicting Hartog's theorem. Let α be the minimal ordinal such that $H(\alpha) = x$. In follows that $\{H(\beta) \mid \beta < \alpha\} = A$ and therefore $H \restriction \alpha$ is a bijection from α into A. Now we can define a well ordering of A using H^{-1}. □

8.3. **cardinals.** Recall that if A can be well ordered, then there α such that $A \approx \alpha$

Definition 8.38. Suppose that A can be well ordered. Denote by $|A|$ to be the minimal ordinal α such that $A \approx \alpha$.

Definition 8.39. An ordinal α is called a cardinal if $\alpha = |\alpha|$. Equivalently, if for every $\beta < \alpha$, $\beta < |\alpha|$.

Clearly, if α, β are cardinals then $\alpha \not\approx \beta$.

Corollary 8.40. *If every set can be well-ordered then for every set A there is a unique cardinal $|A|$ such that $A \approx |A|$.*

Exercise 12. *(1) If $|\alpha| \leq \beta \leq \alpha$ then $|\alpha| = |\beta|$.*
(2) $n \not\approx n+1$ for every n. [Hint: induction.]
(3) If $|\alpha| = n$ then $\alpha = n$.

Corollary 8.41. *ω is a cardinal and every $n \in \omega$ is a cardinal.*

Proof. Otherwise, $|\omega| < \omega$ and therefore $|\omega| = n$ so there $|\omega| < n+1 < \omega$, but then $|n+1| = |\omega| = n$, contradicting the $n \not\sim n+1$. □

So we now have sets which are not countable. But what about uncountable sets? the problem is that $P(\omega)$ might not admit a well order.

Theorem 8.42. *For every ordinal α there is a cardinal κ such that $\alpha < \kappa$.*

Proof. Suppose otherwise, that there is an ordinal α such that for every cardinal κ is at most α. In particular, for every ordinal $\beta \geq \alpha$, $\beta \sim \alpha$. Let $S = \{R \in P(\alpha \times \alpha) \mid R \text{ well-orders } \alpha\}$. S exists by the power set axiom and comprehansion. Define for each $R \in S$, $F(R) = \text{otp}(\alpha, R)$. Then by replacement the following is a set exists $\{E = F(R) \mid R \in S\}$. By our assumption, for every $\beta \geq |\alpha|$, $\beta \sim \alpha$, and we can translate the order $(\beta, \in)$ to a well order R on α such that $\text{otp}(\alpha, R) = \beta$. We conclude that $E = \{\beta \in On \mid \beta \geq |\alpha|\}$. This is a contradiction to the fact that On was already proven not to be a set (Show that the set E cannot be a set!). □

Definition 8.43. For every α, denote by α^+ the minimal cardinal $\alpha < \kappa$. a cardinal of the form α^+ is called a successor cardinal and a cardinal κ such that for every $\alpha < \kappa$, $\alpha^+ < \kappa$ is called a limit cardinal.

Definition 8.44 (The $\aleph$ hierarchy)**.** By transfinite recursion we define $\aleph_\alpha$ for every ordinal $\alpha \in On$. $\omega_0 = \aleph_0 := \omega$ $\omega_{\alpha+1} = \aleph_{\alpha+1} := \aleph_\alpha^+$ and for a limit δ, $\omega_\delta = \aleph_\delta := \sup_{\alpha<\delta} \aleph_\alpha$.

Theorem 8.45. *(1) Every $\aleph_\alpha$ is a cardinal*
(2) For every infinite cardinal κ, there is α such that $\aleph_\alpha = \kappa$.
(3) If $\alpha < \beta$ then $\aleph_\alpha < \aleph_\beta$.
(4) $\aleph_\alpha$ is limit cardinal iff α is a limit ordinal and $\aleph_\alpha$ is a successor cardinal iff α is a successor ordinal.

