Contents

1 Section 1: Propositional Logic 3
   1.1 Propositions and Statements ............................... 3
   1.2 Connectives .................................................. 4
   1.3 Valid Arguments .............................................. 7
   1.4 Logical Identities .......................................... 11

2 Section 2: Set Theory 14
   2.1 Quantifiers ................................................... 14
   2.2 Elementary Set Theory ..................................... 15
   2.3 Common Sets .................................................. 17
   2.4 Operations on Sets ......................................... 18

3 Section 3: Equivalence Relations and Functions 23
   3.1 Orderings ..................................................... 23
   3.2 Equivalence Relations ....................................... 24
   3.3 Functions ..................................................... 27

Course Objectives

This course is meant to introduce you to various aspects of mathematics.

• You will be introduced to concepts in logic, set theory, number theory and abstract algebra.

• You will learn to read and produce formal proofs.

This course will conclude with a unit which applies what we have learned (in particular group theory) to coding theory. Although we won’t touch on them,
the number theory we learn has pertinent applications to cryptography (e.g., see Section 1.6 on public key cryptography in the textbook).
Section 1: Propositional Logic

Mathematics is founded on logic. The first, most basic form of logic is propositional logic (or symbolic logic) due to George Boole – the logic of \textbf{and} and \textbf{or} without quantifiers or mathematical objects. The next layer of logic is first (and higher-order) logic which involves quantifiers, sets and functions. We start with propositional logic.

1.1 Propositions and Statements

Propositional logic is needed to make very basic mathematical arguments. Mathematical propositions, like “7 is prime”, have definite truth values and are the building blocks of propositional logic. Connectives like “and”, “or” and “not” join mathematical propositions into complex statements whose truth depends only on its constituent propositions. You can think of these statements as polynomials in propositions which act like variables. We want to know when two statements are logically equivalent, or when one implies the other. That is, we want to learn how to reason with mathematical statements.

Consider the following statement:

\textit{If the budget is not cut, then a necessary and sufficient condition for prices to remain stable is that taxes will be raised. Taxes will be raised only if the budget is not cut. If prices remain stable, then taxes will not be raised. Hence, taxes will not be raised.}

Is this argument logically sound? That is, is the conclusion “taxes will not be raised” true if the premises of the statement are true? To answer this question, we need logic: “propositional calculus.”

\textbf{Definition 1.1.} A proposition is a sentence or assertion that is true (T) or false (F), but not both.

\textbf{Example 1.2.} Some (non-mathematical) propositions:

- \( p \) = “prices will remain stable”
- \( b \) = “the budget will be cut”
- \( r \) = “taxes will be raised”
**Definition 1.3.** A statement is one of two things:

- a proposition, or
- (two) statements joined by a connective.

The above is a “recursive definition” in that it defines a statement in terms of other statements.

**Example 1.4.** If \( p \) and \( q \) are propositions then \( p \) and \( q \) are also statements. If \( \wedge \) and \( \vee \) are connectives then \( p \wedge q \), \( p \wedge p \), \( p \vee q \), \( (p \wedge q) \vee p \), and so on are all statements. You should think of these as “logical polynomials in \( p \) and \( q \)”.

### 1.2 Connectives

Connectives (a.k.a. truth-functionals, or boolean operators) are functions\(^1\) that take one or more (say up to \( n \)) truth values and output a truth value: i.e., functions of the form

\[
f : \{T, F\}^n \rightarrow \{T, F\}.
\]

We now give a list of common connectives.

**Definition 1.5 (Negation).** Let \( p \) be a proposition (or statement). The negation of \( p \), denoted by \( \neg p \), is the denial of \( p \):

- If \( p \) is \( T \), then \( \neg p \) is \( F \).
- If \( p \) is \( F \), then \( \neg p \) is \( T \).

The definition of negation is summarized by the following truth table.

