MA1100T Basic Discrete Mathematics(T)

(The Torturous Version of MA1100)

Based on lectures by Prof. Goh Jun Le
Notes taken by Zhao Kejun

AY23/24 Semester 1

This set of notes is based on the MA1100T lectures delivered by Prof.Goh Jun Le in AY23/24 Semester 1. The definitions, results and proofs in the notes largely follow those in the lecture slides. However, I have modified them slightly and added some extra proofs, explanations and remarks. Therefore, what was written in this notes is by no means an accurate representation of what was actually lectured, and in particular, all errors are almost surely mine.

Chapter 1 Logic

Chapter 2 Proofs

Chapter 3 Induction

(*The first 3 chapters were omitted mainly because I was too lazy to type them out. In fact, if you have some experience in SG-Cambridge A level H3 Mathematics or equivalent, there is no barrier for you to read and understand the rest of the notes (chapters 4-9). That being said, chapters 1-3 are still important because they set the foundation for this course and provide useful insights into logic and proofs.*)

Chapter 4 Sets

4.1 Naive Set Theory

4.2 Axiomatic Set Theory (Zermelo-Fraenkel Axioms)

Axiom 4.2.1. (*The Empty Set*) There is an empty set, i.e., $\exists x \forall y (y \notin x)$.

Remarks. The empty set is a subset of every set, but is not an element of every set.

Axiom 4.2.2. (*Extensionality*) Two sets are equal if and only if they have the same elements, i.e., $A = B$ is an abbreviation of $\forall x (x \in A \leftrightarrow x \in B)$.

Remarks. Extensionality dictates that there is a unique empty set.

Axiom 4.2.3. (*Pairing*) For any sets x and y, there is a set whose elements are exactly x and y, i.e., $\forall x \forall y \exists z \forall w (w \in z \leftrightarrow (w = y \vee w = x))$

Remarks. $\{x, y\} = \{x\}$ if $x = y$. This axiom thus allows us to construct sets with a single element or two elements. Now, we are able to construct sets such as $\{\emptyset\}$, $\big\{\{\emptyset\}\big\}$, $\Big\{\big\{\{\emptyset\}\big\}, \{\emptyset\}\Big\}$, $\big\{\{\emptyset\}, \emptyset\big\}$ and so on.

Axiom 4.2.4. (*Unary Union*) For each set A, there is a set whose elements are exactly the elements of elements of A, i.e. $\forall A \forall a \forall x \exists W (a \in A \wedge x \in a \rightarrow x \in W)$.

Example. $\bigcup\{A, B\} = \big\{$elements of element of {A,B} $\big\} = \big\{$element of A or element of B $\big\} = A \bigcup B$

Remarks.

1. From the above example, we see that if the set A in the axiom is a finite set, taking the unary union of this set A is just applying the binary union finitely many times.
2. Certainly not all sets are finite. So imbedded in this axiom is the idea of arbitrary union.
3. This axiom, together with 1 and 3, allows us to construct sets with more than 2 elements. For example, one can construct a set with three elements as follows. By axiom 1 and 3, we are able to construct the sets $\emptyset$, $\{\emptyset\}$, $\big\{\{\emptyset\}\big\}$, $\Big\{\big\{\{\emptyset\}\big\}\Big\}$and $\big\{\{\emptyset\}, \emptyset\big\}$, as well as $\bigg\{\big\{\{\emptyset\}, \emptyset\big\}, \Big\{\big\{\{\emptyset\}\big\}\Big\}\bigg\}$ by pairing the last two sets. Then take the unary union of $\bigg\{\big\{\{\emptyset\}, \emptyset\big\}, \Big\{\big\{\{\emptyset\}\big\}\Big\}\bigg\}$, we have the set $\Big\{\emptyset, \{\emptyset\}, \big\{\{\emptyset\}\big\}\Big\}$.
4. The union of the empty set is the empty set.

Axiom 4.2.5. (*Separation*) For each set A and each formula $\phi(x)$, there is a set whose elements are exactly those x in A such that $\phi(x)$ holds.

Remarks.

1. This axiom is not formally stated as we have not defined what a formula is.
2. One cannot allow "arbitrary" formulas or regard a formula as a"variable". Therefore, we cannot quantify ϕ.
3. The above axiom is not of the form $\forall A \forall \phi$... Rather, for each literal formula ϕ, there is an axiom $\forall A \exists B \forall x (x \in B \leftrightarrow (x \in A \wedge \text{a literal } \phi \text{ holds})$. So there are in fact infinitely many axioms, which we call an axiom scheme.

4. Now we can show that the (unary) intersection of a **nonempty** set A exists. Proof: Fix $c \in A$. Then we apply the separation axiom to obtain the set $\{x \in c : (\forall a \in A)(x \in a)\}$. This is essentially the set that contains the sets which are common elements of every element of A.

5. One cannot take all sets with a property and take the intersection. This could lead to paradoxes.

6. We can define the ordered pair (x, y) to be $\big\{\{x\}, \{x, y\}\big\}$ and prove that $\big\{\{x\}, \{x, y\}\big\} = \big\{\{a\}, \{a, b\}\big\}$ if and only if $a = x$ and $b = y$.

   Proof. Suppose $A = \big\{\{x\}, \{x, y\}\big\} = \big\{\{a\}, \{a, b\}\big\} = B$. By extensionality, we have $\{x\} \in A \leftrightarrow \{x\} \in B$ and $\{x, y\} \in A \leftrightarrow \{x, y\} \in B$, i.e. $\{x\}, \{x, y\}$ are elements of B. By extensionality, $\{x\} = \{a\}$ and $\{x, y\} = \{a, b\}$. It follows that $a = x$ and $b = y$. Conversely, if $a = x$ and $b = y$, it is obvious that , $\{x\} = \{a\}$ and $\{x, y\} = \{a, b\}$, and $\big\{\{x\}, \{x, y\}\big\} = \big\{\{a\}, \{a, b\}\big\}$ by extensionality. □

Axiom 4.2.6. (*The Power Set*) For each set A, there is a set whose elements are exactly the subsets of A. $\forall A \exists B \forall a \forall x \big((x \in a \to x \in A) \leftrightarrow a \in B\big)$.

Remarks.

1. For all sets A and B, each ordered pair (a, b) is an element of $\mathcal{P}(\mathcal{P}(A \cup B))$.

   Proof. Since a, b are elements of $A \cup B$, $\{a\}, \{a, b\}$ are elements of $\mathcal{P}(A \cup B)$. Hence, $\{\{a\}, \{a, b\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$ □

2. Therefore, we can define the Cartesian product $A \times B$ to be

$$\{p \in \mathcal{P}(\mathcal{P}(A \cup B)) : (\exists a \in A)(\exists b \in B)(p = (a, b))\}$$

Axiom 4.2.7. (*Infinity*) There is a set X such that $\emptyset \in X$ and for any $x \in X$, we have $x \cup \{x\} \in X$. $\exists X\big(\emptyset \in X \wedge \forall x(x \in X \to x \cup \{x\} \in X)\big)$.

Remarks.

1. The set of natural numbers $\mathbb{N}$ is defined to be the intersection of all sets X as described above. Formally, we can define $\mathbb{N}$ by applying separation and infinity as follows. Fix $x \in X$. (The set X is guarenteed to be nonempty by infinity.) Then apply separation to the set X to obtain $\Big\{x \in X : \forall Y\Big(\big(\emptyset \in Y \wedge \forall y \in Y(y \cup \{y\} \in Y)\big) \to x \in Y\Big)\Big\}$.

2. We can now state the **principle of induction on** $\mathbb{N}$ as follows:
   If $\phi(n)$ is a formula such that $\phi(\emptyset)$ holds, and for all $n \in \mathbb{N}$, $\phi(n) \Rightarrow \phi(n \cup \{n\})$, then $\phi(n)$ holds for all $n \in \mathbb{N}$.

Proof. Consider the set $\{n \in \mathbb{N}\colon \phi(n)\}$. By the assumptions of induction, we have $\emptyset \in \{n \in \mathbb{N}\colon \phi(n)\}$ and for all $n \in \mathbb{N}$, $n \cup \{n\} \in \{n \in \mathbb{N}\colon \phi(n)\}$. Therefore, the set $\{n \in \mathbb{N}\colon \phi(n)\}$ satisfies the conditions for X above. Since $\mathbb{N}$ is the intersection of all sets X, we have $\mathbb{N} \subseteq \{n \in \mathbb{N} : \phi(n)\}$, i.e. $\phi(n)$ holds for all $n \in \mathbb{N}$. □

3. Now we can construct natural numbers: $0 = \emptyset$; $1 = \emptyset \cup \{\emptyset\} = \{\emptyset\}$; $2 = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\}$; $3 = \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$......

4. Laws of arithmetic (addition, multiplication, associativity, distributive law, etc) can be proven using the ZFC axioms.

Axiom 4.2.8. (*Replacement*) For each formula $\phi(x, y)$ and each set A, if for every $x \in A$, there is a unique y such that $\phi(x, y)$ holds, then there is a set whose elements are exactly those y for which there is some $x \in A$ with $\phi(x, y)$, i.e. $\{y : (\exists x \in A)\phi(x, y)\}$ exists.

Remarks.

1. Intuitively, one can think of ϕ as a function with domain A.
2. Replacement justifies the alternate set builder notation: $\{f(x) : x \in A\}$.
3. In axiomatic set theory, saying ϕ is a function presumes that it is a set, which need not to be true. In other words, replacement is much stronger than simply saying that " every function has a range"
4. Replacement does not follow from separation, because in separation, y needs to be in some set in the first place. However, using replacement one can prove separation.

Axiom 4.2.9. (*Foundation*) Every nonempty set A has an element x such that for every $a \in A$, we have $a \notin x$.

Remarks. In particular, there is no x such that $x \in x$ (consider $A = \{x\}$). Whenever the set contains the empty set, foundation is trivially true.

4.3 Relations, Well Orders and Generalised Induction

Definition 4.3.1. (*Binary Relation*) A relation on A is a set which contains only ordered pairs with both coordinates from A. A binary relation on set A is a subset of $A \times A$.

Definition 4.3.2. (*Well Order*) A binary relation $<$ on A is a well-order if:

1. it is transitive, i.e., $x < y$ and $y < z$ implies $x < z$.
2. it satisfies trichotomy, i.e., for every $x, y \in A$, exactly one of $x = y$, $x < y$ and $x > y$ holds

3. every nonempty subset of A has <-least element, i.e.

$$(\forall X \subseteq A)(X = \emptyset \vee (\exists x \in X)(\forall y \in X)(y \not< x))$$

Remarks. This relation "<" should not be interpreted as the usual <. For example, one can prove that a relation < is a well order on the set of negative integers by defining $-1 < -2 < -3 < ....$ Furthermore, one can prove induction on any well-ordered sets. But in fact induction holds for all well-founded sets.

Definition 4.3.3. (*Well-founded Relations*) A relation R on A is well founded if

$$(\forall X \subseteq A)(X = \emptyset \vee (\exists x \in X)(\forall y \in X)((y, x) \notin R))$$

Theorem 4.3.4. (*Generalised Induction*) Suppose a relation R on A is well founded, and $\phi(x)$ is a formula such that for all $y \in A$, if $\phi(x)$ holds for all x such that $(x, y) \in R$, then $\phi(y)$ holds. Then $\phi(x)$ holds for all $x \in A$.

Proof. Suppose towards a contradiction that $\neg\phi(x)$ holds for some $x \in A$. Consider the set $X = \{x \in A : \neg\phi(x)\} \subseteq A$. X is non-empty. By definition of well -foundedness, there exists some $x_0 \in X$ such that for all $y \in X$, $(y, x_0) \notin R$. Then, $\phi(x)$ holds for all x such that $(x, x_0) \in R$ and $\phi(x_0)$ fails, contradicting the assumption. □

Chapter 5 Functions

5.1 What is a function?

Definition 5.1.1. Let A and B be sets. A function $f\colon A \to B$ is an assignment of exactly one element of B to each element of A.
Set theoretically, a function is an ordered triple (A, B, G) such that

1. $G \subseteq A \times B$
2. $(\forall a \in A)(\exists! b \in B)((a, b) \in G)$

Remarks. If f maps an element of the domain to zero elements or more than one element in the codomain, then f is not a function.

Example.

1. For all sets A and x, the constant function c_x on A is defined by $c_x\colon A \to \{x\}, a \mapsto x$
2. For every set A, the identity function on A is defined by $\mathrm{id}_A\colon A \to A, x \mapsto x$.
3. Given sets $A \subseteq B$, the inclusion function from A to B is defined by $f\colon A \to B, x \mapsto x$

5.2 Range of a Function

Definition 5.2.1. The range of a function $f\colon A \to B$ is the set

$$\mathrm{range}(f) = \{b \in B\colon (\exists a \in A)(f(a) = b)\} = \{f(a)\colon a \in A\}$$

Remarks. Because of separation, $\mathrm{range}(f) = \{b \in B\colon (\exists a \in A)((a, b) \in G_f)\}$.
In fact, by replacement, even if f is not a set but is rather a ”rule” defined by some formula ϕ, the range of ϕ on a set is still a set.

5.3 Surjection, Injection, Bijection and Inverse

Definition 5.3.1. A function $f\colon A \to B$ is surjective if

$$(\forall b \in B)(\exists a \in A)(f(a) = b)$$

Definition 5.3.2. A funciton $f\colon A \to B$ is injective if

$$(\forall x, y \in A)(f(x) = f(y) \to x = y)$$

Definition 5.3.3. A function $f\colon A \to B$ is bijective iff it is both surjective and injective.

Remarks.

1. If $f\colon A \to B$ is bijective, we can define the **inverse** of f, denoted $f^{-1}\colon B \to A$, by $f^{-1}(b) = a$ whenever $f(a) = b$.
2. Note that if $f\colon A \to B$ is only injective, the inverse of f does not exist because not all elements in B is assigned a unique element in A.

5.4 Composition of Functions

Definition 5.4.1. Let A, B, C be sets such that $B \subseteq C$. Suppose $f\colon A \to B$ and $g\colon B \to C$. The composition of g and f is a function $g \circ f\colon A \to D$ defined by $g \circ f(a) = g(f(a))$

Remarks. Set theoretically:

$$G_{g \circ f} = \{(a, d) \in A \times D\colon (\exists b \in B)(f(a)) = b \wedge g(b) = d)\}$$

Saying that the composition of some functions equals some value can be thought of as an existential statement.

Example. Let X be a nonempty set. Fix an element $x_0 \in X$. Define

$$f\colon \mathcal{P}(X) \to \mathcal{P}(X) \quad \text{by} \quad f(A) = A - \{x_0\}$$

and

$$g\colon \mathcal{P}(X) \to \mathcal{P}(X) \quad \text{by} \quad g(A) = A \cup \{x_0\}$$

Then $g \circ f = f$ and $f \circ g = f$. This example shows that the composition operation does not allow us to "cancel" a function on both sides and conclude that the remaining one is the identity.

Proposition 5.4.2. For every function $f\colon X \to Y$, $f \circ \mathrm{id}_X$ and $\mathrm{id}_Y \circ f$ are both equal to f.

Proof. By definition, $f \circ \mathrm{id}_X$ is a function from X to Y.So is f. Furthermore, for each $x \in X$, $f(\mathrm{id}_X) = f(x)$. By definition, $\mathrm{id}_Y \circ f$ is a function from X to Y. So is f. Furthermore, for each $x \in X$, $\mathrm{id}_X(f(x)) = f(x)$. □

Proposition 5.4.3. For every bijection $f\colon X \to Y$, $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f = \mathrm{id}_X$.

Proof. By definition, $f \circ f^{-1}$ is a function from Y to Y, and so is id_Y. Furthermore, for each $y \in Y$, there is a unique $x \in X$ such that $y = f(x)$. By definition, $x = f^{-1}(y)$. So $f(f^{-1})(y) = f(x) = y$. (Note that we can apply the function f on both sides of the because $f^{-1}(y) \in X$ is in the domain of f). The second part of the proof is completely analogous. □

Proposition 5.4.4. Suppose X is a nonempty set and $f\colon X \to Y$ is a function. Then f is injective if and only if there is a function $g\colon Y \to X$ such that $g \circ f = \mathrm{id}_X$

Proof. ($\Rightarrow$) Suppose f is injective. Pick $x_0 \in X$ Define $g\colon Y \to X$ as follows.