Proof. For (1), we go by induction of α, the base case and succesoor case are easy by the definition of $\aleph_{\alpha+1}$. For limit δ, suppose toward a contradiction that $|\aleph_\delta| < \aleph_\delta$, then by definition of sup, there is $\alpha < \delta$ such that $|\aleph_\delta| < \aleph_\alpha$. Since δ is limit, we have that $\alpha + 1 < \delta$ and therefore

$$\aleph_\alpha < \aleph_{\alpha+1} \leq \aleph_\delta$$

Which implies by previous exercises that $|\aleph_\alpha| = |\aleph_\delta| < \aleph_\alpha$, contradicting the fact that $\aleph_\alpha$ is a cardinal by the induction hypothesis. As for (2), let κ be a cardinal and let $\delta = \sup\{\gamma \mid \aleph_\gamma \leq \kappa\}$. We claim that $\aleph_\delta = \kappa$. Let us split into cases: if $\delta = \max(\{\gamma \mid \aleph_\gamma \leq \kappa\})$, then $\aleph_\delta \leq \kappa$ and by maximality $\aleph_{\delta+1} = \aleph_\delta^+ > \kappa$. It follows that $\kappa = |\kappa| = \aleph_\delta$. If δ is limit, then again, since $\aleph_\delta = \sup_{\alpha<\delta} \aleph_\alpha$, it follows that $\aleph_\delta \leq \kappa$. It follows again that $\aleph_\delta^+ > \kappa$ and thus $\aleph_\delta = |\kappa| = \kappa$. (3) and (4) are left as exercises. □

9. Zörn's Lemma

Definition 9.1. Let $\langle \Sigma, \leq \rangle$ be a partially ordered set. A chain in Σ is a subset $X \subseteq \Sigma$ such that every $x, y \in X$ are comparable in $\leq$.

Example 9.2. $\{\{0, ..., n\} \mid n \in \mathbb{N}\}$ is a chain in $\langle P(\mathbb{N}), \subseteq \rangle$.

Definition 9.3. Let $\langle \Sigma, \leq \rangle$ be a partially ordered set. A maximal element is Σ is some $\sigma \in \Sigma$ such that there is no $x \in \Sigma$ such that $\sigma < x$.

Example 9.4. On $\mathbb{N} \setminus \{0, 1\}$ define the order $n \prec m$ iff m divides n. Then maximal elements are exactly prime numbers.

Theorem 9.5 ((AC) Zörn's Lemma)**.** *Suppose that $\langle \Sigma, \leq \rangle$ is an ordered set such that:*
(1) $\Sigma \neq \emptyset$.
(2) Every chain in Σ has an upper bound in Σ.
Then Σ has a maximal element.

Theorem 9.6. *Zorn's lemma implies the axiom of choice*

Proof. Let $\mathcal{F}$ be a set of non-empty sets. Define

$$\Sigma = \{f \in P(\mathcal{F} \times \bigcup \mathcal{F}) \mid f \text{ is a choice function on some } X \subseteq \mathcal{F}\}$$

We order Σ by inclusion.

Problem 20. *Prove that for every $f, g \in \Sigma$, $f \subseteq g$ iff* $\mathrm{dom}(f) \subseteq \mathrm{dom}(g)$ *and* $g \restriction \mathrm{dom}(f) = f$.

Let us prove the Σ satisfies the assumptions of Zörn's Lemma. Indeed $\emptyset \in \Sigma$ since it is a choice function on the empty set, hence $\Sigma \neq \emptyset$. Let $C \subseteq \Sigma$ be a chain. We claim that $F := \bigcup C = \bigcup_{f \in C} f$ is an upper bound in Σ. Clearly, F includes f for every $f \in C$ (by definition). Let us prove that $F \in \Sigma$.

???

□

Corollary 9.7. *(AC) Every vector space has a base.*

Proof. □

Theorem 9.8 (Blass)**.** *If every vector space has a base then AC holds.*

Proof that AC implies Zorn's Lemma. □

10. Cardinal Arithmetics

Theorem 10.1. *For every cardinal κ, $\kappa \cdot \kappa = \kappa$.*

Notation 10.2. $\kappa^{<\lambda} = \sup_{\delta<\lambda} \kappa^{\delta}$.