<table>
<thead>
<tr>
<th>( p )</th>
<th>( \neg p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

The negation or “not” gate is depicted by

\(^1\)We will later examine sets and functions more formally.
Definition 1.6 (Conjunction). Let $p$ and $q$ be two propositions. The **conjunction of $p$ and $q$**, denoted by $p \land q$, is another proposition whose truth values are defined by the following table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>$p \land q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Conjunction is also known as “and” and its gate is depicted by

Example 1.7. $r \land p =$ “taxes will be raised and prices will remain stable”

Definition 1.8 (Disjunction). Let $p$ and $q$ be two propositions. The **disjunction of $p$ and $q$**, denoted by $p \lor q$, is another proposition whose truth values are defined by the following table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>$p \lor q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Disjunction is also called “or” and its gate is depicted by

Definition 1.9 (Conditional). Let $p$ and $q$ be two propositions. The **conditional of $p$ and $q$**, denoted by $p \rightarrow q$, is another proposition whose truth values are defined by the following table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>$p \rightarrow q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>
In $p \rightarrow q$, $p$ is called the **antecedent** and $q$ is called the **consequent** of the conditional. The conditional operation can be thought of as “implies.” So $p \rightarrow q$ stands for “$p$ implies $q$,” “$p$ is a sufficient condition for $q$,” “if $p$, then $q$,” “$q$ is a necessary condition for $p$,” “$q$ if $p$” and “$p$ only if $q$.”

One may question why this is the correct truth-table for our intuitive notion of “implies” (particularly the cases where $p = F$). Here are some explanations:

- The way we use implication is through modus ponens: If we know $p \rightarrow q$ is true and we also know that $p$ is true, then we should be able to deduce that $q$ is true. This justifies the first line of the truth table.

- Next, consider what it means for $p \rightarrow q$ to be false. The only time that $p \rightarrow q$ should be false is if we have an instance where $p$ is true, but $q$ is false. We thus have the second line of the truth table.

- Instances where $p$ is false do not provide evidence that $p$ does not imply $q$. To further justify the third and fourth lines of the truth table, observe that one would expect the proposition $r \land s \rightarrow s$ to be always true; in this light, examining the truth table of $r \land s \rightarrow s$, we obtain that when the antecedent $r \land s$ is false no matter what the truth value of the consequent $s$ is (either true or false), we have a true value for $r \land s \rightarrow s$ (in this argument, $r \land s$ stands for $p$ and $s$ stands for $q$).

It is hoped you will find that the above explanations provide a convincing justification of the table. With experience working with implications, it may become more natural. For now, you can also treat this as simply a definition.

**Definition 1.10** (Biconditional). Let $p$ and $q$ be two propositions. The **biconditional of $p$ and $q$**, denoted by $p \leftrightarrow q$, is another proposition whose truth values are defined by the following table:

<table>
<thead>
<tr>
<th>$p$</th>
<th>$q$</th>
<th>$p \leftrightarrow q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

The biconditional operation can be thought of as “if and only if.”
The symbols \( \neg, \land, \lor, \rightarrow, \leftrightarrow \) are called connectives, truth functionals, boolean operators, among other things.

The converse of \( p \rightarrow q \) is \( q \rightarrow p \).

The inverse of \( p \rightarrow q \) is \( \neg p \rightarrow \neg q \).

The contrapositive of \( p \rightarrow q \) is \( \neg q \rightarrow \neg p \). An implication is “logically equivalent” to its contrapositive (see Definition 1.15 for the definition of “logical equivalence”).

### 1.3 Valid Arguments

**Example 1.11.**

*If the budget is not cut, then a necessary and sufficient condition for prices to remain stable is that taxes will be raised. Taxes will be raised only if the budget is not cut. If prices remain stable, then taxes will not be raised. Hence, taxes will not be raised.*

Let \( p \), \( b \) and \( r \) be the following propositions:

- \( p \) = “prices will remain stable”
- \( b \) = “the budget will be cut”
- \( r \) = “taxes will be raised”

We have the following premises:

- If the budget is not cut, then a necessary and sufficient condition for prices to remain stable is that taxes will be raised: \( \neg b \rightarrow (p \leftrightarrow r) \)
- Taxes will be raised only if the budget is not cut: \( r \rightarrow \neg b \)
- If prices remain stable, then taxes will not be raised: \( p \rightarrow \neg r \)

The conclusion is:

- Hence, taxes will not be raised: \( \neg r \).

Is this argument valid?