- For every $y \in Y$ such that $f(x) = y$ for some $x \in X$, define $g(y) = x$. This is well-defined because f being an injection guarantees that x is unique.

- For all other $y \in Y$, define $g(y) = x_0$.

One cannot simply writing " define $gf(x) = x$" because

- for every $y \in Y$ such that $y = f(x)$, there may be multiple x such that $f(x) = y$. This could result in y being mapped back to multiple x, i.e. g(f(x)) have multiple values

- there may be $y \in Y$ which are not mapped to be any $x \in X$. We need to define g for such y as well.

($\Leftarrow$) Suppose there is a function $g\colon Y \to X$ such that $g \circ f = \mathrm{id}_X$. Pick $x_1, x_2 \in X$. Suppose $f(x_1) = f(x_2)$. Apply the function g on both sides to obtain $g \circ f(x_1) = g \circ f(x_2)$. Thus, $\mathrm{id}_X(x_1) = \mathrm{id}_X(x_2)$, i.e., $x_1 = x_2$ □

Remarks. Being one-to-one is equivalent to having a left inverse.

Proposition 5.4.5. Suppose X is a nonempty set and $f\colon X \to Y$ is a function. Then f is surjective if there is a function $g\colon Y \to X$ such that $f \circ g = \mathrm{id}_Y$

Proof. Given $y \in Y$. By assumption, there is $g(y) \in X$ such that $f(g(y)) = y$. So f is onto. □

Remarks. One may attempt to prove the forward direction as follows:

- Given $y \in Y$, we want to define $g(y)$ such that $f(g(y)) = y$. Since f is onto, there is some $x \in X$ such that $y = f(x)$. So we only need to let $x = g(y)$.

The issue here is that there may be multiple such x such that $y = f(x)$, and we need to choose exactly one such x and define it to be $g(y)$. Also, we need to do this for every $y \in Y$. It turns out that in Zermelo Fraenkel set theory, ($\Rightarrow$) is not provable. ($\Rightarrow$) is equivalent to the axiom of choice.

Proposition 5.4.6. Suppose X is a nonempty set and $f\colon X \to Y$ is a function. The following statements are equivalent:

1. f is a bijection
2. there is a function $g\colon Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.
3. there are functions $g, h\colon Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ h = \mathrm{id}_Y$.

Proof. (1)$\Rightarrow$(2): This follows from **Proposition 4.3**
(2)$\Rightarrow$(3): (3) is a weaker version of (2)
(3)$\Rightarrow$(1): This follows from **Proposition 4.4** and **Proposition 4.5** □

Remarks. To prove a function is a bijection, one can attempt to produce its left inverse g and right inverse h.

Proposition 5.4.7. *(Associativity of function composition)* Let $f\colon X \to Y$, $g\colon Y' \to Z$, $h\colon Z' \to W$ be functions, where $Y \subseteq Y'$ and $Z \subseteq Z'$. Then $h \circ (g \circ f) = (h \circ g) \circ f$.

Proof. First verify that the two functions are well defined and the domains and codomains are equal. Then check the graphs are the same. □

5.5 Restriction of Functions

Definition 5.5.1. A function $f\colon A \to Y$ is a restriction of a function $g\colon X \to Y$ if $A \subseteq X$ and for each $a \in A$, we have $f(a) = g(a)$.

Proposition 5.5.2. f is a restriction of $g\colon X \to Y$ if $f = g \circ \mathrm{id}_A$ for some $A \subseteq X$.(We denote $g \circ \mathrm{id}_A$ by $g \restriction A$)

Proof. By definition, f is a function from A to Y. Furthermore, for each $a \in A$, $f(a) = g \circ (\mathrm{id}_A)(a) = g(a)$ □

Proposition 5.5.3. If $g\colon X \to Y$ is a function, then there is some $A \subseteq X$ such that $g \restriction A$ is one-to-one

Proof. Let $A = \emptyset$ and we are done. □

Proposition 5.5.4. If $g\colon X \to Y$ is onto, then there is some $A \subseteq X$ such that $g \restriction A$ is a bijection. (Assume that there exists a function $f\colon Y \to X$ such that $g \circ f = \mathrm{id}_Y$, i.e. g has a left inverse.)

Remarks. To prove the above, for each $y \in Y$, we want to choose exactly one $x \in X$ such that $g(x) = y$. Again, the issue with choices arises. For now, we assume the existence of right inverse.

Proof. Define $A = \mathrm{range}(f)$. Then $A \subseteq X$.
To prove $g \restriction A$ is injective: Pick $a_1, a_2 \in A$. Suppose $g(a_1) = g(a_2)$. By definition, there are $y_1, y_2 \in Y$ such that $a_1 = f(y_1)$ and $a_2 = f(y_2)$. Then, $g(a_1) = g(f(y_1)) = g(a_2) = g(f(y_2))$. Since $g \circ f = \mathrm{id}_Y$, we have $y_1 = y_2$. So $a_1 = f(y_1) = f(y_2) = a_2$ as desired.
To prove $g \restriction A$ is surective: Given $y \in Y$. We have $f(y) \in A$ and $g(f(y)) = y$ as desired. □

5.6 Image of a function on a subset of the domain

Definition 5.6.1. Suppose $g\colon X \to Y$ is a function and $A \subseteq X$. The image of A under g, denoted by $g[A]$, is defined to be range$(g \restriction A)$, or equivalently, $\{g(a) \in Y : a \in A\}$.
If $B \subseteq Y$, the preimage of B under g, denoted by $g^{-1}[B]$, is defined to be $\{x \in X\colon g(x) \in B\}$

Remarks. One should take note that

- $g^{-1}[B]$ is defined even if g^{-1} does not exist.
- if g^{-1} exists, the preimage of B under g agrees with the image of B under g^{-1}.
- if B is a singleton $\{y\}$, one can write $g^{-1}(y)$. The set $g^{-1}(y)$ is called the fiber of y.

Proposition 5.6.2. Suppose $g\colon X \to Y$ is a function and $A, B \subseteq X$, then $g[A \cup B] = g[A] \cup g[B]$

Proof. $y \in g[A \cup B]$ iff there is some $x \in A \cup B$ such that $g(x) = y$. This holds iff there is some $x \in A$ or $x \in B$ such that $g(x) = y$, i.e., $y \in g[A]$ or $y \in g[B]$. This holds iff $y \in A \cup B$. □

Remarks. This result can be easily generalised for arbitrary unions over arbitrary index sets.

Proposition 5.6.3. Suppose $g\colon X \to Y$ is a function and $A, B \subseteq X$, then $g[A \cap B] \subseteq g[A] \cap g[B]$

Proof. Suppose $y \in g[A \cap B]$, i.e., there is some $x \in A \cap B$ such that $g(x) = y$. Since $x \in A$ and $x \in B$, we have $y \in g[A]$ and $y \in g[[B]$. So $y \in g[A] \cap g[B]$ □

Proposition 5.6.4. Suppose $g\colon X \to Y$ is a function and $A \subseteq X$, then $A \subseteq g^{-1}[g[A]]$

Proof. Fix $x \in A$. Then $g(x) \in g[A]$. So $x \in \{x \in X\colon g(x) \in g[A]\}$. So $x \in g^{-1}[g[A]]$. □

Chapter 6 Number Theory

6.1 Infinitude of primes

Proposition 6.1.1. For every $b \in \mathbb{N}^+$, there is a prime number greater than b.

Proof. Consider the number $b! + 1$. If $b! + 1 > b$ is a prime number, then we are done. Otherwise, there is some prime number $p < b! + 1$ such that $p \mid b! + 1$. If $p \leq b$, then $p \mid b!$, so $p \nmid b! + 1$. So, we must have $p > b$ □

6.2 Bounded Sets

Definition 6.2.1. A set $X \subseteq \mathbb{N}$ is bounded if there is some $b \in \mathbb{N}$ such that for all $x \in X$, we have $x \leq b$

Proposition 6.2.2. Every nonempty bounded set has a unique maximum element. (We denote the maximum element of X by $\max(X)$)

Proposition 6.2.3. The union of two bounded sets is bounded.

Proof. Consider sets A and B which are bounded by a and b respectively. Let $c = \max(a, b)$, then for each $x \in A$ or $X \in B$, we have $x \leq c$. □

6.3 Greatest Common Divisor

Definition 6.3.1. Let $a, b \in \mathbb{Z}$ be nonzero. A positive integer k is a gcd of a and b if $k|a$ and $k|b$, and if $d|a$ and $d|b$, then $|d| \leq k$.

Proposition 6.3.2. Given any nonzero integers a, b, their gcd exists. Furthermore, the gcd is unique.

Proof. Consider the set $S = \{|d| \in \mathbb{N}^+ : d|a \text{ and } d|b\}$. Then,

- S is nonempty because $1 \in S$
- S is bounded by $\min(a, b)$ because d divides both a and b

By Proposition 2.2, S has a unique maximum element.

□

6.4 The Division Theorem

Proposition 6.4.1. For every $a \in \mathbb{Z}$ and $b \in \mathbb{N}^+$, there are unique $q, r \in \mathbb{Z}$ such that $a = bq + r$ and $0 \leq r < b$.

Proof. We prove existence and uniqueness separately.
(*Existence*) Consider the set $S = \{r \in \mathbb{N} \colon (\exists q \in \mathbb{Z})(a = bq + r)\}$.($q$ cannot be chosen arbitrarily. $a - bq$ must be non-negative).To show that S is nonempty, consider $q = -|a|$.

$$a - bq = a + b|a| \geq -|a| + b|a| = (b-1)|a| \geq 0$$

since b is positive. Note that if $b = 0$, S could be empty (when a is negative).
By well-ordering, let r be the least element of S. It remains to prove that $r < b$. Suppose $r \geq b$, then $r - b \in S$. This is because if $r = a - bq$, then $r - b = a - b(q+1)$. But $r - b < r$, contradicting the minimality of r
(*Uniqueness*) Suppose $bq_1 + r_1 = bq_2 + r_2$, where $q_1, q_2, r_1, r_2 \in \mathbb{Z}$ and $0 \leq r_1, r_2 \leq b$. Rearranging we obtain

$$b(q_1 - q_2) = r_2 - r_1$$

So $b|(r_2 - r_1)$. If $r_2 - r_1$ is nonzero, then $b \leq |r_2 - r_1|$. But $0 \leq r_1, r_2 < b$ implies $|r_2 - r_1| < b$. So $r_2 = r_1$. Finally, $b(q_1 - q_2) = 0$, so $q_1 = q_2$ since $b > 0$. □

Remarks. In the proof above, by the definition of S, q is up to us to choose. By concluding S has a least element r, we already have a corresponding q.

Remarks. The existence part of the division theorem allows us to prove statements of the form $(\forall a \in \mathbb{Z} P(a))$ by cases.

6.5 Bezout's Identity

Definition 6.5.1. (*Ideals*) A nonempty subset I of $\mathbb{Z}$ is an ideal if:

1. for all $a, b \in I$, $a - b \in I$, and
2. for all $n \in \mathbb{Z}$ and $a \in I$, $na \in I$

It follows that 0 is an element of every ideal.

Remarks. The second property can be proved from the first one.

Proposition 6.5.2. (*Every ideal in $\mathbb{Z}$ is principal, i.e. generated by a single integer.*)
For every ideal $I \subseteq \mathbb{Z}$, there is some $k \in I \cap \mathbb{N}$ such that $I = \{nk : n \in \mathbb{Z}\}$

Proof. If $I = \{0\}$, then we can take $k = 0$. Otherwise, consider $S = I \cap \mathbb{N}^+$. If S is nonempty, we are done. If S is empty, then I must contain some negative integer s. By definition $(-1)s \in I$ and $-s \in \mathbb{N}^+$, so $-s \in S$.
By well ordering, let k be the least element of S. We claim that $I = \{nk \colon n \in \mathbb{Z}\}$

($\supseteq$) Since $k \in I$, $nk \in I$ for all $n \in \mathbb{Z}$ by the second condition in Definition 5.1. So $\{nk\colon n \in \mathbb{Z}\} \subseteq I$.
($\subseteq$) Fix $a \in I$. Given $k \in \mathbb{N}^+$ By the division theorem, there are integers n and r such that $a = nk + r$ and $0 \leq r < k$. It remains to show that $r = 0$. Since $n \in \mathbb{Z}$ and $k \in I$, we have $nk \in I$. Furthermore, since $a \in I$, we have $r = a - nk \in I$. If r were to be nonzero, then $r \in S$. By the minimality of k, $k \leq r$, contradicting to $r < k$. Hence, $r = 0$. □

Proposition 6.5.3. (*Bezout's Identity*) For every nonzero $a, b \in \mathbb{Z}$, there is some $k \in \mathbb{N}^+$ such that:

- $k|a$ and $k|b$, and
- there are $n, m \in \mathbb{Z}$ such that $k = na + mb$.

Proof. Consider $I = \{na + mb\colon n, m \in \mathbb{Z}\}$. One can easily check that I is an ideal. Since ideals in Z are principal, fix $k \in I \cap \mathbb{N}$ such that $I = \{nk : n \in \mathbb{Z}\}$. Note that $k \neq 0$ because I contains a nonzero number (such as a). So $k \in \mathbb{N}^+$. Since $a, b \in I$, we have $k|a$ and $k|b$. Since $k \in I$, there are $n, m \in \mathbb{Z}$ such that $k = na + mb$. □

Proposition 6.5.4. For every nonzero $a, b \in \mathbb{Z}$, there are $n, m \in \mathbb{Z}$ such that $na + mb = \gcd(a, b)$. Furthermore, if $d|a$ and $d|b$, then $d|gcd(a, b)$

Proof. By proposition 6.5.3, there is some $k \in \mathbb{N}^+$ and $n, m \in \mathbb{Z}$ such that $k = na + mb$ and $k|a$ and $k|b$. It follows that $k \leq \gcd(a, b)$. Since $\gcd(a, b)|a$ and $\gcd(a, b)|b$, we have $\gcd(a, b)|k = na + mb$. So $\gcd(a, b) \leq k$. Thus, $\gcd(a, b) = k = na + mb$. □

Remarks. It follows from the proofs of Propositions 6.5.2 and 6.5.3 that $\gcd(a, b)$ is the smallest positive linear combination of a and b.

Proposition 6.5.5. (*Euclid's Lemma*) If $a, b \in \mathbb{N}^+$ and p is a prime number with $p|ab$, then $p|a$ or $p|b$.

Proof. Consider $I = \{n \in \mathbb{Z}\colon p \mid na\}$. We claim that I is an ideal. (clearly, I is nonempty since $0 \in I$)

- Fix $n_1, n_2 \in I$. If $p \mid n_1 a$ and $p \mid n_2 a$, then $p \mid n_1 a - n_2 a$.
- If $p \mid n_1 a$, then for all $n \in \mathbb{Z}$, we have $p \mid n(n_1 a)$

Since every ideal in $\mathbb{Z}$ is principal, fix $k \in I \cap \mathbb{N}$ such that $I = \{nk\colon n \in \mathbb{Z}\}$. Since $p \mid pa$, We now have $p, k \in \{n \in \mathbb{Z}\colon p \mid na\} = \{mk\colon m \in \mathbb{Z}\}$. So $p = mk$ for some $m \in \mathbb{Z}$. So $k|p$. Since p is a prime number, either $k = 1$ or $k = p$.

- Case 1: If $k = 1$, then $p \mid a$ because $p \mid ka$
- Case 2: If $k = p$, since $p \mid ab$(by assumption), $b \in I$, so $b = km$ for some $m \in \mathbb{Z}$, i.e., $p \mid b$

□

6.6 Fundamental Theorem of Arithmetic

Definition 6.6.1. (*Finite support*) Let $P \subseteq \mathbb{N}^+$. A function $e\colon P \to \mathbb{N}^+$ has **finite support** if the set $\{p \in P\colon e(p) \neq 0\}$ is bounded. Let $\mathcal{F}$ denote the set of all functions $e\colon P \to \mathbb{N}^+$ with finite support.