Corollary 10.3 (AC)**.** *If κ, λ are infinite then:*

(1) $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$.
(2) Suppose that for every $\alpha < \kappa$, X_α is a set such that $|X_\alpha| \leq \kappa$. Then $|\cup_{\alpha<\kappa} X_\alpha| \leq \kappa$
(3) For $\delta \leq \kappa$, $\kappa^{\delta} = |[\kappa]^{\delta}|$ where $\kappa^{\delta} = \{X \in P(\kappa) \mid |X| = \delta\}$.
(4) $\kappa^{<\omega} = \kappa$.

Proof. For (2), for each $\alpha < \kappa$ choose a function $f_\alpha : \kappa \to X_\alpha$ which is onto. Then define a function $f : \kappa \times \kappa \to \cup_{\alpha<\kappa} X_\alpha$ by $f(\alpha, \beta) = f_\alpha(\beta)$. Then f is onto and therefore $|\cup_{\alpha<\kappa} X_\alpha| \leq \kappa \cdot \kappa = \kappa$. For (3), The function $F(f) = Im(f)$ is an onto fnction from ${}^{\delta}\kappa$ to $[\kappa]^{\delta}$. For the other direction, ${}^{\delta}\kappa \subseteq \{R \in P(\kappa \times \delta) \mid |R| = \delta\}$. Since $|P(\kappa \times \delta)| = |P(\kappa)$ we get that $|{}^{\delta}\kappa| \leq |[\kappa]^{\delta}|$. For (4), note that $\kappa^n = \kappa$ for every $n \geq 1$ (by induction and since $\kappa \cdot \kappa = \kappa$) and therefore $\kappa^{<\omega} = \sup_{n<\omega} \kappa^n = \kappa$ □

It follows that $\kappa^{<\delta} = |[\kappa]^{<\delta}|$ where $[A]^{<\delta} = \{B \subseteq A \mid |B| < \delta\}$. Also from (1) we see that only the exponent operation is left unsettled. As we will see later, ZFC cannot determine theses values. However, there are some cases which are settled, in the rest of this chapter we investigate what restrictions ZFC pose one these values:

Theorem 10.4. *If $\lambda \geq \omega$ and $2 \leq \kappa \leq \lambda$ then*

$$\kappa^\lambda = 2^\lambda$$

Proof. $2^\lambda \leq \kappa^\lambda \leq (2^\kappa)^\lambda = 2^{\kappa\cdot\lambda} = 2^\lambda$. □

In case $\lambda < \kappa$ we can say a bit more about κ^λ but we need the following definition:

Definition 10.5. Let α be an **ordinal**. We define $cf(\alpha)$ to be the minimal γ such that there is an cofinal/unbounded function $f : \gamma \to \alpha$[14].

Example 10.6. $cf(\omega) = \omega$, $cf(\omega_1) = \omega_1$ (since if $\alpha < \omega_1$ and $f : \alpha \to \omega_1$, we have that $\sup(f)$ is a countable union of countable sets so $|\sup(f)| = \omega$. It follows that $\sup(f) < \omega_1$.

Remark 10.7. (1) $cf(\alpha) \leq \alpha$.
(2) $cf(\alpha + 1) = 1$.
(3) there is always $f : cf(\alpha) \to \alpha$ which is cofinal and strictly increasing.

Exercise 13. *If α is a limit ordinal and $f : \alpha \to \beta$ is cofinal and strictly increasing then $cf(\alpha) = cf(\beta)$.*

Exercise 14. *For every limit ordinal α, $cf(\aleph_\alpha) = \alpha$.*

Corollary 10.8. *$cf(cf(\beta)) = cf(\beta)$.*

Definition 10.9. a limit ordinal κ called *regular* if $cf(\kappa) = \kappa$, otherwise it is called *singular*.

Corollary 10.10. *If κ is regular then κ is a cardinal.*

Example 10.11. ω is regular and ω_1 is regular. $cf(\aleph_\omega) = \omega < \aleph_\omega$ is singular.

Theorem 10.12 (AC)**.** *For every κ, κ^+ is regular.*

Proof. Otherwise, there is a function $f : \lambda \to \kappa^+$ for some $\lambda \leq \kappa$. For every $\alpha < \lambda$, let $X_\alpha = f(\alpha)$, then $|X_\alpha| \leq \kappa$ and therefore $|\kappa^+| = |\cup_{\alpha<\lambda} X_\alpha| \leq \kappa$, contradiction. □

Is there a limit regular cardinal greater than $\aleph_0$?