A statement is called a tautology (or logical identity) if it is always true (i.e., it is true for all possible truth-value assignments of its propositions).
Example 1.12. \( s = p \lor \neg p \) is a tautology.

\[
\begin{array}{c|c|c}
   p & \neg p & p \lor \neg p \\
   \hline
   T & F & T \\
   F & T & T \\
\end{array}
\]

A statement is called a **contradiction** (or fallacy) if it is always false.

Example 1.13. \( s = p \land \neg p \) is a contradiction.

\[
\begin{array}{c|c|c}
   p & \neg p & p \land \neg p \\
   \hline
   T & F & F \\
   F & T & F \\
\end{array}
\]

**Definition 1.14** (Logical Implication). Let \( s \) and \( q \) be two statement forms involving the same set of propositions. We say that \( s \) **logically implies** \( q \) and write \( s \Rightarrow q \) if whenever \( s \) is true, \( q \) is also true (i.e., if every assignment of truth values making \( s \) true also makes \( q \) true).

**Definition 1.15** (Logical Equivalence). Let \( s \) and \( q \) be two statement forms involving the same set of propositions. We say that \( s \) **logically equivalent** \( q \) and write \( s \Leftrightarrow q \) if both \( s \) and \( q \) have identical truth tables (i.e., for all truth assignments of their propositions).

Note that \( \Rightarrow \) and \( \Leftrightarrow \) are logical relationships between statements, while \( \rightarrow \) and \( \leftrightarrow \) are connective operators used to make statements.

**Example 1.16.**

- \( s = p \rightarrow p \) is a tautology.

\[
\begin{array}{c|c}
   p & p \rightarrow p \\
   \hline
   T & T \\
   F & T \\
\end{array}
\]

- If we suppose \( s = (p \rightarrow q) \rightarrow p \) and \( r = p \lor q \), then \( s \Rightarrow r \).

\[
\begin{array}{c|c|c|c|c|c}
   p & q & p \rightarrow q & s = (p \rightarrow q) \rightarrow p & r = p \lor q \\
   \hline
   T & T & T & T & T \\
   T & F & F & T & T \\
   F & T & T & F & T \\
   F & F & T & F & F \\
\end{array}
\]

Notice that whenever is \( s \) is true, \( r \) is also true. Thus, \( s \) logically implies \( r \).
Example 1.17. Consider the two statements $s = p \lor q$ and $r = \neg (p \leftrightarrow q)$. First note that $r$ has the same truth-table as a boolean operator called “exclusive or” (xor). The importance of xor is that it corresponds to binary addition.

Now, look at the truth tables below. For every truth-assignment that makes $r$ true, we also have that $s$ is true. Thus $r$ logically implies $s$; $r \Rightarrow s$.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$q$</th>
<th>$r = \neg (p \leftrightarrow q)$</th>
<th>$s = p \lor q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

Exercise 1.18.

- Show that $(p \rightarrow q) \iff \neg p \lor q$.
- Show that $(p \leftrightarrow q) \iff (\neg p \lor q) \land (\neg q \lor p)$.

Theorem 1.19. Let $s$ and $r$ be two statements in the same propositions.

(i) $s \iff r$ if and only if $s \leftrightarrow r$ is a tautology.

(ii) $s \Rightarrow r$ if and only if $s \rightarrow r$ is a tautology.

(iii) $s \iff r$ if and only if $s \Rightarrow r$ and $r \Rightarrow s$.

Proof.

(i) If $s \iff r$ then $s$ and $r$ take the same values for every truth-assignment to their propositions. Thus, for a given truth-assignment, we either have $s = r = T$ in which case $s \iff r$ is true, or $s = r = F$ in which case $s \iff r$ is true. Since $s \iff r$ is true for all truth assignments, it is a tautology. A similar argument works for the converse.

(ii) If $s \Rightarrow r$ then whenever a truth-assignment yields $s = T$, then the same truth assignment gives $r = T$. Thus, for a given truth assignment, $s$ is either false and hence $s \rightarrow r$ is true, or $s$ is true and consequently $r$ and $s \Rightarrow r$ are true. A similar analysis of the cases tells us that if $s \rightarrow r$ is a tautology then $s \Rightarrow r$. 
(iii) We know \( s \equiv r \) means that \( s \) and \( r \) always take the same truth-value on any assignment to their proposition. So, it’s clear that \( s \equiv r \) implies \( s \Rightarrow r \) and \( r \Rightarrow s \).