Theorem 6.6.2. (*Fundamental Theorem of Arithmetic*) For each $a \in \mathbb{N}^+$, there is a unique $e_a \in \mathcal{F}$ such that $a = \prod\{p^{e_a(p)}\colon e_a(p) \neq 0\}$

Proof. We shall prove the existence and uniqueness separately.
(*Existence*) Base case: When $a = 1$, define $e_1(p) = 0$ for all $p \in P$. Thus, $\{p \in P\colon e_1(p) \neq 0\} = \varnothing$, which is bounded. So, $e_1 \in \mathcal{F}$ and we have $\prod\{p^{e_a(p)}\colon e_a(p) \neq 0\} = \prod \varnothing = 1$.
Inductive step: Suppose for all $a' < a$, there is some $e_{a'} \in \mathcal{F}$ such that $a' = \prod\{p^{e_{a'}(p)}\colon e_{a'}(p) \neq 0\}$. If $a \in P$, define

$$e_a(p) = \begin{cases} 1 & \text{if } p = a \\ 0 & \text{otherwise} \end{cases}$$

One can easily check e_a is in $\mathcal{F}$, and $a = a^{e_a(a)}$.
If $a \notin P$, there is some $q \in P$ such that $a = kq$ for some $k \in \mathbb{N}^+/\{1\}$. Notice that $k < a$, By the induction hypothesis, there is some $e_k \in \mathcal{F}$ such that $k = \prod\{p^{e_k(p)}\colon e_k(p) \neq 0\}$. We define

$$e_a(p) = \begin{cases} e_k(p) + 1 & \text{if } p = q \\ e_k(p) & \text{otherwise} \end{cases}$$

We claim that $e_a\colon P \to \mathbb{N}^+$ has finite support(i.e.,$\{p \in P\colon e_a(p) \neq 0\}$ is bounded). We know $\{p \in P\colon e_k(p) \neq 0\}$ is bounded, and suppose it has a bound b. Clearly, $q \in \{p \in P\colon e_a(p) \neq 0\}$. If $q \leq b$, then $\{p \in P\colon e_a(p) \neq 0\}$ is bounded by b. Otherwise it is bounded by q. Therefore, $e_a \in \mathcal{F}$. Notice that $e_a(p) \neq 0$ whenever $e_k(p) \neq 0$, so one can easily verify that

$$a = qk = \prod\{p^{e_a(p)-e_k(p)}\colon e_a(p) \neq 0\} \prod\{p^{e_k(p)}\colon e_a(p) \neq 0\} = \prod\{p^{e_a(p)}\colon e_a(p) \neq 0\}$$

(*Uniqueness*) Suppose towards a contradiction that for some $a \in \mathbb{N}^+$, there are distinct functions $e_a, e'_a \in \mathcal{F}$ such that $a = \prod\{p^{e_a(p)}\colon e_a(p) \neq 0\} = \prod\{p^{e'_a(p)}\colon e'_a(p) \neq 0\}$. Fix $e_a \neq e'_a$, fix some $q \in P$ such that $e_a(q) \neq e'_a(a)$. WLOG, suppose $e_a(q) < e'_a(q)$. Now we can divide both products by $q^{e_a(q)}$ to obtain

$$\prod\{p^{e_a(p)}\colon e_a(p) \neq 0 \wedge p \neq q\} = q^{e'_a(p)-e_a(p)} \prod\{p^{e'_a(p)}\colon e'_a(p) \neq 0 \wedge p \neq q\}$$

So q must divide the LHS. By Euclid's Lemma, q must divide some prime number not equal to q, which yields the desired contradiction.

□

Proposition 6.6.3. The function $a \mapsto e_a$ is a bijection from $\mathbb{N}^+$ to $\mathcal{F}$.

Proof. First, by FTA, the function $a \mapsto e_a$ from $\mathbb{N}^+$ to $\mathcal{F}$ is well defined.
To prove $a \mapsto e_a$ is onto, consider $e \in \mathcal{F}$. Define $a = \prod\{p^{e(p)} \colon e(p) \neq 0\}$. This product is well defined because $e \in \mathcal{F}$. Also, by FTA, since $a \in \mathbb{N}^+$, $a = \prod\{p^{e_a(p)} \colon e_a(p) \neq 0\}$ and e_a is unique. Hence, we have $e = e_a$.
To prove $a \mapsto e_a$ is one-to-one, consider $a, a' \in \mathbb{N}^+$. Suppose $e_a = e_{a'}$, we have

$$a = \prod\{p^{e_a(p)} \colon e_a(p) \neq 0\} = \prod\{p^{e_{a'}(p)} \colon e_{a'}(p) \neq 0\} = a'$$

□

Remarks.

- The proof for surjectivity uses the uniqueness of e_a. If e_a is characterised by a certain property and we want to show e equals to e_a, it suffices to prove that e has the same property as e_a.

- The FTA does not give us constructive definition of e_a. It only asserts that there is unique e_a with certain properties. So in this proof what matters is the property of e in the set $\mathcal{F}$.

Proposition 6.6.4. For every $a, b \in \mathbb{N}^+$, we have $e_{ab}(p) = e_a(p) + e_b(p)$ for all $p \in P$

Proof. By the definition of e_a and e_b, we have

$$a = \prod\{p^{e_a(p)} \colon e_a(p) \neq 0\} \quad \text{and} \quad b = \prod\{p^{e_b(p)} \colon e_b(p) \neq 0\}$$

Multiplying them together, we have

$$ab = \prod\{p^{e_a(p)+e_b(p)} \colon e_a(p) \neq 0 \vee e_b(p) \neq 0\} = \prod\{p^{e_a(p)+e_b(p)} \colon e_a(p) + e_b(p) \neq 0\}$$

Let e denote the function $e_a + e_b$, i.e., for each $p \in P$, $e_a(p) + e_b(p) = e(p)$. We claim that $e(p)$ has finite support, i.e., the set $\{p \in P \colon e_a(p) + e_b(p) \neq 0\}$ is bounded. Notice that

$$\begin{aligned}\{p \in P \colon e_a(p) + e_b(p) \neq 0\} &= \{p \in P \colon e_a(p) \neq 0 \vee e_b(p) \neq 0\} \\ &= \{p \in P \colon e_a(p) \neq 0\} \cup \{p \in P \colon e_b(p) \neq 0\}\end{aligned}$$

We conclude that $\{p \in P \colon e_a(p) + e_b(p) \neq 0\}$ is bounded because the union of two bounded sets is bounded(**Proposition 6.2.3**). Therefore, we have $e \in \mathcal{F}$. By the uniqueness of e_{ab}, we have $e_a + e_b = e = e_{ab}$ □

Remarks. The subtlety in this proof is that in order to apply the unique property of e_{ab}, one must prove that $e = e_a + e_b$ is actually in $\mathcal{F}$ because the uniqueness property only applies to elements in $\mathcal{F}$.

Proposition 6.6.5. For every $a, b \in \mathbb{N}^+$, we have $a \mid b$ if and only if $e_a(p) \leq e_b(p)$ for all $p \in P$

Proof. (⇒) Suppose $b = ka$. By **Proposition 6.6.4**, we have $e_b(p) = e_{ak}(p) = e_a(p) + e_k(p)$ for all $p \in P$. Since $e_k(p) \geq 0$ for all $p \in P$, we have $e_a(p) \leq e_b(p)$ for all $p \in P$.
(⇐) Suppose $e_a(p) \leq e_b(p)$ for all $p \in P$. Define

$$k = \prod\{p^{e_b(p)-e_a(p)} \colon e_b(p) \neq 0\}$$

Notice that, since $0 \leq e_a(p) \leq e_b(p)$, whenever $e_a(p) \neq 0$, we have $e_b(p) \neq 0$ and whenever $e_a(p) = 0$, we have $p^{e_a(p)} = 1$. Thus,

$$a = \prod\{p^{e_a(p)} \colon e_a(p) \neq 0\} = \prod\{p^{e_a(p)} \colon e_b(p) \neq 0\}$$

It follows that

$$\begin{aligned} ka =& \prod\{p^{e_b(p)-e_a(p)} \colon e_b(p) \neq 0\} \prod\{p^{e_a(p)} \colon e_b(p) \neq 0\} \\ =& \prod\{p^{e_b(p)} \colon e_b(p) \neq 0\} \\ =& b \end{aligned}$$

□

Remarks. Two subtleties in this proof

- When defining k, we only restrict $e_b(p)$ to be non-zero. Noting that $e_a(p) \neq 0$ is stronger than $e_b(p) \neq 0$, we must allow $e_a(p) = 0$. Otherwise, we may completely neglect those prime factors which divide b but do not divide a. We also do not restrict $e_b(p) - e_a(p) \neq 0$ because we need the all terms with $e_b(p) \neq 0$ for later use.
- We changed $a = \prod\{p^{e_a(p)} \colon e_a(p) \neq 0\}$ to $a = \prod\{p^{e_a(p)} \colon e_b(p) \neq 0\}$ to make the condition in the set builder notation match with that of b. We can do this precisely because $e_a(p) \neq 0$ is stronger than $e_b(p) \neq 0$ and, in cases where $e_a(p) = 0$ and $e_b(p) \neq 0$, changing the condition has no effect on the product as $p^{e_a(p)} = 1$.

Proposition 6.6.6. (*gcd revisited*) For every $a, b \in \mathbb{N}^+$, there is some $k \in \mathbb{N}^+$ such that

- $k \mid a$ and $k \mid b$ and
- if $d \mid a$ and $d \mid b$, then $d \mid k$.

Proof. Define $f \colon P \to \mathbb{N}^+$ by $f(p) = \min\{e_a(p), e_b(p)\}$. We claim that f has finite support, i.e., the set $\{p \in P \colon f(p) \neq 0\}$ is bounded.
By definition, $f(p) \leq e_a(p)$ and $f(p) \leq e_b(p)$ for all $p \in P$. Therefore, $f(p) \neq 0$ is stronger than $e_a(p) \neq 0$. So, $\{p \in P \colon f(p) \neq 0\} \subseteq \{p \in P \colon e_a(p) \neq 0\}$. It follows that $\{p \in P \colon f(p) \neq 0\}$ is bounded. Hence,we have $f \in \mathcal{F}$.
Since the function $a \mapsto e_a$ is a surjection from $\mathbb{N}^+$ to $\mathcal{F}$, we can fix some $k \in \mathbb{N}^+$ such that $f = e_k$. Then, $e_k(p) \leq e_a(p)$ and $e_k(p) \leq e_b(p)$ for all $p \in P$. Therefore, $k \mid a$ and $k \mid b$.
If $d \mid a$ and $d \mid b$, then $e_d(p) \leq e_a(p)$ and $e_d(p) \leq e_b(p)$ for all $p \in P$. Therefore, $e_d(p) \leq \min\{e_a(p), e_b(p)\} = f(p) = e_k(p)$ for all $p \in P$. We conclude that $d \mid k$. □

Remarks. We need to prove f has finite support before applying the properties of the functions with finite support.

6.7 Modular Arithmetic

Definition 6.7.1. (*The Remainder Function*) Fix $b \in \mathbb{N}^+$. Denote the set $\{0, 1, 2, ..., b-1\}$ by $[b]$. $R_b \colon \mathbb{Z} \to [b]$ is the function which maps $a \in \mathbb{Z}$ to the remainder of a when a is divided by b.

Remarks. By the division theorem, $R_b \colon \mathbb{Z} \to [b]$ is well defined.

Proposition 6.7.2. $R_b \colon \mathbb{Z} \to [b]$ is onto but not one-to-one.

Proof. For every $c \in [b]$, we have $c \in \mathbb{Z}$ such that $R_b(c) = c$. So R_b is onto. Notice that $R_b(c) = R_b(c+b)$ and $c \neq c+b$, so R_b is not one-to-one. □

Proposition 6.7.3. $R_b(a) = R_b(a')$ if and only if $b \mid (a - a')$

Proof. ($\Rightarrow$) Suppose $R_b(a) = R_b(a') = r$. Fix $q, q' \in \mathbb{Z}$ such that $r = a - bq = a' - bq'$. Rearranging, we have $a - a' = b(q - q')$. Thus, $b \mid (a - a')$
($\Leftarrow$) (By contrapositive) Suppose $a - bq = r$ and $a' - bq' = r'$. WLOG, assume $r > r'$. Then, $(a - a') = b(q - q') + (r - r')$. Since $0 < r' < r < b$, we have $0 < r - r' < b$. Thus, b does not divide $a - a'$. □

Definition 6.7.4. (*+ on* $[b]$) Addition on [b] is the function $+_b \colon [b] \times [b] \to [b]$ such that

$$R_b(a) +_b R_b(a') = R_b(a + a')$$

Proposition 6.7.5. $+_b$ is well defined.

Proof. First, since R_b is onto, for each c and each d in $[b]$, there is $a \in \mathbb{Z}$ such that $c = R_b(a)$ and there is $a' \in \mathbb{Z}$ such that $d = R_b(a')$. And since R_b is well defined, for each $(c, d) \in [b] \times [b]$, $R_b(a + a')$ will output some value.
Second, we need to prove that for each $(c, d) \in [b] \times [b]$, $c +_b d$ has at most one value.
Suppose $c = R_b(a_1) = R_b(a_2)$ and $d = R_b(a'_1) = R_b(a'_2)$. We shall prove that $R_b(a_1 + a'_1) = R_b(a_2 + a'_2)$. Fix $q_1, q_2, q'_1, q'_2 \in \mathbb{Z}$ such that

$$c = a_1 - bq_1 = a_2 - bq_2 \quad \text{and} \quad d = a'_1 - bq'_1 = a'_2 - bq'_2 \tag{1}$$

Rearranging, we have

$$a_1 + a'_1 = b(q_1 + q'_1) + (c + d) \quad \text{and} \quad a_2 + a'_2 = b(q_2 + q'_2) + (c + d) \tag{2}$$

Subtracting the two equations, we obtain

$$(a_1 + a_1') - (a_2 + a_2') = b((q_1 + q_1') - (q_2 + q_2')) \tag{3}$$

Clearly, $b \mid (a_1 + a_1') - (a_2 + a_2')$. By **Proposition 6.7.3**, we are done.

□

Remarks. Two subtleties in this proof

- For each $c \in [b]$, there are in fact infinitely many a such that $R_b(a) = c$. Likewise for d. So we need to justify that regardless of the choices of a and a', the output value is the same, i.e., the output only depends on c and d and it is unique.
- One cannot conclude from Equation (2) that $R_b(a_1 + a_1') = c + d = R_b(a_2 + a_2')$ because $c + d$ may be greater or equal to b. To resolve this issue, one could use either Proposition 6.7.3 or the division theorem.

6.8 Congruence Classes

Definition 6.8.1. Fix $b \in \mathbb{N}^+$. We say that a and a' are congruent $\mod b$ if $b \mid (a - a')$. (Recall that $R_b(a) = R_b(a')$ if and only if $b \mid (a - a')$.)

Definition 6.8.2. Define a function $C_b \colon \mathbb{Z} \to \mathcal{P}(\mathbb{Z})$ by

$$C_b(a) = \{a' \in \mathbb{Z} \colon b \mid (a - a')\} = \{a' \in \mathbb{Z} \colon R_b(a) = R_b(a')\}$$

Remarks. Let $Q_b \subseteq \mathbb{Z}$ denote the range of C_b. We restrict codomain of C_b to Q_b so that C_b is onto.

Proposition 6.8.3. If $b \mid (a - a')$, then $C_b(a) = C_b(a')$.