Definition 10.13. A cardinal κ is called
(1) Weakly inaccessible if it regular and a limit cardinal.
(2) Strongly inaccessible if it is regular and

$$\forall\lambda < \kappa . 2^\lambda < \kappa$$

weakly and strongly inaccessible cardinals are so-called "large cardinals", these are cardinals which ZFC cannot prove their existence.

Lemma 10.14 (Konig's Lemma)**.** *Let κ be an infinite cardinal, and assume that $cf(\kappa) \leq \lambda$, then $\kappa^\lambda > \kappa$*

[14] $f : \gamma \to \alpha$ is cofinal/unbounded if $Im(f)$ is inbounded in α.

Proof. Let $f : \lambda \to \kappa$ be cofinal. Suppose toward a contradiction that there is $G : \kappa \to {}^{\lambda}\kappa$ which is onto. Define $g : \lambda \to \kappa$ by

$$g(\alpha) = \min(\kappa \setminus \{G(\mu)(\alpha) \mid \mu < f(\alpha)\})$$

To see that $g \notin Im(G)$, let $\rho < \kappa$ then there is $\beta < \lambda$ such that $\rho < f(\beta)$. Hence $g(\beta) \notin \{G(\mu)(\beta) \mid \mu < f(\beta)\}$ and in particular $g(\beta) \neq G(\rho)(\beta)$, hence $g \neq G(\rho)$. This is a contradiction to the fact he G is onto. □

Corollary 10.15. *For any infinite cardinal κ, $cf(2^\kappa) > \kappa$.*

Proof. Note that $(2^\kappa)^\kappa = \kappa$, hence by the contrapositive of Konig's lemma, we get $cf(2^\kappa) > \kappa$. □

10.1. **The continuum function.** The function $\alpha \mapsto 2^{\aleph_\alpha}$ is called the continuum function and as we will see, its values are highly undetermined by ZFC.

Definition 10.16 (AC)**.** The Generalized Continuum Hypothesis (GCH) is the statement that for every α, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

Under GCH, all the values of κ^λ (and therefore the continuum function) can easily be computed,

Theorem 10.17 (AC+GCH)**.** *Let λ, κ be infinite cardinals. Then:*
(1) If $\lambda \geq \kappa$, then $\kappa^\lambda = \lambda^+$.
(2) If $cf(\kappa) \leq \lambda < \kappa$ then $\kappa^\lambda = \kappa^+$.
(3) If $\lambda < cf(\lambda)$ then $\kappa^\lambda = \kappa$

Proof. It remains to prove 3, so $\kappa \leq \kappa^\lambda = \sup_{\delta<\kappa} \delta^\lambda \leq \sup_{\delta<\kappa} \delta^+ = \kappa$ □

Let us define the *beth function*:

Definition 10.18. $\beth_0 = \aleph_0$, $\beth_{\alpha+1} = 2^{\beth_\alpha}$ and for limit δ, $\beth_\delta = \sup_{\alpha<\delta} \beth_\alpha$.

Exercise 15. *GCH is equivalent to the statement that for every α, $\beth_\alpha = \aleph_\alpha$.*

To summarize what we know about the continuum function, we have the following theorem:

Theorem 10.19. *(1) $\kappa < \lambda \Rightarrow 2^\kappa \leq 2^\lambda$. (Monotonicity)*
(2) $cf(2^\kappa) > \kappa$. (Konig's lemma)
(3) If κ is limit then $2^\kappa = (2^{<\kappa})^{cf(\kappa)}$.