Now assume that both \( s \Rightarrow r \) and \( r \Rightarrow s \). Since \( s \Rightarrow r \), whenever \( s \) is true, \( r \) is also true. Whenever a truth assignment gives \( s = F \), then since \( r \Rightarrow s \), \( r \) must also be false. Thus \( s \) and \( r \) must have the same value for all truth-assignments. That is \( s \equiv r \).

\[ \square \]

**Definition 1.20.** An argument with premises \( p_1, \ldots, p_n \) and conclusion \( q \) is valid if \( p_1 \land \cdots \land p_n \Rightarrow q \).

**Example 1.21.** The argument in our motivating example was

\[
(\neg b \rightarrow (p \leftrightarrow r)) \land (r \rightarrow \neg b) \land (p \rightarrow \neg r) \Rightarrow \neg r.
\]

In order to show that it is a valid argument, it suffices to show that

\[
s = [(\neg b \rightarrow (p \leftrightarrow r)) \land (r \rightarrow \neg b) \land (p \rightarrow \neg r)] \rightarrow \neg r
\]

is a tautology. We can do this by examining all \( 2^3 = 8 \) possible values for \( b, p \) and \( r \), and verify that, on each truth-assignment, \( s \) is true.

Here’s a second approach: We will prove that \( s \) is a tautology by contradiction. In a proof by contradiction, you make an unfounded assumption and look at the logical implications of that assumption. If, using that assumption, we can show something absurd (i.e., false), then we will know that the assumption is false.

In our case, we’ll assume that \( s \) is not a tautology. If we can prove a false statement, then we’ll know that our assumption was wrong: we’ll have proven that \( s \) is a tautology.

Assume that \( s \) is not a tautology. Let

\[
q_1 = \neg b \rightarrow (p \leftrightarrow r)
\]

\[
q_2 = r \rightarrow \neg b
\]

\[
q_3 = p \rightarrow \neg r
\]

so that \( s = (q_1 \land q_2 \land q_3) \rightarrow \neg r \). Since \( s \) is not a tautology, there must be a truth-assignment making \( \neg r = F \) and \( q_1 = q_2 = q_3 = T \). Since \( \neg r = F \),
we know $r = T$ in our truth-assignment. Further, since $q_3 = T$, we have $T = p \rightarrow \neg r = p \rightarrow F$ and therefore, we must have $p = F$. As $T = q_2 = r \rightarrow \neg b$ and $r = T$, we conclude that $\neg b = T$ and hence $b = F$.

So, we’ve determined the truth assignment that makes $(q_1 \land q_2 \land q_3) \rightarrow \neg r = F$; it’s $b = F$, $p = F$ and $r = T$. However, with this truth assignment, $q_1 = T \rightarrow (F \leftrightarrow T) = T \rightarrow F = F$ which shows that $s = T$. This contradicts our choice of truth-assignment as one that makes $s = F$. Thus our assumption must have been wrong: $s$ is a tautology. Consequently, the original argument is valid.

1.4 Logical Identities

The logical identities listed in the theorem below allow you to manipulate a statement into another form that is logically equivalent to the original.

**Theorem 1.22.** The following logical identities hold for all statements $p, q, r$.

- $p \land p \Leftrightarrow p$ \hspace{1cm} Idempotence
- $p \lor p \Leftrightarrow p$
- $p \land \neg p \Leftrightarrow F$ \hspace{1cm} Contradiction
- $p \lor \neg p \Leftrightarrow T$ \hspace{1cm} Tautology
- $p \land F \Leftrightarrow F$
- $p \land T \Leftrightarrow p$
- $p \lor T \Leftrightarrow T$
- $p \lor F \Leftrightarrow p$
- $p \land q \Leftrightarrow q \land p$ \hspace{1cm} Commutativity
- $p \lor q \Leftrightarrow q \lor p$
- $p \land (q \lor r) \Leftrightarrow (p \land q) \land r$ \hspace{1cm} Associativity
- $p \lor (q \land r) \Leftrightarrow (p \lor q) \lor r$
- $\neg(p \land q) \Leftrightarrow (\neg p) \lor (\neg q)$ \hspace{1cm} DeMorgan’s Laws
- $\neg(p \lor q) \Leftrightarrow (\neg p) \land (\neg q)$
- $\neg \neg p \Leftrightarrow p$ \hspace{1cm} Double Negation
- $p \land (q \lor r) \Leftrightarrow (p \land q) \lor (p \land r)$ \hspace{1cm} Distributivity
- $p \lor (q \land r) \Leftrightarrow (p \lor q) \land (p \lor r)$
\[(p \rightarrow q) \iff (\neg q \rightarrow \neg p)\]
Contrapositive
\[p \land (p \lor q) \iff p\]
Absorption
\[p \lor (p \land q) \iff p\]