Proof. By Proposition 6.7.3, we have $R_b(a) = R_b(a')$.
($\subseteq$) Take $c \in C_b(a)$, then $R_b(a) = R_b(c)$. Since $R_b(a) = R_b(a')$, we have $R_b(c) = R_b(a')$. So $c \in C_b(a')$.
($\supseteq$) $C_b(a) \supseteq C_b(a')$ follows analogously. □

Theorem 6.8.4. (*Universal Property of* $C_b \colon \mathbb{Z} \to Q_b$) Suppose X is a set and $f \colon \mathbb{Z} \to X$ is a function such that if $b \mid (a - a')$, then $f(a) = f(a')$. Then there exists a unique function $g \colon Q_b \to X$ such that $f = g \circ C_b$. (One should take note that $f(a)$ could be equal to $f(a')$ even if $b \nmid (a - a')$.)

Example. Let's first consider some special f and find out their corresponding g.

1. If $f \colon \mathbb{Z} \to X$ is a constant function that maps all integers to $x \in X$, then $g \colon Q_b \to X$ is also a constant function which maps all $C \in Q_b$ to $x \in X$.

2. If f is $C_b\colon \mathbb{Z} \to Q_b$, then $g\colon Q_b \to Q_b$ is the identity function.

3. If f is $R_b\colon \mathbb{Z} \to [b]$, then $g\colon Q_b \to [b]$ is such that $g(C_b(a)) = R_b(a)$

Proof. We shall prove the existence and uniqueness separately.
(*Existence*) Define $g\colon Q_b \to X$ by $g(C) = f(a)$ for each $a \in C$. We shall prove that g is well defined. Consider $C \in Q_b$ and take $a \in \mathbb{Z}$ such that $C = C_b(a)$. Fix $a_1, a_2 \in C_b(a)$. Then, $b \mid (a - a_1)$ and $b \mid (a - a_2)$. By assumption, $f(a) = f(a_1)$ and $f(a) = f(a_2)$. Therefore, $f(a_1) = f(a_2)$. So the value of $g(C)$ is unique for all $a \in C$. Therefore, g is well defined. Furthermore, since $a \in C_b(a)$, one can check that $g \circ C_b(a) = g(C_b(a)) = f(a)$.
(*Uniqueness*) Suppose we have $h\colon Q_b \to X$ such that $f = h \circ C_b$. For each $C_b(a) \in Q_b$, we have $h(C_b(a)) = f(a) = g(C_b(a))$. □

Remarks. Every $C \in Q_b$ is of the form $C_b(a)$. So instead of writing " consider $C \in Q_b$ take $a \in \mathbb{Z}$ such that $C = C_b(a)$", one can simply write "take $C_b(a) \in Q_b$".

Remarks. From the earlier example, $C_b\colon \mathbb{Z} \to Q_b$ is an example of f. So the universal property is saying that the function C_b is the 'source' of all functions f, i.e., all functions f with a certain property can be obtained from a function C_b with the same property (by composing some g to C_b).

Proposition 6.8.5. For every function $g\colon Q_b \to X$, the function $g \circ C_b\colon \mathbb{Z} \to X$ is an example of f.

Proof. By Proposition 6.8.3, whenever $b \mid (a - a')$, we have $C_b(a) = C_b(a')$. It follows that $g(C_b(a)) = g(C_b(a'))$, .i.e., $g \circ C_b(a) = g \circ C_b(a')$ □

Proposition 6.8.6. Fix $b \in \mathbb{N}^+$ and function $C_b\colon \mathbb{Z} \to Q_b$. For each set X there is a bijection between the following two sets of functions:

$$G = \{g\colon g \text{ is a function from } Q_b \text{ to } X\}$$
$$F = \{f\colon f \text{ is a function from } Z \text{ to } X \text{ such that if } b \mid (a - a'), \text{ then } f(a) = f(a')\}$$

given by $g \mapsto g \circ C_b$.

Proof. First, we prove that $g \mapsto g \circ C_b$ is well defined. For each $g \in G$, $g \circ C_b$ is the only output. Furthermore, by Proposition 6.8.5, we have $g \circ C_b \in F$. So $g \mapsto g \circ C_b$ is well defined.
To prove $g \mapsto g \circ C_b$ is one-to-one, fix $g_1, g_2 \in G$. Suppose $f = g_1 \circ C_b = g_2 \circ C_b$. By Theorem 6.8.4, there is a unique $g \in G$ such that $f = g \circ C_b$. Thus, we have $g_1 = g = g_2$.
To prove $g \mapsto g \circ C_b$ is onto, consider $f \in F$. By Theorem 6.8.4, there is a unique $g \in G$ such that $f = g \circ C_b$.

□

Chapter 7 Equivalence Relations

7.1 Equivalence Relation and Quotient Map

Definition 7.1.1. (*Equivalence Relations*) A relation $\sim$ on a set A is an equivalence relation if:

- it is reflexive, i.e., $(\forall a \in A)(a \sim a)$
- it is symmetric, i.e., $(\forall a, b \in A)(a \sim b \to b \sim a)$
- it is transitive, i.e., $(\forall a, b, c \in A)((a \sim b \wedge b \sim c) \to a \sim c)$

Proposition 7.1.2. Suppose f is a fixed function with domain A. Define a relation on A by

$$a \sim b \text{ if } f(a) = f(b).$$

Then, $\sim$ is an equivalence relation.

Definition 7.1.3. (*Equivalence Class*) For each $a \in A$, the set $\{a' \in A \colon a' \sim a\}$ is called the equivalence class of a. We denote it by $[a]_\sim$

Proposition 7.1.4. Suppose $\sim$ is an equivalence relation on A and $a, b \in A$. If $a \sim b$, then $[a]_\sim = [b]_\sim$.

Proof. ($\subseteq$) Fix $a' \in [a]_\sim$. Then, $a \sim a'$ and $a' \sim a$ by the symmetric property. Since $a \sim b$, we have $a' \sim b$ by transitivity. So $a \in [b]_\sim$. ($\supseteq$) This direction follows analogously. □

Definition 7.1.5. (*Quotient*) Suppose $\sim$ is an equivalence relation on A. The quotient, denoted by $A/\sim$, is defined to be the set

$$\{\{a' \in A \colon a' \sim a\} \in \mathcal{P}(A) \colon a \in A\} = \{[a]_\sim \in \mathcal{P}(A) \colon a \in A\}$$

In other words, $A/\sim \subseteq \mathcal{P}(A)$ is the set of all equivalence classes.

Definition 7.1.6. (*Quotient Map*) Suppose $\sim$ is an equivalence relation on A. The quotient map $\pi \colon A \to A/\sim$ is defined by $\pi(a) = [a]_\sim$.

Proposition 7.1.7. $\pi \colon A \to A/\sim$ is onto but not one-to-one (unless $\sim$ is equality).

Theorem 7.1.8. (*Universal Property of* $\pi \colon A \to A/\sim$) For every set X and every function $f \colon A \to X$ such that if $a \sim b$, then $f(a) = f(b)$, there is a unique function $g \colon A/\sim \to X$ such that $f = g \circ \pi$.

Proof. Define $g\colon A/\sim \to X$ by $g([a]_\sim) = f(a)$ for all $a \in [a]_\sim$.
To prove g is well defined, consider $[a]_\sim \in A/\sim$. Fix $a_1, a_2 \in [a]_\sim$. Then, $a_1 \sim a_2$ and $a_2 \sim a_1$. By symmetry and transitivity, we have $a_1 \sim a_2$. By assumption, $f(a_1) = f(a_2)$. Furthermore, since $a \in [a]_\sim$, $g \circ \pi(a) = g(\pi(a)) = g([a]_\sim) = f(a)$
To prove g is unique, suppose $h\colon A/\sim \to X$ is such that $f = h \circ \pi$. For every $[a]_\sim \in A/\sim$, we have $h([a]_\sim) = f(a) = g([a]_\sim)$ as desired. □

7.2 Treating $\sim$ as Equality

Notation 7.2.1. For all sets A and B, $\text{Maps}(A, B)$ denote the set of all functions from A to B.

Proposition 7.2.2. Suppose $\sim$ is an equivalence relation on A. For each set X the function defined by

$$\begin{aligned} \text{Maps}(A/\sim, X) &\to \{f \in \text{Maps}(A, X)\colon (\forall a, a' \in A)(a \sim a' \to f(a) = f(a'))\} \\ g &\mapsto g \circ \pi \end{aligned}$$

is a bijection.

Proof. First, one can check $g \circ \pi$ is a function from A to X. Furthermore, for all $a, a' \in A$, if $a \sim a'$, then $[a]_\sim = [a']_\sim$. Thus, if $a \sim a'$, then $g \circ \pi(a) = g([a]_\sim) = g([a']_\sim) = g \circ \pi(a')$. So $g \circ \pi \in \{f \in \text{Maps}(A, X)\colon (\forall a, a' \in A)(a \sim a' \to f(a) = f(a'))\}$. Lastly, it is clear that for each $g \in \text{Maps}(A/\sim, X)$, $g \circ \pi$ is the only output.
To prove injectivity, take $g, g' \in \text{Maps}(A/\sim, X)$. Suppose for all $a \in A$, $g \circ \pi(a) = g' \circ \pi(a)$. Then, $g([a]_\sim) = g'([a]_\sim)$ for each $a \in A$. So, $g([a]_\sim) = g'([a]_\sim)$ for each $[a]_\sim \in A/\sim$ as desired.
To prove surjectivity, take $f \in \{f \in \text{Maps}(A, X)\colon (\forall a, a' \in A)(a \sim a' \to f(a) = f(a'))\}$. By the universal property, there is a unique function $g \in \text{Maps}(A/\sim, X)$ such that $f = g \circ \pi$. □

Remarks. If we wish to treat $\sim$ as equality, then we should work in the quotient set.:

1. If we want to work with some $f \in F$, it is equivalent to work with the corresponding $g \in G$. Notice that functions on $A/\sim$ treats $\sim$ as equality, because all it sees are $\sim$ classes, not elements of A.

2. All functions f treat $\sim$ as equality since $a \sim a' \to f(a) = f(b)$. By the universal property, all such f come from π.

7.3 Constructing functions with domain $A/\sim$

In order to define a function from $A/\sim$ to X, it suffices to define a function $f\colon A \to X$ such that if $a \sim a'$, then $f(a) = f(b)$, and then apply the universal property to find g.

$(g([a]_\sim) = f(a))$.
In fact, by Proposition 7.2.2, every function $g\colon A/\sim \to X$ can be obtained in this way.

7.4 Equivalence relations and partitions

Proposition 7.4.1. Let $\sim$ be an equivalence relation on A. For every $a, a' \in A$, the following are equivalent:

$$(1)\, a \sim a' \quad (2)\, [a]_\sim = [a']_\sim \quad (3)\, (\exists b \in A)(a, a' \in [b]_\sim) \quad (4)\, [a]_\sim \cap [a']_\sim \neq \emptyset$$

Proof. (1)⇒(2): By propositon 7.1.4 .
(2)⇒(3): Suppose $[a]_\sim = [a']_\sim$. Then, by reflexivity, we have $a \in [a]_\sim$ and $a' \in [a']_\sim = [a]_\sim$. Thus, $a, a' \in [a]_\sim$ as desired.
(3)⇒(4): Suppose $a, a' \in [b]_\sim$, then $a \sim b$ and $a' \sim b$. By symmetric property, we have $b \sim a$ and $b \sim a'$. Thus, we have $b \in [a]_\sim \cap [a']_\sim$ as desired.
(4)⇒(1): Suppose $[a]_\sim \cap [a']_\sim \neq \emptyset$. Then, there exists $b \in [a]_\sim \cap [a']_\sim$ such that $a \sim b$ and $a' \sim b$. By symmetry and transitivity, we have $a \sim a'$ as desired.

□

Remarks. This proposition essentially tells us that if two equivalence classes are different as sets, then they are disjoint. ((2) and (4))

Definition 7.4.2. A set $P \subseteq \mathcal{P}(A) - \{\emptyset\}$ is a partition of A if

1. $\bigcup P = A$
2. $(\forall C, D \in P)(C = D \vee C \cap D = \emptyset)$.

Proposition 7.4.3. For every equivalence relation $\sim$ on A, the quotient set $A/\sim$ is a partition of A.

Proof. First, $\emptyset \notin A/\sim$. This holds because for each $a \in A$, $a \in [a]_\sim$ by reflexivity.
Second, we need to show $\bigcup A/\sim = A$. ($\subseteq$) Fix $x \in \bigcup A/\sim$. Then, there is some $a \in A$ such that $x \in [a]_\sim$. Since $[a]_\sim \subseteq A$, we have $x \in A$. ($\supseteq$) Fix $a \in A$. By reflexivity, we have $a \in [a]_\sim$. Thus, $a \in \bigcup A/\sim$.
Third, by proposition 7.4.1, equivalence classes are disjoint. □

Proposition 7.4.4. Suppose P is a partition of A, and the relation $\sim$ on A is defined by

For all $a, b \in A$, $a \sim b$ if there exists some $C \in P$ such that $a, b \in C$.

Then, $\sim$ is an equivalence relation on A. Furthermore, $P = A/\sim$.

Proof. First, we prove that $\sim$ is an equivalence relation.
Reflexivity: Fix $a \in A$. Since P is a partition of A, we have $\bigcup P = A$. Thus, $a \in \bigcup P$. So there is a $C \in P$ such that $a \in C$. Hence, $a \sim a$ as desired.
Symmetry: Suppose $a \sim b$, i.e., there is a $C \in P$ such that $a, b \in C$. Then, it is trival that $b \sim a$.
Transitivity: Suppose $a \sim b$ and $b \sim c$. Then, there are $C_1, C_2 \in P$ such that $a, b \in C_1$ and $b, c \in C_2$. Since $C_1 \cap C_2 \neq \emptyset$ (because they have a common element b), we have $C_1 = C_2 = C$. Thus, $a, c \in C$ as desired.
Second, we shall prove that $P = A/\sim = \{\{a' \in A\colon a' \sim a\} \in \mathcal{P}(A)\colon a \in A\}$.
($\subseteq$) Fix an arbitrary $C \in P$ and suppose $a \in C$. By definition, for all $a' \in A$, if $a' \in C$, then $a' \sim a$. This means that C is an equivalence class. So we have $C \in A/{\sim}$.
($\supseteq$) Fix $[a]_\sim \in A/\sim$. $a' \sim a$ for all $a' \in [a]_\sim$. So there exists $C \in P$ such that $a \in C$ and for all $a' \in [a]_\sim$, $a' \in C$. Furthermore, for all $x \in C$, $x \sim a$. Thus, $x \in [a]_\sim$. It follows that $C = [a]_\sim$. Thus, $[a]_\sim \in P$. □

Remarks. This result tells us that given a partition of A, we can define an equivalence relation on A, and the quotient set of this equivalence relation is the partition that we start with.

Proposition 7.4.5. The function

$$\mathcal{G}\colon \{\text{equivalence relations on } A\} \to \{\text{partitions of } A\}$$
$$\sim \mapsto A/\sim$$

is a bijection.

Proof. First, check that $\mathcal{G}$ is well defined. By proposition 7.4.3, for each $\sim$ on A, there is a $A/\sim \in \{\text{partitions of } A\}$. By definition, the quotient set is unique for each $\sim$ on A.
$\mathcal{G}$ is onto because by proposition 7.4.4, for each partition P, we can define an equivalence relation on A such that P is the quotient set.
To prove that $\mathcal{G}$ is one-to-one, recall that by proposition 7.4.1, if $\sim$ is an equivalence relation on A, then $a \sim b$ if and only if there is some $C \in A/\sim$ such that $a, b \in C$.
Based on proposition 7.4.4, we define a new function

$$\mathcal{H}\colon \{\text{partitions of } A\} \to \{\text{equivalence relations on } A\}$$

by

$$\mathcal{H}(P) = \sim \text{ if for all } a, b \in A,\ a \sim b \text{ whenever there is } C \in P \text{ such that } a, b \in C$$

By proposition 7.4.4, $\mathcal{H}$ is well defined. Furthermore, we have

$$\mathcal{H} \circ \mathcal{G}(\sim) = \mathcal{H}(A/\sim) = \sim$$

So $\mathcal{H} \circ \mathcal{G} = \mathrm{id}_{\{\text{equivalence relations on } A\}}$. By proposition 5.4.4, $\mathcal{G}$ is one-to-one.