Proof. We need to prove (3), $\kappa = \sup_{i<cf(\kappa)} \kappa_i$. So the map

$$X \subseteq \kappa \mapsto \langle X \cap \kappa_i \mid i < cf(\kappa) \rangle$$

is a $1-1$ function from $P(\kappa)$ to ${}^{cf(\kappa)}([\kappa]^{<\kappa})$. Hence

$$2^\kappa \leq (2^{<\kappa})^{cf(\kappa)} \leq (2^\kappa)^{cf(\kappa)} = 2^{\kappa \cdot cf(\kappa)} = 2^\kappa$$

□

In case κ is regular, (3) is not very interesting and as we will see, constrains $(1),(2)$ are the only limitations ZFC pose on the continuum function in ZFC. However, (3), suggests that for singular cardinals the situation is very different and depends heavily on thecontinuum function restricted to cardinals below it and on the exponent values. For example we have the following corollary:

Corollary 10.20. *If κ is singular, and the continuum function is eventually constant below κ with value λ, then $2^\kappa = \lambda$.*

Definition 10.21. A cardinal κ is strong limit if $\forall \nu < \kappa . 2^\nu < \kappa$.

Note that a strong limit cardinal is in particular a limit cardinal.

Exercise 16. *(1) Prove that there is a strong limit cardinal and that the least such carinal is of cofinality ω.*
(2) Prove that if κ is strong limit then:

$$\forall \nu, \lambda < \kappa . \lambda^\nu < \kappa$$

(3) If κ is strong limit then $2^\kappa = \kappa^{cf(\kappa)}$

Definition 10.22. The *Singular Cardinal Hypothesis* is the statement:

$$\text{For every strong limit singular cardinal } \kappa,\ 2^\kappa = \kappa^+$$

There is another formulation which implies the above, which involves all singular cardinals:

$$\text{For every singular cardinal } \kappa,\ 2^{cf(\kappa)} < \kappa \Rightarrow \kappa^{cf(\kappa)} = \kappa^+$$

We will leave it as an exercise to prove that the second formulation determines the continuum function for all singular cardinals. While the second version implies the first, it is known that the two formulations are not equivalent.

11. Appendix

11.1. **Induction and Recursion.**
Induction and recursion and extremely related techniques, however, they have totally different purposes:

Important: *Induction* is a proof technique while *Recursion* is a definition technique.

11.2. **Recursion.**
As we said, recursion is a definition technique, but what does it define? sequences:

Definition 11.1. A *sequence* of elements of a set A is an list of elements of A enumerated by the natural numbers.[15]

Example 11.2. The following are examples of sequences:
(1) The sequence $a_n = n$ is the sequence $0, 1, 2, 3, 4, ...$

[15] The real definition of a sequence involves the concept of functions which we will study later.

(2) The sequence $b_n = \frac{1}{n+1}$ is the sequence $1, \frac{1}{2}, \frac{1}{3}, ...$
(3) The sequence $c_n = (-1)^n$ is the sequence $1, -1, 1, -1, 1, ...$
(4) The sequence d_n of the sum of angles in degrees of a polygon with $n+3$ vertexes, is the sequence $180^o, 360^o, 540^o, ...$ and actually $d_n = (n+1) \cdot 180^o$.

Definition 11.3. A *recursive* definition of a sequence has two parts:
(1) Initial values of the sequence: A definition of the first few values of the sequence.
(2) The recursive condition: A formula to compute the next element in the sequence from the previous elements.

Remark 11.4. The number of previous elements required to define the next element is called the *depth* of the recursion. The depth of the recursion determined how many initial values should we specify.

Example 11.5. (1) $a_0 = 0$, $a_{n+1} = a_n + 1$, the depth is 1.
(2) An arithmetic sequence is a sequence of the form $a_0 = a$ and $a_{n+1} = a_n + d$, for some given a, d. For example: $a_0 = 5$ and $a_{n+1} = a_n - 7$.
(3) A geometric sequence is a sequence of the form $a_0 = a$ and $a_{n+1} = a_n \cdot q$ for some given a, q for example $a_0 = 5$ and $a_{n+1} = a_n \cdot (-7)$.
(4) $a_0 = a_1 = 1$ and $a_{n+1} = a_n + a_{n-1}$. Here the depth is 2. This is called the Fibonacci sequence.
(5) $0! = 1$ and $(n+1)! = n! \cdot (n+1)$.
(6) $a_1 = \emptyset$ and $a_{n+1} = \{a_n\}$. We are allowed to start the enumeration from a natural number grater than 0.

11.3. **induction.** One of the most common techniques for proving **Universal statements** of the form $\forall n \in \mathbb{N}...$ is a proof by induction. Let us explain our goal and the idea behind induction.