**Proof.** Write out the truth tables replacing \(\iff\) with \(\leftrightarrow\) and check that you get a tautology in each case. \(\square\)

**Example 1.23.** The logic circuit for the statement \((a \land c) \lor [\neg a \lor b]\) has three inputs \((a, b, c)\) and one output (the value of the statement) and four gates \((\land, \lor, \neg, \lor)\).

We can find a logically equivalent statement with fewer gates:
\[
(a \land c) \lor [\neg a \lor b] \iff [(a \land c) \lor (\neg a)] \lor b \quad \text{using associativity}
\]
\[
\iff [(\neg a) \lor (a \land c)] \lor b \quad \text{using commutativity}
\]
\[
\iff [(\neg a \lor a) \land (\neg a \lor c)] \lor b \quad \text{using distributivity}
\]
\[
\iff (\top \land (\neg a \lor c)) \lor b \quad \text{using tautology}
\]
\[
\iff [\neg a \lor c] \lor b \quad \text{using the sixth identity}
\]

This abruptly ends our discussion of propositional logic. We’ve learned about propositions, statements, and connectives. We have one theorem which relates the conditional and biconditional connectives to logical implication and logical equivalence. This theorem lets us prove statements of the form \(s \Rightarrow r\) by showing that \(s \rightarrow r\) is a tautology – something that can be checked in the truth-table. We’ve also given an example of how one can prove a statement is tautological without using the truth-table (see Example 1.21). Finally, we have a list of rules for manipulating statements into other logically equivalent statements. Using the distributive property, we see that we can expand statements easily, but that logically equivalent statements, even when expanded, can take different forms.

In our last example, Example 1.23, we suggest that we might be able to factor statements to get them into some minimal form. This is, in fact, a difficult task. You might want to read wikipedia’s articles on circuit minimization and Karnaugh maps.

The breakthrough connection between electrical switching circuits and symbolic logic was made by Claude Elwood Shannon in 1938, playing a catalyst role in the digital revolution. Shannon went on in the late 1940s to found the field of information theory, making him regarded as the father of the information age.
As a final note, I suggest you read Section 3.3 of the textbook. It contains a lot of practical advice on how to prove statements in mathematics. This is a skill that we’ll try to hone throughout the rest of this course.
Section 2: Set Theory

Throughout, the symbols “|” and “:” will both stand for “such that” and will be used interchangeably.

2.1 Quantifiers

There is one intermediary topic that we need to address as we pass from propositional logic to set theory – the topic of quantifiers. There are two quantifiers used in mathematics:

- \( \exists \), there exists.
- \( \forall \), for all.

Much like how propositional logic takes a narrow view of the mathematical world, logicians often choose to work in restricted regions of mathematics. For instance, someone studying arithmetic may assume that all variables represent positive integers. Thus a statement like

\[ \exists x, \forall y, \text{ } y \text{ divides } x \Rightarrow (y = 1 \lor y = x) \]

which is interpreted as, “there exists an integer \( x \) such that every integer dividing \( x \) is either 1 or \( x \).” In other words, this statement asserts the existence of a prime number. You’ll note that the order of \( \exists \) and \( \forall \) is important: the statement

\[ \forall y, \exists x, \text{ } y \text{ divides } x \Rightarrow (y = 1 \lor y = x) \]

says that “for every integer \( y \), there is some integer \( x \) which, if divisible by \( y \), means \( y \) was 1 or equal to \( x \).” While not very useful, this statement is true. Given an integer \( y \), we could simply pick \( x = y \) and the implication will be satisfied. For each \( y \), we could also pick \( x \) such that \( y \) does not divide \( x \), and this will also satisfy the implication.