□

7.5 Range isomorphic to quotient by some equivalence relation

Fix a set B and a function $\sigma\colon A \to B$ which is onto. Define an equivalence relation $\sim$ on A by: $a \sim a'$ if $\sigma(a) = \sigma(a')$.

Theorem 7.5.1. For every set X and every function $f\colon A \to X$ such that if $a \sim b$, then $f(a) = f(b)$, there is a unique function $h\colon B \to X$ such that $f = h \circ \sigma$.

Proof. Define $h\colon B \to X$ by

$$\text{for each } b \in B,\ h(b) = f(a)$$

where $a \in A$ is such that $\sigma(a) = b$. σ being onto guarantees the existence of such a.
To prove that h is well defined, fix $b \in B$ and consider $a, a' \in A$ such that $\sigma(a) = b = \sigma(a')$. Then, by definition, $a \sim a'$. Therefore, by assumption, $f(a) = f(a')$. Furthermore, one can check that $h \circ \sigma(a) = h(b) = f(a)$ as desired.
To prove that h is unique, suppose $h'\colon B \to X$ is such that $f = h' \circ \sigma$. Then, for each $b \in B$, we have $h'(b) = h' \circ \sigma(a) = f(a) = h(b)$ as desired. □

Corollary 7.5.2. There is a unique function $h\colon B \to A/\sim$ such that $\pi = h \circ \sigma$.

Proof. Apply theorem 7.5.1 by replacing X with $A/\sim$ and f with π. □

Theorem 7.5.3. (*Universal property for* π) For every set X and every function $f\colon A \to X$, such that if $a \sim b$, then $f(a) = f(b)$, there is a unique function $g\colon A/\sim \to X$ such that $f = g \circ \pi$.

Corollary 7.5.4. There is a unique function $g\colon A/\sim \to B$ such that $\sigma = g \circ \pi$.

Proof. Apply the universal property by replacing X with B and f with σ. □

Putting Corollary 7.5.4 and 7.5.2 together, we have the following result

Proposition 7.5.5. *$h\colon B \to A/\sim$ and $g\colon A/\sim \to B$ are bijections which are inverses of each other.*

Proof. $\mathrm{id}_{A/\sim} \circ \pi = \pi = h \circ \sigma = h \circ (g \circ \pi) = (h \circ g) \circ \pi$. So $h \circ g = \mathrm{id}_{A/\sim}$, since π is onto.
Similarly, $\mathrm{id}_B \circ \sigma = \sigma = g \circ \pi = g \circ (h \circ \sigma) = (g \circ h) \circ \sigma$. So $g \circ h = \mathrm{id}_B$, since σ is onto.
By proposition 5.4.3 and 5.4.6, g and h are bijections which are inverses of each other. □

Remarks.

1. This means that B and $A/\sim$ are "isomorphic". Intuitively, B and $A/\sim$ are of the same "size".

2. This reasoning (two objects satisfying the same universal property can be proved to be isomorphic) can be applied in other contexts.

7.6 Construction of $\mathbb{Z}$ from $\mathbb{N}$

Assume that we have defined $\mathbb{N}$ and addition $+$ on $\mathbb{N}$. We have not defined - on $\mathbb{N}$ because $\mathbb{N}$ is not closed under - (or we can only "partially define" - on $\mathbb{N}$).

Definition 7.6.1. Define relation $\sim$ on $\mathbb{N} \times \mathbb{N}$ by

$$(m,n) \sim (p,q) \quad \text{if} \quad m+q = p+n$$

(Note that the above definition uses addition but not subtraction.)

Proposition 7.6.2. $\sim$ defined above is an equivalence relation.

Definition 7.6.3. (*The Set of Integeres*) The set of integers, denoted by $\mathbb{Z}$, is defined to be the set $\mathbb{N} \times \mathbb{N}/\sim$, where $\sim$ is the equivalence relation defined above.

Remarks. Notice that the set of natural numbers is not a subset of integers. But the set $\{[(n,0)]_\sim : n \in \mathbb{N}\}$ behaves like $\mathbb{N}$, i.e., it has all the properties that $\mathbb{N}$ has (when we restrict $+_\mathbb{Z}$, $\cdot_\mathbb{Z}$, etc to it.)

Definition 7.6.4. (*Addition on* $\mathbb{Z}$) Define $+_\mathbb{Z} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ as

$$[(m,n)]_\sim +_\mathbb{Z} [(p,q)]_\sim = [(m+p, n+q)]_\sim$$

Proposition 7.6.5.

1. $+_\mathbb{Z} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ is well defined.
2. $+_\mathbb{Z}$ is associative and commutative.
3. $[(0,0)]_\sim$ is an additive identity.
4. For each $[m,n]_\sim \in \mathbb{Z}$, there is a unique $[(p,q)]_\sim \in \mathbb{Z}$ such that $[(m,n)]_\sim +_\mathbb{Z} [(p,q)]_\sim = [(0,0)]_\sim$.

7.7 Construction of $\mathbb{Q}$ from $\mathbb{Z}$

Where $\mathbb{Z}$ extends $\mathbb{N}$ by adding additive inverse, we can think of $\mathbb{Q}$ as extending $\mathbb{Z}$ by adding multiplicative inverses.

Definition 7.7.1. Define a relation $\approx$ on $\mathbb{Z} \times (\mathbb{Z} - \{0_\mathbb{Z}\})$ by

$$(a,b) \approx (c,d) \quad \text{if} \quad a \cdot_\mathbb{Z} d = c \cdot_\mathbb{Z} b$$

One can check that $\approx$ is an equivalence relation.

Definition 7.7.2. (*The Set of Rational Numbers*) The set of rational numbers, denoted by $\mathbb{Q}$, is defined to be the set $(\mathbb{Z} \times (\mathbb{Z} - \{0_\mathbb{Z}\}))/\approx$.

Remarks. $\mathbb{Z}$ is not a subset of $\mathbb{Q}$, but the set $\{[(a, 1_\mathbb{Z})]_\approx : a \in \mathbb{Z}\}$ behaves like $\mathbb{Z}$.

Chapter 8 Cardinality and choice

8.1 Equinumerous Sets

Definition 8.1.1. Two sets X and Y are equinumerous, written as $X \approx Y$, if there is a bijection from X to Y.

Proposition 8.1.2. $\approx$ is reflexive, symmetric and transitive.

Remarks. $\approx$ is not an relation. If $\approx$ were to be a relation, then it is defined on the "set" of all sets, which is not a set.

Recall that $[n] = \{0, 1, 2, ..., n-1\}$.

Definition 8.1.3. (*Finite Sets*) A set X is finite if $X \approx [n]$ for some $n \in \mathbb{N}$. Otherwise X is infinite.

Remarks. Notice that when we list elements of a finite set A without repetition, we are in fact choosing a bijection between A and $[n]$.

Example.

8.2 Cantor's Theorem

Theorem 8.2.1. For every set X, we have $X \not\approx \mathcal{P}(X)$.

Proof. Suppose $f\colon X \to \mathcal{P}(X)$ is a function. Consider

$$A = \{x \in X : x \notin f(x)\} \in \mathcal{P}(X)$$

If there is some $x \in A$ such that $f(x) = A$, then $x \in f(x)$. But for all $x \in A$, we have $x \notin f(x)$, a contradiction.
If there is some $x \in X - A$ such that $f(x) = A$, then $x \notin f(x)$. Thus, $x \in A$, a contradiction. □

Proposition 8.2.2. If X is finite and $f\colon X \to X$ is onto, then f is one-to-one.

Proposition 8.2.3. There is an injection from X to $\mathcal{P}(X)$.

Proof. □

8.3 Pigeonhole principle

Theorem 8.3.1. Every one-to-one function $f\colon [n] \to [n]$ must be onto.

Proof. We proceed by induction on $\mathbb{N}$.
Base case: Vacuously true.
Inductive step: Suppose $n \in \mathbb{N}$ is such that every injective function $f\colon [n] \to [n]$ is onto. Consider an injection $f\colon [n+1] \to [n+1]$. We consider two cases:
Case 1: For all $m \in [n]$, $f(m) \neq n$. Then one can check that $f \restriction [n]\colon [n] \to [n]$ is an injection. By induction hypothesis, $f \restriction [n]$ maps onto [n]. Sicne f is injective, we must have $f(n) = n$. Therefore, f is onto.
Case 2: There exists some $m \in [n]$ such that $f(m) = n$. Since f is injective, $f(n) \neq n$. So $f(n) \in [n]$. We define $g\colon [n] \to [n]$ by

$$g(k) = \begin{cases} f(n) & \text{if } f(k) = n \\ f(k) & \text{otherwise} \end{cases}$$

First, one can check that g is well defined because f is well defined. To check g is one-to-one, suppose $g(k) = g(k')$. If $g(k) = g(k') = f(k) = f(k')$. Since f is one-to-one, we have $k = k'$. If $g(k) = g(k') = f(n)$, then $f(k) = f(k') = n$ (by definition). It follows that $k = k'$. Since $g\colon [n] \to [n]$ is one-to-one, by the induction hypothesis, g is onto. It follows that f is also onto. □

Corollary 8.3.2. For every $m < n$, there is no injection from $[n]$ to $[m]$. In particular, $[n] \not\approx [m]$.

Proof. Suppose $f\colon [n] \to [m]$ is injective. Then, $f \restriction [m]$ is also an injection from $[m]$ to $[m]$. By the Pigeonhole Principle, $f \restriction [m]$ is onto.
Notice that $f(m) \in [m] = \text{range}(f \restriction [m])$, and since $f \restriction [m]$ is onto, there is some $k < m$ such that $f(k) = f(m)$. So f is not injective, a contradiction.

□

Proposition 8.3.3. If B is finite and $A \subsetneq B$, then there is no injection from B to A. Therefore, no finite set is equinumerous to a proper subset of itself. See proposition 8.10.8 for applications.

Proof. Suppose $f\colon B \to A$ is a function. Since B is finite, we can fix $n \in \mathbb{N}$ and a bijection $g\colon B \to [n]$. Suppose towards a conntradiction that f is injective, then

$$g \circ f \circ g^{-1}\colon [n] \to [n]$$

is injective as well. By the Pigeonhole Principle, $g \circ f \circ g^{-1}$ is onto.
Fix $b \in B - A$. $g(b) \in [n]$. So, there is some $m \in [n]$ such that $g \circ f \circ g^{-1}(m) = g(b)$. Since g is injective, we have $f \circ g^{-1}(b) = b$. But b is not in the codomain of f, yielding a contradiction. □

Remarks. The contrapositive of this statement tells us that if a set is equinumerous to a proper subset of itself, then it is infinite. Hence, $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ are infinite.

8.4 Cardinality of finite sets

Proposition 8.4.1. Every finite set is equinumerous to a unique $[n]$.

Proof. If A is finite, then $A \approx [n]$ for some $n \in \mathbb{N}$. If $A \approx [n]$ and $A \approx [m]$, then by symmetry and transitivity, we have $[n] = [m]$. By Corollary 8.3.2 we have $m = n$. □

Definition 8.4.2. (*Cardinality*) If $A \approx [n]$, we say that the cardinality of A, denoted by $|A|$, is equal to n.

Proposition 8.4.3. If $A \subsetneq [n]$, then A is finite and $|A| < n$.

Proof. We proceed by induction on $\mathbb{N}$.
Base Case: $[0]$ has no proper subsets, so the statement is vacuously true.
Inductive Step: Suppose every proper subset of $[n]$ is equinumerous to $[m]$ for some $m < n$. Consider $A \subsetneq [n+1]$.
Case 1: $n \notin A$. If $A = [n]$, then we are done because $n < n+1$. If $A \subsetneq [n]$, then we are done by the induction hypothesis.
Case 2: $n \in A$. Then $A - n \subsetneq [n]$ (because $A \subsetneq [n+1]$). By the induction hypothesis, $A - n \subsetneq [n] \approx [m]$ for some $m < n$. It follows that $A \approx [m+1]$.

□

Proposition 8.4.4. If B is finite and $A \subsetneq B$, then A is finite and $|A| < |B|$.

Proof. Since B is finite, we can fix some $n \in \mathbb{N}$ and a bijection $g\colon B \to [n]$. Take $b \in B - A$, we have $g(b) \notin g[A]$ because g is one-to-one. So, we have $g[A] \subsetneq [n]$. By proposition 8.4.3, $g[A]$ is finite. Since $A \approx g[A]$, we have A, A is finite and $|A| = |g[A]| < n$. □

Proposition 8.4.5. If A and B are finite, then $|A| \leq |B|$ if and only if there is an injection from A to B. (Every subset of a finite set is finite.)

Proof. ($\Leftarrow$) Fix an injection $f\colon A \to B$. Then $f[A] \subsetneq B$.
If $f[A] = B$, then A is finite and $|A| = |B|$.
If $f[A] \subsetneq B$. By proposition 8.4.4, $f[A]$ is finite and $|A| = f[A] < |B|$ (as witnessed by the bijection $f \restriction A\colon f[A]$).
($\Rightarrow$) Fix $m, n \in \mathbb{N}$ such that $|A| = m \leq n = |B|$. We can construct an injection from A to B by

$$A \approx [m] \hookrightarrow [n] \approx B$$

□

Lemma 8.4.6. If A is nonempty and there is an injection from A to B, then there is a surjection from B to A. (8.6.6)

Proof. Suppose $f\colon A \to B$ is an injection, then by proposition 5.4.4, there is a function $g\colon B \to A$ such that $g \circ f = \mathrm{id}_A$. By proposition 5.4.5, g is a surjection. □

Lemma 8.4.7. Suppose A is a set and $B \subseteq \mathbb{N}$. For every surjection $g\colon B \to A$, there is an injection $f\colon A \to B$ such that $g \circ f = \mathrm{id}_A$. (8.6.8)

Proof. Define $f\colon A \to B$ by $f(a) = \min(g^{-1}(a))$. $g^{-1}(a)$ is nonempty because g is onto. By well ordering, $\min(g^{-1}(a))$ exists. So f is well defined.
To prove f is injective, suppose $f(a) = \min(g^{-1}(a)) = \min(g^{-1}(a')) = f(a')$. By definition, we have $f(a) \in g^{-1}(a)$ and $f(a) \in g^{-1}(a)$. Hence, $a = g(f(a)) = g(f(a')) = a'$ as desired. Finally, one can check $g \circ f = \mathrm{id}_A$. □

Proposition 8.4.8. If A and B are nonempty finite sets, then $|A| \leq |B|$ if and only if there is a surjection from B to A.

Proof. ($\Rightarrow$) Since $|A| \leq |B|$, by proposition 8.4.5, there is an injection from A to B. By lemma 8.4.6, there is a surjection from B to A.
($\Leftarrow$) Fix a surjection from B to A. Since B is finite, fix a bijection from $[n]$ to B. By composition (of the above two functions), we have a surjection from $[n]$ to [A]. Since $[n]$ is a subset of $\mathbb{N}$, it follows from lemma 8.4.7 that there is an injection from A to $[n]$. By proposition 8.4.5, $|A| \leq n = |B|$ as desired. □

8.5 Bounded versus Finite

Definition 8.5.1. *$A \subseteq \mathbb{R}$ is bounded in $\mathbb{R}$ if there are $u, l \in \mathbb{R}$ such that for all $a \in A$, $l \leq a \leq u$. $\max(A)$ is defined to be the number $m \in A$ such that for all $a \in A$, $a \leq m$.*

Remarks. Analogously, we can define "bounded ", "max" and "min" in $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, provided that we have total ordering.

Proposition 8.5.2. If A is a nonempty finite subset of $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$, then $\max(A)$ and $\min(A)$ exist. In particular, A is bounded. (Contrapositive: Every set which is unbounded in $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ is infinite.)

Proposition 8.5.3. In $\mathbb{N}$, every bounded set is finite. A subset of $\mathbb{N}$ is finite if and only if it is bounded in $\mathbb{N}$.