Suppose we would like to prove a claim of the form

"For every **natural number** n, q(n) (some property of n)"

This is extremely important that the statement speaks about natural numbers. In order to prove such statement, we can use a proof by induction. The point is to prove an infinite chain of implications:

$$q(0) \Rightarrow q(1) \Rightarrow q(2) \Rightarrow ...q(n) \Rightarrow q(n+1) \Rightarrow ...$$

This is done by proving **for a general** n that $q(n) \Rightarrow q(n+1)$, this is called *the inductive step*. Then the final step is to prove $q(0)$ which is called the *base of the induction*. If we proved both the base of the induction and the induction step then we can now derive the property for every natural number since:

- $q(0)$ is true by the base.
- $q(0) \Rightarrow q(1)$, then $q(1)$ is true.
- $q(1) \Rightarrow q(2)$, then $q(2)$ is true, and so on.

Practically, since $q(n) \Rightarrow q(n+1)$ is a universal implication, a *proof by induction* for the claim $\forall n \in \mathbb{N}.q(n)$ has the following structure:

(1) The base of the induction: Proof for $q(0)$.
(2) Induction hypothesis: "Suppose that $q(n)$ holds", here n is a general variable.
(3) Induction step: We need to prove that $q(n+1)$ holds, under the given induction assumption that $q(n)$ holds.

Example 11.6. Prove by induction the following claims:

(1) $\forall n \in \mathbb{N}.n^2 \geq n$.
Proof. **The induction base:** We need to prove that for $n = 0$, $0^2 \geq 0$, this is indeed true since $0^2 = 0$.
The induction hypothesis(Abbreviated I.H.): Let n be any natural number, and suppose that $n^2 \geq n$.
The induction step: We need to prove that $(n+1)^2 \geq n+1$. Indeed,
$$(n+1)^2 = n^2 + 2n + 1 \underset{\text{Since } n \geq 0}{\geq} n^2 + 1 \underset{I.H.}{\geq} n + 1$$
□

(2) $\forall n \geq 1(n+1 \leq 2n)$.[16]
Proof. **The induction base:** We need to prove the claim for $n = 1$. Indeed,
$$1 + 1 = 2 \leq 2 = 2 \cdot 1$$
The induction Hypothesis: Suppose that for a general $n \geq 1$, $n+1 \leq 2n$.
The induction step: We need to prove that $(n+1)+1 \leq 2(n+1)$. Indeed,
$$(n+1) + 1 \underset{I.H}{\leq} 2n + 1 \leq 2n + 2 = 2 \cdot (n+1)$$
□

(3) $\forall n > 3.2^n < n!$.
Proof. **The induction base:** We need to prove the claim for $n = 4$, indeed $2^4 = 16 \leq 24 = 4!$.
The induction hypothesis: Suppose that for a general $n > 3$, $2^n < n!$.
The induction step: We need to prove that $2^{n+1} < (n+1)!$. Indeed,
$$2^{n+1} = 2^n \cdot 2 \underset{I.H.}{\leq} n! \cdot 2 \underset{\text{Since } n>3}{\leq} n! \cdot (n+1) \underset{\text{Recursive def}}{=} (n+1)!$$
□

(4) A general term for an arithmetic sequence. Suppose that $a_n = a_{n-1} + d$ is an arithmetic sequence. Then for every $n \in \mathbb{N}$, $a_n = a_0 + d \cdot n$. (homework: geometric sequence and sum of squares)
Proof. **The induction base:** For $n = 0$, we need to prove that $a_0 = a_0 + d \cdot 0$. This is clearly true.

[16] We can start the induction from a natural number greater than 0, this only changes the base of the induction.