Proof. If b bounds X, then $X \subseteq [b+1]$. Since $[b+1]$ is finite, X is finite. □

Remarks. In other orders such as the standard ordering on $\mathbb{R}$, bounded sets may not be finite.

Proposition 8.5.4. Each of the following are sufficient for A to be infinite.

1. A is unbounded in $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$.

2. there is an injection from an infinite set to A.

3. A is equinumerous with a proper subset of itself.

See corollary 8.10.6 for applications

Proof. (1) is not necessary for A to be infinite. (A can be infinite but bounded.)
To prove (2): Suppose towards a contradiction that A is finite, X is infinite and there is an injection $f\colon X \to A$. Then $f[x] \subseteq A$. So $f[X]$ is finite, which means X is finite.
(2) is necessary because every infinite set is equinumerous to itself (in particular, $\mathrm{id}\colon A \to A$ is an injection.)
(3) is true by proposition 8.3.3. □

8.6 Countability

Definition 8.6.1. A set is countable if it is finite (i.e., equinumerous to some $[n]$) or equinumerous to $\mathbb{N}$. (When a set is equinumerous to $\mathbb{N}$, we say it is countably infinite.)

Proposition 8.6.2. $\mathcal{P}(\mathbb{N})$ is uncountable.

Proof. We shall show that $\mathcal{P}(\mathbb{N})$ is both infinite and not equinumerous to $\mathbb{N}$.
By Cantor's Theorem, $\mathbb{N} \not\approx \mathcal{P}(\mathbb{N})$. Also, there is an injection from $\mathbb{N}$ to $\mathcal{P}(\mathbb{N})$. Since $\mathbb{N}$ is infinite, so is $\mathcal{P}(\mathbb{N})$.

□

Proposition 8.6.3. A set A is countable if and only if there is an injection $f\colon A \to \mathbb{N}$.

Proof. ($\Rightarrow$) If $A \approx \mathbb{N}$, then there is a bijection from A to $\mathbb{N}$. If A is finite, then $A \approx [n] \hookrightarrow \mathbb{N}$, which yields an injection from A to $\mathbb{N}$.
($\Leftarrow$) Fix an injection $f\colon A \to \mathbb{N}$. If $\mathrm{range}(f)$ is finite, then $A \approx f[A] \approx \mathrm{range}(f)$ is finite. Otherwise, we construct a function $g\colon \mathbb{N} \to \mathrm{range}(f)$ by recursion on $n \in \mathbb{N}$.

$$g(n) = \min(\mathrm{range}(f) - \{g(m)\colon m < n\})$$

Fisrt, we prove g is well defined and one-to-one by strong induction.
Suppose for all $m < n$, $g(m)$ is well defined and $g(m) \neq g(m')$ for all $m' < m$.
Then, by the induction hypothesis, $g \restriction [n]\colon [n] \to \{g(m)\colon m < n\}$ is a bijection. Hence, $\{g(m)\colon m < n\}$ is finite. Since $\mathrm{range}(f)$ is infinite, $\mathrm{range}(f) - \{g(m)\colon m < n\}$ is nonempty. (Otherwise, $\mathrm{range}(f) \subseteq \{g(m)\colon m < n\}$, implying that $\mathrm{range}(f)$ is finite.) By well ordering, $\min(\mathrm{range}(f) - \{g(m)\colon m < n\})$ exists and it is not an element of $\{g(m)\colon m < n\}$. Hence, $g(n)$ is well defined and injective.

Second, we prove $g\colon \mathbb{N} \to \text{range}(f)$ is onto. Suppose towards a contradiction that $\text{range}(f) - \text{range}(g)$ is nonempty. Fix

$$l \in \min(\text{range}(f) - \text{range}(g))$$

Define the set $S = \{m \in \mathbb{N}\colon g(m) \geq l\}$ (to mirror the construction of g).
We claim S is nonempty. Otherwise $\text{range}(g) \subseteq [l]$, which means $\text{range}(g)$ is finite. However, since g is injective, we have $\mathbb{N} \approx \text{range}(g)$, contradicting the fact that $\mathbb{N}$ is infinite.
Fix any $n \in S$. Now we have

- $l \in \text{range}(f)$
- $l \notin \{g(m)\colon m < n\}$ (Since $l \notin \text{range}(g)$)
- $l < g(n)$ (Since $g(n) \geq l$ and $l \notin \text{range}(g)$)

Therefore, $l \in \text{range}(f) - \{g(m)\colon m < n\}$, but $l < g(n) = \min(\text{range}(f) - \{g(m)\colon m < n\})$, yielding a contradiction.
One can check that $g^{-1} \circ f$ is a bijection, so $A \approx \mathbb{N}$.

□

Corollary 8.6.4. Every subset of a countable set is countable. The intersection of countable sets is countable.

Proof. If $A \subseteq B$ and B is countable, then we can fix an injection $f\colon B \to \mathbb{N}$. Then $f \circ \iota\colon A \to \mathbb{N}$ is an injection. □

Corollary 8.6.5. If $X \subseteq \mathbb{N}$ is infinite, then $X \approx \mathbb{N}$.(see Corollary 8.10.6 for application)

Proof. This follows from the proof for ($\Leftarrow$) in proposition 8.6.3: Fix an infinite $X \subseteq \mathbb{N}$. We can construct a bijection $g\colon \mathbb{N} \to X$ by replacing $\text{range}(f)$ with X. Thus, $X \approx \mathbb{N}$. □

Proposition 8.6.6. A nonempty set A is countable if and only if there is a surjection from $\mathbb{N}$ to A.

Proof. ($\Rightarrow$) Since A is countable, by proposition 8.6.3, there is an injection from A to $\mathbb{N}$. By Lemma 8.4.6, there is a surjection form $\mathbb{N}$ to A. ($\Leftarrow$) By Lemma 8.4.7, there is an injection from A to $\mathbb{N}$. By proposition 8.6.3, A is countable. □

Proposition 8.6.7. The union of two countable sets is countable.

Proof. Suppose A and B are countable. If one of them is empty, then clearly their union is countable. Otherwise, fix (by proposition 8.6.6) surjections $f_1\colon \mathbb{N} \to A$ and $f_2\colon \mathbb{N} \to B$. Define $g\colon \mathbb{N} \to A \cup B$ by

$$g(n) = \begin{cases} f_1(\frac{n}{2}) & \text{if n is even} \\ f_2(\frac{n-1}{2}) & \text{if n is odd} \end{cases}$$

To prove that g is not onto, fix some $x \in A \cup B$.
If $x \in A$, there is some $n_1 \in \mathbb{N}$ such that $f_1(n_1) = x$. Then, $2n_1$ is even. So, $x \in \text{range}(g)$.
If $x \in B$, there is some $n_2 \in \mathbb{N}$ such that $f_2(n_2) = x$. Then, $2n_2 + 1$ is odd. So, $x \in \text{range}(g)$.
By proposition 8.6.6, $A \cup B$ is countable.

□

Exercise: Prove the above by considering injections

Remarks. By induction, the finite union of countable sets is countable. To prove that the countable union of countable sets is countable, we need the axiom of choice.

8.7 Size of $\text{Maps}([n], X)$

For each set X and each $n \in \mathbb{N}$, how many sequences of length n are there with values in X. Equivalently, what is the size of $\text{Maps}([n], X)$?

Example. $\text{Maps}([1], X) \approx X$. $\text{Maps}([2], X) \approx X \times X$ ($f \mapsto (f(0), f(1))$ is a bijection).

Lemma 8.7.1. For all sets A, B, X, Y, if $A \approx X$ and $B \approx Y$, then $A \times B \approx X \times Y$.

Proof. Fix bijections $f\colon A \to X$ and $g\colon B \to Y$. Consider $(a, b) \mapsto (f(a), g(b))$. One can check that this is a bijection. □

Lemma 8.7.2. $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$.

Proof. Define $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $f(a, b) = \frac{1}{2}(a + b)(a + b + 1) + b$. One can check that f is a bijection. (See Cantor's paring function). □

Proposition 8.7.3. If X is countably infinite, so is $X \times X$. If X and Y are countably infinite, then so is $X \times Y$.

Proof. By Lemma 8.7.1 and 8.7.2, $X \approx \mathbb{N} \approx \mathbb{N} \times \mathbb{N} \approx X \times X$ and $X \times Y \approx \mathbb{N} \times \mathbb{N} \approx \mathbb{N}$ □

Exercise: Disprove: For all sets A, B, X, Y, if $A \approx X$ and $B \approx Y$, then $A \cup B \approx X \cup Y$

Proof. A counterexample is $A = \{0, 1\}$, $X = \{2, 3\}$, $B = \{0\}$, $Y = \{4\}$. □

Proposition 8.7.4. For all sets X and all $n \in \mathbb{N}$, $\text{Maps}([n + 1], X) \approx \text{Maps}([n], X) \times X$. It follows that if X is countably finite, then for each $n \in \mathbb{N}^+$, $\text{Maps}([n], X)$ is countably finite.

Proof. Given $f \in \text{Maps}([n+1], X)$, map it to $(f \restriction [n], f(n)) \in \text{Maps}([n], X) \times X$. This mapping is one-to-one because if $f \neq f'$, then either $f(n) \neq f'(n)$ or $f \restriction [n] \neq f' \restriction [n]$. To show it is onto, take $(g, x) \in \text{Maps}([n], X) \times X$. Define $f(n) = x$ and $f(k) = g(k)$ for $k \in [n]$. One can check $f \in \text{Maps}([n+1], X)$.
By induction, if X is countably finite, then for each $n \in \mathbb{N}^+$, $\text{Maps}([n], X)$ is countably finite.

□

Remarks. Note that the Cartesian product is not associative.

Lemma 8.7.5. If $A \approx B$, then $\text{Maps}(A, X) \approx \text{Maps}(B, X)$.

Proof. Fix a bijection $f : B \to A$. Given $g \in \text{Maps}(A, X)$, map it to $g \circ f \in \text{Maps}(B, X)$.
To show this mapping is one-to-one, suppose $g \circ f = g' \circ f$. Then, $g(f(b)) = g'(f(b))$ for all $b \in B$. Since f is a surjection, we have $g(a) = g'(a)$ for all $a \in A$.
To show this mapping is onto, take $h \in \text{Maps}(B, X)$. Then, $g \circ f^{-1} \in \text{Maps}(A, X)$ and $h \circ f^{-1} \circ f = h$.

□

Proposition 8.7.6. If A is nonempty and finite, X is countably infinite, then $\text{Maps}(A, X)$ is countably infinite as well.

Proof. Fix $n \in \mathbb{N}$ such that $A \approx [n]$. By Lemma 8.7.5, $\text{Maps}([n], X) \approx \text{Maps}(A, X)$. By proposition 8.7.4, $\text{Maps}([n], X)$ is countably infinite, so $\text{Maps}(A, X)$ is countably infinite as well. □

Proposition 8.7.7. If X is countably infitnite, then $\text{Maps}(X, [2])$ is uncountable. It follows that if X and Y are countably infinite, then $\text{Maps}(X, Y)$ is uncountable.

Proof. First, we will construct a bijection f between $\mathcal{P}(X)$ and $\text{Maps}(X, [2])$ as follows: given $A \in \mathcal{P}(X)$, map it to $g \in \text{Maps}(X, [2])$ such that

$$g(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

This is clearly well defined. To prove it is one-to-one, suppose $f(A) = f(A') = g$. For each $x \in X$, $x \in A$ and $x \in A'$ iff $g(x) = 1$; $x \notin A$ and $x \notin A'$ iff $g(x) = 0$. Thus, $A = A'$.
To prove it is onto, fix $g \in \text{Maps}(X, [2])$. Define $A = \{x \in X : g(x) = 1\} \subseteq X$. One can check $f(A) = g$.
Thus, we have $\text{Maps}(X, [2]) \approx \mathcal{P}(X)$.
Second, we shall prove there is no surjection from X to $\mathcal{P}(X)$. Suppose towards a contradiction that we have a surjection from X to $\text{Maps}(X, [2])$. Then, by composing this surjection with the bijection $f : \text{Maps}(X, [2]) \to \mathcal{P}(X)$, we have a surjection from X to $\mathcal{P}(X)$, contradicting the proof of Cantor's Theorem. □

Notation 8.7.8. Let X be a **nonempty and countable** set. $X^{<\mathbb{N}}$ denote the set of all finite sequences with values in X, i.e., $X^{<\mathbb{N}} = \bigcup_{n\in\mathbb{N}} \text{Maps}([n], X)$

Proposition 8.7.9. $\mathbb{N}^{<\mathbb{N}}$ is countably infinite.

Proof. Since the set P of primes is an unbounded subset of $\mathbb{N}$, by proposition 8.5.3, it is an infinite subset of $\mathbb{N}$. Thus, we have $P \approx \mathbb{N}$. Fix a bijection $f\colon \mathbb{N} \to P$. Define a function from $\mathbb{N}^{<\mathbb{N}}$ to $\mathbb{N}$ by

$$(a_0, a_1, a_2, ..., a_n) \mapsto f(0)^{a_0+1} f(1)^{a_1+1} \cdots f(n)^{a_n+1}$$

This is injective because by FTA, the exponent on each prime number in the prime factorisation is unique. So by proposition 8.6.3, $\mathbb{N}^{<\mathbb{N}}$ is countable.
To prove $\mathbb{N}^{<\mathbb{N}}$ is infinite, observe that there is an injection from $\mathbb{N}$ to $\mathbb{N}^{<\mathbb{N}}$. (map each $n \in \mathbb{N}$ to the sequence n of length 1.). By proposition 8.5.4, $\mathbb{N}^{<\mathbb{N}}$ is infinite. □

Remarks.

1. Intuitively, the function f indexes every prime (in an ascending order).
2. We added 1 on each exponent to ensure that the function is one-to-one. For example, if we do not add 1, then (1) and (1, 0) are both mapped to 2.
3. Note that the function we defined is not onto. (e.g. Intuitively, 10 is not in the range).
4. Alternatively, one can prove $\mathbb{N}^{<\mathbb{N}}$ is countably infinite by adapting the proof for proposition 8.6.3. (But this is certainly more troublesom than just constructing another injection from N to $\mathbb{N}^{<\mathbb{N}}$).

Proposition 8.7.10. $X^{<\mathbb{N}}$ is countably infinite.

Proof. Since X is countably and nonempty, fix a surjection $g\colon \mathbb{N} \to X$. Define $h\colon \mathbb{N}^{<\mathbb{N}} \to X^{<\mathbb{N}}$ by

$$(a_0, a_1, a_2, ..., a_n) \mapsto (g(a_0), g(a_1), ..., g(a_n))$$

h is well-defined and onto because g is well defined and onto.
Since $\mathbb{N}^{<\mathbb{N}}$ is countably infinite, fix a bijection f from $\mathbb{N}$ to $\mathbb{N}^{<\mathbb{N}}$. Then, $h \circ f$ is a surjection from $\mathbb{N}$ to $X^{\mathbb{N}}$. By proposition 8.6.6, $X^{\mathbb{N}}$ is countable.
To prove $X^{\mathbb{N}}$ is infinite, notice that there is an injection from $\mathbb{N}$ to $X^{\mathbb{N}}$. (Fix $x_0 \in X$. Map each $n \in \mathbb{N}$ to the constant sequence x_0 with length n.) □

8.8 Axiom of Choice

Axiom 8.8.1. For every set X with $\emptyset \notin X$, there is a choice function $F\colon X \to \bigcup X$ such that for all $S \in X$, we have $F(S) \in S$

Remarks.

1. Intuitively, f is looking into every element S of X and then choose an element of S for us.
2. For finite X, Choice is provable in ZF. (By induction on n. Base case: $X = \{S\}$. since S is nonempty, there exists $x \in S$. Define $f(S) = x$.)
3. For X such that every $S \in X$ is a singleton, Choice is provable in ZF. (If S is finite, then Choice is not provable in ZF.)
4. For $X \subseteq \mathcal{P}(\mathbb{N})$, Choice is provable in ZF. (Since each $S \in X$ is a nonempty subset of $\mathbb{N}$, we can use well-ordering to choose the minimum element of each S.)
5. One should be concerned with the issue of Choice when there are **infinitely many** sets and we wnat to choose an element from each of them.

Proposition 8.8.2. A countable union of countable sets is countable.

Proof. Suppose $(A_i)_{i\in\mathbb{N}}$ is a sequence of countable sets. WLOG, assume each A_i is nonempty. Consider the set

$$X = \{\{g \in \text{Maps}(\mathbb{N}, A_i)\colon g \text{ is onto}\} \colon i \in \mathbb{N}\}$$

Since each A_i is nonempty and countable, $\{g \in \text{Maps}(\mathbb{N}, A_i)\colon g \text{ is onto}\} \neq \emptyset$ (by proposition 8.6.6). So $\emptyset \notin X$. By Choice, there is a function $F\colon X \to \bigcup X$ such that for each $i \in \mathbb{N}$, we have

$$F(\{g \in \text{Maps}(\mathbb{N}, A_i)\colon g \text{ is onto}\}) \in \{g \in \text{Maps}(\mathbb{N}, A_i)\colon g \text{ is onto}\}$$

Let $F(\{g \in \text{Maps}(\mathbb{N}, A_i)\colon g \text{ is onto}\}) = g_i$. Define $G\colon \mathbb{N}\times\mathbb{N} \to \bigcup_{i\in\mathbb{N}} A_i$ by

$$G(i, n) = g_i(n)$$

G is well defined because each g_i is well defined. To prove G is onto, fix any $a \in \bigcup_{i\in\mathbb{N}} A_i$. Then, there exists $i \in \mathbb{N}$ such that $a \in A_i$. Since g_i is onto, there exists $n \in \mathbb{N}$ such that $g_i(n) = a$. Thus, there is $(i, n) \in \mathbb{N}\times\mathbb{N}$ such that $G(i, n) = a$.
By composing a bijection from $\mathbb{N}$ to $\mathbb{N}\times\mathbb{N}$ with G, we obtain a surjection from $\mathbb{N}$ to $\bigcup_{i\in\mathbb{N}} A_i$, proving that $\bigcup_{i\in\mathbb{N}} A_i$ is countable. □

8.9 Cantor-Schroder-Bernstein

Notation 8.9.1. If there is an injection from A to B, we write $A \preceq B$. If there is an injection from A to B but $A \not\approx B$, we write $A \prec B$.

If $A \approx B$, then $A \prec B$ and $B \prec A$.

Proposition 8.9.2. If A and B are countable sets and $A \preceq B$ and $B \preceq A$, then $A \approx B$.

Proof. We consider two cases.
If A and B are both countably infinite, then $A \approx \mathbb{N} \approx B$ and we are done.
Otherwise, WLOG, suppose A is finite. Fix injections $f\colon A \to B$ and $g\colon B \to A$. Then the composition $g \circ f\colon A \to A$ is an injection. Since A is finite, by the Pigeonhole Principle, $g \circ f$ is onto. This means that for each $a \in A$, there is some $a' \in A$ such that $g(f(a')) = a$, i.e., there is some $b = f(a') \in B$ such that $g(b) = a$. So g is onto. Thus, $B \approx A$ as witnessed by the bijection g. □

Exercise: Prove that if A is finite and $B \preceq A$, then B is finite.

Proof. Fix an injection $f\colon B \to A$. Then $f[B] \subseteq A$. So $f[B]$ is finite. So B is finite as witnessed by the bijection $f \restriction B\colon B \to f[B]$.

□

Theorem 8.9.3. (*Cantor-Schroder-Bernstein*) For all sets A and B, if $A \preceq B$ and $B \preceq A$, then $A \approx B$.

Discussion. We fix injections $f\colon A \to B$ and $g\colon B \to A$. If f is onto, then we are done.
Otherwise, the set $B - \operatorname{range}(f)$ is nonempty. We wish to obtain a new injection $h\colon A \to B$ which includes $B - \operatorname{range}(f)$ in its range.
For each $b \in B - \operatorname{range}(f)$, let $h(g(b)) = b$. Now, each $b \in B - \operatorname{range}(f)$ is in the range of h.
One may attempt to define h by

$$h(a) = \begin{cases} g^{-1}(a) & \text{if } a \in g[B - \operatorname{range}(f)] \\ f(a) & \text{otherwise} \end{cases}$$

However, we notice that the sets $f[g[B - \operatorname{range}(f)]]$ and $f[A - g[B - \operatorname{range}(f)]]$ are disjoint since f is injective. Also notice that $\operatorname{range}(h) = (B - \operatorname{range}(f)) \cup f[A - g[B - \operatorname{range}(f)]]$. Therefore, for each $b \in B - \operatorname{range}(f)$, $f(g(b))$ is not in $\operatorname{range}(h)$.)
No matter, we define $h(g(f(g(b)))) = f(g(b))$ for each $b \in B - \operatorname{range}(f)$. Now, $f(g(b))$ is in $\operatorname{range}(h)$.
Repeat this for infinitely many times, $\operatorname{range}(h)$ will eventually cover the set B.

Proof. Fix injections $f\colon A \to B$ and $g\colon B \to A$. Define a sequence $(C_n)_{n\in\mathbb{N}}$ of subsets of A by recursion:

$$C_0 = g[B - \operatorname{range}(f)], \quad C_{n+1} = g[f[C_n]] \quad \text{for } n \in \mathbb{N}$$

Define $h\colon A \to B$ by

$$h(a) = \begin{cases} g^{-1}(a) & \text{if } a \in C_n \text{ for some } n \\ f(a) & \text{otherwise} \end{cases}$$

(Here, $h(a) = g^{-1}(a)$ means that for each $b \in C_{n-1}$, $h(g(b)) = b$. In other words, $g^{-1}(a)$ denotes the preimage of a under g.)
Nottice that for all n, $C_n \subseteq \operatorname{range}(g)$, and g is one-to-one. Thus, each $a \in C_n$ has a unique preimage under g. So h is well defined.

To prove h is one-to-one, suppose $h(a) = h(a')$. Consider the following cases:
Case 1: If $a, a' \notin \bigcup_{n\in\mathbb{N}} C_n$, then we are done (because f is one-to-one.)
Case 2: If $a \in \bigcup_{n\in\mathbb{N}} C_n$ and $a' \notin \bigcup_{n\in\mathbb{N}} C_n$. Then, we have $h(a) = g^{-1}(a) = h(a') = f(a')$. So $a = g(f(a'))$.
Case 2a: If $a \in C_0$, then $g(f(a')) \in C_0 = g[B - \text{range}(f)]$. Since g is one-to- one, we have $f(a') \in B - \text{range}(f)$, a contradiction.
Case 2b: If $a \in C_{n+1}$, then $g(f(a')) \in C_{n+1} = g[f[C_n]]$. Since g is one-to-one, we have $f(a') \in f[C_n]$. Since f is one-to-one, we have $a' \in C_n$, contradicting the assumption.
Therefore, in case 2, it is impossible to have $h(a) = h(a')$.
Case 3: If $a, a' \in \bigcup_{n\in\mathbb{N}} C_n$, then $g^{-1}(a) = g^{-1}(a')$. Since g is well defined, we have $a = a'$.

To prove h is onto, we first note that $\text{range}(h) = g^{-1}[\bigcup_{n\in\mathbb{N}} C_n] \bigcup f[A - \bigcup_{n\in\mathbb{N}} C_n]$. Clearly, $\text{range}(h) \subseteq B$. We shall prove $B \subseteq \text{range}(h)$. Fix $b \in B$, we consider two cases:
Case 1: If $g(b) \in \bigcup_{n\in\mathbb{N}} C_n$, then $b \in g^{-1}[\bigcup_{n\in\mathbb{N}} C_n] \subseteq \text{range}(h)$.
Case 2: If $g(b) \notin \bigcup_{n\in\mathbb{N}} C_n$, then $g(b) \in A - \bigcup_{n\in\mathbb{N}} C_n$ and $b \notin g^{-1}[\bigcup_{n\in\mathbb{N}} C_n]$. We want to prove $b \in f[A - \bigcup_{n\in\mathbb{N}} C_n]$, and it suffices to prove that $f^{-1}(b) \notin C_n$ for all n.
Fix any $k \in \mathbb{N}$. Observe that $g(b) \notin C_n$ for all $n \in \mathbb{N}$, and in particular, $g(b) \notin C_{k+1} = g[f[C_k]]$. Since g is injective, $b \notin f[C_k]$. Since f is injective, it follows that $f^{-1}(b) \notin C_k$. □

Remarks. In order to show $A \approx B$, it suffices to show that $A \preceq B$ and $B \preceq A$.

Corollary 8.9.4. If $A \approx C$ and $A \subseteq B \subseteq C$, then $A \approx B \approx C$.

Proof. Since $A \subseteq B$, $A \preceq B$ as witnessed by the inclusion function. Since $A \approx C$ and $B \subseteq C$, we can fix a bijection $f \colon C \to A$ and injection $\iota \colon B \to C$. Then $f \circ \iota \colon B \to A$ is an injection. Thus, $B \preceq A$. By Cantor-Schroder-Bernstein, $A \approx B$. Thus, $C \approx A \approx B$ □

8.10 Zorn's Lemma , Ideals and Comparability

Given any sets A and B, is it true that either $A \preceq B$ or $B \preceq A$? By previous results, this is true for countable sets.

Definition 8.10.1. (*Chain*) A set $\mathcal{C}$ is a chain if for every $X, Y \in \mathcal{C}$, either $X \subseteq Y$ or $Y \subseteq X$.

Lemma 8.10.2. (*Zorn's Lemma*) (Assuming Choice) Suppose $\mathcal{S}$ is a set such that for every chain $\mathcal{C} \subseteq \mathcal{S}$, we have $\bigcup \mathcal{C} \in S$. Then, there is some $M \in S$ which is maximal, i.e., for every $X \in \mathcal{S}$, M is not a proper subset of X.

Remarks.

1. Assuming ZF, Zorn's Lemma is equivalent to Choice. The proof of Zorn's Lemma from Choice is out of the scope of this course.

2. Note that ZL does not assert that for every $X \in \mathcal{S}$, $X \subseteq M$. $X \subseteq M$ means that $\forall x(x \in X \to x \in M)$. M is not a proper subset of X means that $(\exists m(m \in M \wedge m \notin X)) \vee (\forall x(x \in X \to x \in M))$. Certainly, $X \subseteq M$ implies M is not a proper subset of X, but the converse is not true.

3. Note that the empty set is a chain, and is also a subset of any set. If $\mathcal{S}$ satisfies the assumptions of ZL, then we have $\bigcup \emptyset = \emptyset \in \mathcal{S}$. Whenever we want to apply ZL to $\mathcal{S}$, we will not check $\emptyset \in \mathcal{S}$. This is immaterial since $\emptyset$ is never maximal unless $S = \{\emptyset\}$.

4. In general, this works for any partial order. (we can define maximal element in any partial order.) A partial order is a relation that is reflexive, anti-symmetric (if $a \leq b$ and $b \leq a$, then $a = b$.) and transitive. In partial order, we don't require any two elements to be comparable.

5. We can think of chains as sets that are closed under (binary)union, i.e., given $X, Y \in \mathcal{C}$, $X \cup Y \in \mathcal{C}$.

6. It is **false** that for finite $\mathcal{S}$, if M is maximal, then M has the max cardinality.
   One counterexample is $\mathcal{S} = \{\emptyset, \{1\}, \{0\}, \{1,2\}\}$. $\{0\}$ and $\{1,2\}$ are both maximal, but $\{0\}$ does not have max cardinality.
   It is also **false** that if M has maximal cardinality, then it is maximal. For example, consider $\mathcal{P}(\mathbb{N})$. $\mathbb{N} - \{0\} \in \mathcal{P}(\mathbb{N})$ has the max cardinality, but $\mathbb{N} - \{0\}$ is not maximal.

7. In general, it is false that the maximal element is the unary union of some $\mathcal{C} \subseteq \mathcal{S}$.

Example.

1. If $\mathcal{S} = \mathcal{P}(X)$ for some set X, then it satisfies ZL. We can take M to be X.

2. Fix sets A and B. Let $\mathcal{S}$ be the set of graphs of injections f such that $\mathrm{dom}(f) \subseteq A$ and $\mathrm{range}(f) \subseteq B$. Set theoretically,

   $$\mathcal{S} = \{f \in \mathcal{P}(A \times B) \colon ((a,b),(a,b') \in f \to b = b') \wedge ((a,b),(a',b) \in f \to a = a')\}$$

   where

   $$f = \{(a, f(a)) \colon a \in A\} = \{(a,b) \in A \times B \colon b = f(a)\}$$

   $\mathcal{S}$ satisfies the assumptions of ZL. Intuitively, the maximal element of $\mathcal{S}$ is our best attempt to construct an injection from A to B. In this case, the maximal element is not unique.

3. If $\mathcal{S}$ is the set of all finite subsets of $\mathbb{N}$, it does **not** satisfy the assumptions of ZL: Consider the chain $\mathcal{C} = \{[n] \colon n \in \mathbb{N}\} \subset \mathcal{S}$. We have $\bigcup \mathcal{C} = \mathbb{N} \notin \mathcal{S}$.

Definition 8.10.3. (*Ideal*) In a ring R, an ideal I is a nonempty subset of R such that:

1. if $a, b \in I$, then $a + b \in I$.

2. if $r \in R$ and $a \in I$, then $r \cdot a \in I$.

Trivially, R is an ideal itself. Other ideals are said to be proper.

Proposition 8.10.4. (Assuming Choice) In every ring R, there is a proper ideal M which is maximal, i.e., there is no proper ideal I such that $M \subsetneq I$.

Proof. Let $\mathcal{S}$ be the set of all proper ideals in R. By ZL, it suffices to prove that $\mathcal{S}$ satisfies the assumptions of ZL.
Fix any chain $\mathcal{C} \subseteq \mathcal{S}$. We claim that $\bigcup \mathcal{C} \in \mathcal{S}$, i.e., $\bigcup \mathcal{C}$ is a proper ideal.
To prove $\bigcup \mathcal{C}$ is proper ($\bigcup \mathcal{C} \neq R$), it suffices to show that $1 \notin \bigcup \mathcal{C}$. Suppose towards a contradiction that $1 \in \bigcup \mathcal{C}$, then 1 lies in some $I \in \mathcal{C}$. By the definition of ideal, this means that for each $r \in R$, we have $1 \cdot r = r \in I$. So $I = R$, contradicting the fact that I is a proper ideal.
To prove that $\bigcup \mathcal{C}$ is an ideal, suppose $a, b \in \bigcup \mathcal{C}$. Fix $I, J \in \mathcal{C}$ such that $a \in I$ and $b \in J$. WLOG, suppose $J \subseteq I$. So we have $a, b \in I$. Since I is an ideal, we have $a + b \in I$, implying $a + b \in \bigcup \mathcal{C}$. Similarly, for each $r \in R$, we have $r \cdot a \in I$, implying $I \in \mathcal{C}$. ($a + b$ and $r \cdot a$ are elements of element of $\mathcal{C}$).

□

Example. In $\mathbb{Z}$, a proper ideal is maximal if and only if it is the set of multiples of some prime. For example, the set of even integers is a maximal ideal, but the set of multiples of 6 is not maximal as it is strictly contained in the set of even integers.

Theorem 8.10.5. (*Comparability*) (Assuming Choice) For all sets A and B, either $A \preceq B$ or $B \preceq A$.