The induction hypothesis: Suppose that for a general n, $a_n = a_0 + dn$.
The induction step: We need to prove that $a_{n+1} = a_0 + d(n+1)$. Using the recursive definition of a_n, we have that:

$$a_{n+1} = a_n + d \underset{I.H.}{=} a_0 + dn + d = a_0 + d(n+1)$$

□

(5) The partial sum of a arithmetic sequence. Suppose that $a_n = a_{n-1} + d$ is an arithmetic sequence. Then for every $N \in \mathbb{N}$,

$$\sum_{i=0}^{N} a_i = a_0 + a_1 + ... + a_N = (N+1)(a_0 + dN/2)$$

Proof. **The induction base:** We need to prove the formula for $N = 0$, $a_0 = (0+1)(a_0 + d \cdot 0/2)$. This is clear.
The induction hypothesis: Suppose that the formula is true for a general N, namely, we assume that truth of the equality

$$\sum_{i=0}^{N} a_i = a_0 + a_1 + ... + a_N = (N+1)(a_0 + dN/2)$$

The induction step:

$$\sum_{i=0}^{N+1} a_i = \underbrace{a_0 + ... + a_N}_{\sum_{i=0}^{N} a_i} + a_{N+1} \underset{I.H.}{=} (N+1)(a_0 + dN/2) + a_{N+1} \underset{\text{Previous exercise}}{=}$$

$$= (N+1)(a_0 + dN/2) + a_0 + d(N+1) = (N+2)a_n + (N+1)d(N/2+1) =$$
$$(N+2)a_0 + (N+2)d(N+1)/2 = (N+2)(a_0 + d(N+1)/2)$$

□

For example, consider the arithmetic sequence $a_n = n$ (here $a_0 = 0$ and $d = 1$) then we can apply the formula to conclude that

$$0 + 1 + 2 + ... + 1000 = 1001(0 + 1 \cdot 1000/2) = 1001 \cdot 500 = 500,500$$

(6) Prove that for any given n lines in the plane, no two are parallel, and no three intersect at a single point[17], have exactly $\frac{n(n-1)}{2}$ points of intersection. (homework: the sum of angles of a polygon)
Proof. Let d_n denote the number of intersection points of n non-concurrent lines. We firs construct a recursive formula for d_n. Clearly, $d_1 = 0$ (and $d_2 = 1$, $d_3 = 3$). Given n non-concurrent lines, they have d_n intersection points. Adjoining a new line to them, it intersect each of the lines exactly once (since it is not parallel to any of them) and the points of intersection are different since no three lines intersect at a point. Hence

$$d_{n+1} = \underbrace{d_n}_{\text{The intesection points of the old lines}} + \underbrace{n}_{\text{the intersections with the new line}}$$

[17]Such lines are called *non-concurrent* lines.

Now let us prove by induction that $d_n = \frac{n(n-1)}{2}$.

The induction base: Indeed $d_1 = 0 = \frac{0\cdot(-1)}{2}$.

The induction hypothesis: Suppose that for a general n, $d_n = \frac{n(n-1)}{2}$.

The induction step: We need to prove that $d_n = \frac{(n+1)n}{2}$. We use the recursive description of d_{n+1},

$$d_{n+1} = d_n + n = \frac{n(n-1)}{2} + n = n(\frac{n-1}{2} + 1) = n\frac{n-1+2}{2} = \frac{n(n+1)}{2}$$

□

(7) Define the recursive sequence $a_0 = \emptyset$, $a_{n+1} = P(a_n)$. Then for every $n \in \mathbb{N}$, $a_n \subseteq a_{n+1}$.

Proof. **The induction base:** For $n = 0$ we need to prove that $a_0 \subseteq a_1$. By definition $a_0 = \emptyset$, and we have already prove that the empty set is included in every set. In particular $a_0 = \emptyset \subseteq a_1$.

The induction hypothesis: Suppose that for a general n, $a_n \subseteq a_{n+1}$.

The induction step: We need to prove that $a_{n+1} \subseteq a_{n+2}$. This is an inclusion proof, so let $X \in a_{n+1}$. We need to prove that $X \in a_{n+2}$. By definition, $a_{n+1} = P(a_n)$, and by the assumption, $X \in P(a_n)$. By definition of the power set, $X \subseteq a_n$. By the induction hypothesis, $a_n \subseteq a_{n+1}$. We already saw that if $a \subseteq b \wedge b \subseteq c$ then $a \subseteq c$. It our case, we conclude that $X \subseteq a_{n+1}$. Again by the definition of the power set, $X \in P(a_{n+1}) = a_{n+2}$, as wanted.

□