Proof. If $A \preceq B$, then we are done. Now, assume $A \not\preceq B$, we shall prove that $B \preceq A$.
Let $\mathcal{S}$ be the set of graphs of injections f such that $\mathrm{dom}(f) \subseteq A$ and $\mathrm{range}(f) \subseteq B$. Set theoretically,

$$\mathcal{S} = \{f \in \mathcal{P}(A \times B) \colon ((a, b), (a, b') \in f \to b = b') \wedge ((a, b), (a', b) \in f \to a = a')\}$$

Fix any chain $\mathcal{C} \subseteq \mathcal{S}$. We shall prove that $\bigcup \mathcal{C} \in \mathcal{S}$.
First, notice that every element of $\mathcal{C}$ is a subset of $A \times B$. So every element of element of $\mathcal{C}$ is an element of $A \times B$. Thus, $\bigcup \mathcal{C} \subseteq A \times B$.
We claim that $\bigcup \mathcal{C}$ is a function. Fix $(a, b), (a, b') \in \bigcup \mathcal{C}$, we shall prove $b = b'$. Fix $f, f' \in \mathcal{C}$ such that $(a, b) \in f$ and $(a, b') \in f'$. Since $\bigcup \mathcal{C}$ is a chain, WLOG, suppose $f' \subseteq f$. So $(a, b), (a, b') \in f$. We know that f is a function, so we must have $b = b'$.
The proof that $\bigcup \mathcal{C}$ is injective is analogous to the proof for that $\bigcup \mathcal{C}$ is well defined.
By ZL, fix $g \in \mathcal{S}$ such that g is maximal. Since $A \not\preceq B$, $\mathrm{dom}(g)$ must be a proper subset of A. Now we claim that $\mathrm{range}(g) = B$. If not, pick any $b \in B - \mathrm{range}(g)$ and any $a \in A - \mathrm{dom}(g)$. Then, we can extend g into a injection by mapping a to b, contradicting the maximality of

g (because $g \subsetneq g \cup \{(a, b)\}$).
It follows that g is a bijection from $\mathrm{dom}(g)$ to B. Therefore, $g^{-1}\colon B \to \mathrm{dom}(g)$ is a witness that $B \preceq A$.

□

Discussion. The proof uses the maximality of g in a crucial way. Intuitively, imagine we are constructing an injection from A to B, but get stuck somewhere because we have used up all elements of B. Now, this g we get is maximal, and by taking its inverse, we have a injection from B to A.

Corollary 8.10.6. A set A is infinite if and only if $N \preceq A$. (The forward direction assumes Choice)

Proof. ($\Rightarrow$) By Comparability, either $N \preceq A$ (in which case we are done) or $A \preceq \mathbb{N}$. If $A \preceq \mathbb{N}$, fix an injection $f\colon A \to \mathbb{N}$. Since $\mathrm{range}(f) \subseteq \mathbb{N}$ and $\mathrm{range}(f)$ is infinite, by proposition corollary 8.6.5, $\mathrm{range}(f) \approx \mathbb{N}$. Notice that since f is injective, $A \approx \mathrm{range}(f)$. Thus, $A \approx \mathbb{N}$, and in particular, $N \preceq A$.
($\Leftarrow$) This is true by proposition 8.5.4 □

Remarks. An interpretation of this result is that $\mathbb{N}$ is the smallest infinite set.

Definition 8.10.7. (*Dedekind Infinite*) A set is Dedekind infinite if it is equinumerous to a proper subset of itself.

Proposition 8.10.8. A set X is Dedekind infinite if and only if it is infinite. (The backward direction assumes Choice)

Proof. ($\Rightarrow$) This is a direct consequence of proposition 8.3.3.
($\Leftarrow$) By proposition 8.10.6, we can fix an injection $f\colon \mathbb{N} \to A$. Define a function $g\colon A \to A - \{f(0)\}$ by

$$g(a) = \begin{cases} a & \text{if } a \notin \mathrm{range}(f) \\ f(n+1) & \text{if } a = f(n) \end{cases}$$

g is well defined because f is well defined and **one-to-one**.
To prove g is one-to-one, suppose $g(a) = g(a')$. If $a, a' \notin \mathrm{range}(f) - \{f(0)\}$, then we have $a = a'$. If $a, a' \in \mathrm{range}(f) - \{f(0)\}$, fix n, n' such that $a = f(n)$ and $a' = f(n')$. Then $g(a) = g(a')$ implies $f(n+1) = f(n'+1)$. Since f is one-to-one, we have $n = n'$. Thus, $a = f(n) = f(n') = a'$. If $a \in \mathrm{range}(f)$ and $a' \notin \mathrm{range}(f) - \{f(0)\}$, then $g(a) \in \mathrm{range}(f) - \{f(0)\}$ and $g(a') \notin \mathrm{range}(f) - \{f(0)\}$). So we cannot have $g(a) = g(a')$.
To prove g is onto, fix any $a \in A - \{f(0)\}$. If $a \notin \mathrm{range}(f)$, then $g(a) = a$. If $a \in \mathrm{range}(f) - \{f(0)\}$, there is a unique $n \in \mathbb{N}^+$ such that $f(n) = a$ (since f is one-to-one). Let $a' = f(n-1)$. By definition, $g(a') = f(n) = a$.

□

Exercise. Prove that $[0, 1] \approx (0, 1)$.

Proof. Define $f\colon [0,1] \to (0,1)$ by

$$f(x) = \begin{cases} \frac{1}{2} & \text{if } x = 0 \\ \frac{1}{4}x & \text{if } x \in \{\frac{1}{2^n} : n \in \mathbb{N}\} \\ x & \text{otherwise} \end{cases}$$

One can check this is a bijection. □

Discussion. The proof for proposition 8.10.8 and the proof in the exercise use the same idea: we take an infinite sequence and shift every term of the sequence by one (or two). In the exercise, we shift $\frac{1}{2^n}$ to $\frac{1}{2^{n+2}}$. In the proposition, since the injection f gives us an infinite sequence in A, we shift $f(n)$ to $f(n+1)$.

Remarks. Assuming Choice, this gives us an equivalent way to define infinite sets.

Chapter 9 The Real Numbers

9.1 Construction of $\mathbb{R}$ via Dedekind Cuts

Definition 9.1.1. (*Dedekind cut*) A Dedekind cut is a nonempty set $C \subsetneq \mathbb{Q}$ such that

1. for all $q \in C$, if $r \in \mathbb{Q}$ and $r < q$, then $r \in C$ as well
2. C has no maximum, i.e., for all $q \in C$, there is some $q' \in C$ such that $q' > q$.

Remarks.

1. One should take note of the difference between supremum and maximum. The maximum of a set A must always be in A, whereas the supremum may not be an element of A.
2. C is countable because $\mathbb{Q}$ is countable. But the set of all Dedekind cuts is uncountable and is equinumerous to $\mathcal{P}(\mathbb{Q})$.
3. Notice that any set of Dedelind cuts forms a chain. (However, Zorn's Lemma is not useful in this case because for Dedekind cuts we have an ordering that satisfies trichotomy. If we want the maximal Dedekind cut, that will be the union of all Dedekind cuts in the chain, and we don't even need the conclusion of Zorn's Lemma. Zonr's Lemma is more likely to be useful when we have a set whose elements are not proper subsets of each other and we want to choose a maximal one.)
4. C may have a supremum in $\mathbb{R}$. It may or may noy have a supremum in $\mathbb{Q}$.

Definition 9.1.2. ($\mathbb{R}$) The set of all real numbers, denoted by $\mathbb{R}$, is the set of all Dedekind cuts. (Notice that $\mathbb{R} \subseteq \mathcal{P}(\mathbb{Q})$)

Remarks. Intuitively, a real number r is the set of all rational numbers strictly less that r.

Definition 9.1.3. (*Ordering on* $\mathbb{R}$) For Dedekind cuts $C, D \in \mathbb{R}$, $C < D$ if $C \subsetneq D$.

Proposition 9.1.4. The ordering on $\mathbb{R}$ satisfies trichotomy, i.e., for every two Dedekind cuts C and D, exactly one of the following holds:

1. $C = D$
2. $C \subsetneq D$
3. $D \subsetneq C$

Proof. If $C = D$, then it is clear that $C \subsetneq D$ and $D \subsetneq C$ do not hold.
Suppose that $C \neq D$. We shall prove exactly one of $C \subsetneq D$ and $D \subsetneq C$ holds.
Suppose towards a contradiction that none of the two holds, then there exists $q \in C$ such that $q \notin D$ and there exists $q' \in D$ such that $q' \notin C$. Since $q \notin D$, we must have $q \geq q'$. Similarly, we must have $q' \geq q$. Thus, $q = q'$, a contradiction.
It is not possible that both of $C \subsetneq D$ and $D \subsetneq C$ hold since if $C \subsetneq D$, then there exists $q \in C$ such that $q \notin D$, contradicting $D \subsetneq C$. □

9.2 Dedekind Completeness

Definition 9.2.1. $A \subseteq \mathbb{R}$ is bounded in $\mathbb{R}$ if there are $u, l \in \mathbb{R}$ such that for all $a \in A$, $l \leq a \leq u$. $\max(A)$ is defined to be the number $m \in A$ such that for all $a \in A$, $a \leq m$.

Theorem 9.2.2. (*Dedekind Completeness*) Every nonempty (upper) bounded subset A of $\mathbb{R}$ has a supremum $s \in \mathbb{R}$, i.e.,

1. for all $a \in A$, $a \leq s$ (s is an upper bound of A.)
2. for all $r < s$, there is some $a \in A$ such that $a > r$. (r is not an upper bound of A)

Equivalently, the set $\{s \in \mathbb{R} \colon (\forall a \in A)(a \leq s)\}$ has a minimum element.

Proof. We claim that $\sup(A) = \bigcup A$.
First, we shall prove $\bigcup A \in \mathbb{R}$, i.e., $\bigcup A$ is a Dedekind cut.
$\bigcup A$ is nonempty because A is nonempty and each element of A is nonempty.
Since A is bounded, there is some $D \in \mathbb{R}$ such that for each $C \in A$, we have $C \subseteq D \subsetneq \mathbb{Q}$.
Suppose $q \in \bigcup A$, $r \in \mathbb{Q}$ and $r < q$. Fix $C \in A$ such that $q \in C$. Since C is a Dedekind cut, we have $r \in C$. So $r \in \bigcup A$.
Suppose $q \in \bigcup A$. Fix $C \in A$ such that $q \in C$. Since C is a Dedekind cut, there is some $q' \in C$ such that $q' > q$. So we have $q' \in \bigcup A$ and $q' > q$.
Second, we shall prove $\bigcup A$ is the least upper bound of A.
By definition of union, for all $C \in A$, we have $C \subseteq \bigcup A$. So $\bigcup A$ is an upper bound of A.
Suppose $C \subsetneq \bigcup A$. Fix $q \in \bigcup A - C$, and fix $D \in A$ such that $q \in D$. Since $q \in D - C$, D is not a subset of C. By trichotomy, we have $C \subsetneq D$. So $\bigcup A$ is the least upper bound of A.
□

Remarks.

1. We denote the supremum of A by $\sup(A)$. Every $A \subseteq \mathbb{R}$ has a unique supremum.
2. $\sup(A)$ does not have to be an element of A. For example, the open interval $(0, 1)$ has supremum 1. Dedekind Completeness is powerful in the sense that no matter how complicated the subset of $\mathbb{R}$ is (e.g. the interval $(0, 1)$ with a lot of holes in it), the existence of the supremum is guaranteed.
3. To prove $\sup(A) = s$, first one needs to check s is an upper bound. Then, one can take an arbitrary $r < s$ and produce an $a > r$.

Proposition 9.2.3. There is some positive $r \in \mathbb{R}$ such that $r^2 = 2$

Proof. Consider the set $A = \{x \in \mathbb{R} \colon x^2 < 2\}$. We claim that $r = \sup(A)$ and $r^2 = 2$. We wish to prove that if $r^2 > 2$, it is not the least upper bound. If $r^2 < 2$, then it is not an upper bound. □

Proposition 9.2.4. $\mathbb{Q}$ is not Dedekind complete.

Proof. Consider the nonempty bounded set $S = \{q \in \mathbb{Q} \colon q^2 < 2\}$. We claim that if r is the supremum of S in $\mathbb{Q}$, then $r^2 = 2$. But there is not rational number whose square is 2, so S has no supremum in $\mathbb{Q}$. □

9.3 The Archimedean Property

Definition 9.3.1. (*The Archimedean Property*) In any ordered field F with additive identity 0_F and multiplicative identity 1_F, we call the set $\{0_F, 1_F, 1_F +_F 1_F, ...\}$ the set of natural numbers in F. F has the Archimedean property if the set of natural numbers in F is unbounded in F.

Proposition 9.3.2. The failure of the Archimedean property is equivalent to the existence of some infinitesimal x, i.e., $x > 0$ yet for all $n \in \mathbb{N}^+$, $x < \frac{1}{n}$.

Proof. The Archimedean property fails if and only if there is some $y > 0$ such that for all $n \in \mathbb{N}^+$, we have $0 < n < y$. For every $n \in \mathbb{N}^+$, $0 < n < y$ if and only if $0 < \frac{1}{y} < \frac{1}{n}$. So y witnesses the failure of the Archimedean property if and only if $\frac{1}{y}$ is a positive infinitesimal. □

Remarks. As lonng as we have a infinitesimal, then there are infinitely many infinitesimals.(divide by a positve integer repeatedly.)

Proposition 9.3.3. $\mathbb{Q}$ has the Archimedean property.

Proof. For each $\frac{m}{n} \in \mathbb{Q}$, we have $\frac{m}{n} < |m| + 1$. So $\mathbb{N}$ is unbounded in $\mathbb{Q}$. □

Proposition 9.3.4. $\mathbb{R}$ has the Archimedean property.

Proof. Suppose towards a contradiction that $\mathbb{N}$ is bounded in R. Since $\mathbb{N}$ is nonempty and bounded, $\sup(\mathbb{N}) \in \mathbb{R}$ exists.

Since $\sup(\mathbb{N}) - 1 < \sup(\mathbb{N})$, $\sup(\mathbb{N}) - 1$ is not an upper bound of $\mathbb{N}$. So there exists some $n \in \mathbb{N}$ such that $n > \sup(\mathbb{N}) - 1$.

Now $n + 1 \in \mathbb{N}$ and $n + 1 > \sup(\mathbb{N})$, contradicting the fact that $\sup(\mathbb{N})$ is an upper bound of $\mathbb{N}$.

□

Remarks. The proofs that $\mathbb{R}$ and $\mathbb{Q}$ have the Archimedean property does not follow from the axioms of ordered field. In fact, there are ordered fields where the Archimedean property fails.

Example. Not all ordered field have the Archimedean property. Here is an example:

9.4 The notions of density

Definition 9.4.1. The ordering on F is dense if for every two distinct $x, y \in F$, there is some $z \in F$ in between them.

Definition 9.4.2. We say $\mathbb{Q}$ is dense in $\mathbb{R}$ if for every two distinct $x, y \in \mathbb{R}$, there is $z \in \mathbb{Q}$ in between them.

Proposition 9.4.3. The standard ordering on $\mathbb{R}$ and $\mathbb{Q}$ are dense.

Proposition 9.4.4. The standard ordering in the set of irrationals is dense. (Note that the set of irrationals is not an ordered field.)

Proof. Given two distinct irrationals x, y, consider $z = \frac{x+y}{2}$. If z is irrational, then we are done. Otherwise, consider $w = \frac{z+x}{2}$. If w is irrational, then we are done. Otherwise, since w and z are two distinct rational numbers, there is an irrational number in between them (T1 Q5). □

Proposition 9.4.5. $\mathbb{Q}$ is dense in $\mathbb{R}$ if and only if $\mathbb{R}$ has the Archimedean property.

Proof. ($\Rightarrow$) Given $r \in \mathbb{R}$, since $r < r+1$ and $\mathbb{Q}$ is dense in $\mathbb{R}$, there is some rational $\frac{m}{n}$ such that $r < \frac{m}{n} < r+1$. Then, $|m| \geq \frac{m}{n} > r$.
($\Leftarrow$) Take $0 < x < y$ and $x, y \in \mathbb{R}$. By the Archimedean property, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < y - x$. Thus, $ny > 1 + nx$. By T10 Q3, there exists a positive integer $m = \lfloor nx \rfloor + 1$, and we have $nx < m < ny$. Thus, $x < \frac{m}{n} < y$ as desired. □