In your real analysis course, you’ll learn the difference between continuity and uniform continuity. In these definitions, a subtle change in the order of the quantifiers makes a significant difference.

Unlike the examples above, we won’t restrict our variables to integers. Our variables will be allowed to be of any mathematical type, so we often need to specify extra conditions on a given variable. Let \( \mathbb{N} = \{0, 1, 2, \ldots \} \) be the set
of natural numbers. Here’s how the existence of a prime number reads if we’re working in all of mathematics:

$$\exists x, x \in \mathbb{N} \Rightarrow \forall y, y \in \mathbb{N} \Rightarrow [y \text{ divides } x \Rightarrow (y = 1 \lor y = x)].$$

Here $x$ is allowed to range over all polynomials, sets, integers, and so on. If it happens to be a natural number then the remaining part of the statement is required to be true as well. For brevity, we often drop the $\Rightarrow$’s in favour of some notation before the comma:

$$\exists x \in \mathbb{N}, \forall y \in \mathbb{N}, y \text{ divides } x \Rightarrow (y = 1 \lor y = x).$$

Finally, there is a De Morgan’s law for quantifiers:

$$\neg(\exists x, P(x)) \iff \forall x, \neg P(x)$$

$$\neg(\forall x, P(x)) \iff \exists x, \neg P(x).$$

### 2.2 Elementary Set Theory

Set theory is another crucial pillar in the foundation of mathematics. It would be very easy to say that a set is a collection of mathematical objects and leave it at that. But then you’d never be able to address Russell’s paradox:

Let $X$ be the set of all sets which do not contain themselves:

$$X = \{Y \mid Y \notin Y\}.$$  Is $X$ a member of itself? If it is, then it shouldn’t be. If it’s not, then it should.

To resolve this paradox, there are very strict conditions on which sets we can form. The rules for building sets are given in a list of axioms referred to as ZFC (Zermelo-Fraenkel-Choice). Before we get into the list of axioms, some notation:

(i) $x \in X$ means $x$ is in the set $X$. We also say $x$ is an element of $X$ or a member of $X$.

(ii) $\emptyset$ is the empty set; this set is assumed to exist. It has the property that $\forall x, \neg(x \in \emptyset)$. By De Morgan’s law for quantifiers, we can restate this as $\neg(\exists x, x \in X)$. 
(iii) $X \subseteq Y$ is an abbreviation for $\forall x, x \in X \Rightarrow x \in Y$. The notation $X \subset Y$ means $X \subseteq Y \land X \neq Y$ and we say $X$ is a proper (or strict) subset of $Y$. Mathematicians often sloppily use $\subset$ when they mean $\subseteq$ (the audacity!), so be cautious.

**Example 2.1.** Let $X = \{1, 2, \{3\}\}$.

| $1 \in X$ | $\{1\} \subseteq X$ |
| $2 \in X$ | $\{1, 2\} \subseteq X$ |
| $3 \notin X$ | $\{1, 2, 3\} \not\subseteq X$ |
| $\{3\} \in X$ | $\{\{3\}\} \subseteq X$ |

In the following, the most readable and important axioms are starred.

(1*) Extensionality.

$$\forall X, \forall Y, (\forall Z, Z \in X \Leftrightarrow Z \in Y) \Rightarrow X = Y.$$  

Two sets are equal if they contain the same elements. In order to prove two sets are equal, we often break the argument into two steps – we show $X \subseteq Y$ and $Y \subseteq X$. An argument proving $X \subseteq Y$ would go

“Let $x \in X$. Yada, yada, yada. Therefore $x \in Y$.”

(2) Regularity.

$$\forall X, (\exists a, a \in X) \Rightarrow (\exists Y, Y \in X \land \neg(\exists Z, Z \in Y \land Z \in X)).$$

Every non-empty set is disjoint (i.e., has no common elements) from one of its members. In other words, if $X$ is a non-empty set, then for some $Y \in X$, $Y$ has no elements in common with $X$.

(3*) Subsets. For all statements $\phi$,

$$\forall X, \exists Y, \forall x, x \in Y \Leftrightarrow (x \in X \land \phi(x)).$$

Given any set, you can restrict that set to a subset. Notationally, we express $Y$ as

$$Y = \{x \in X \mid \phi(x)\}.$$  

This axiom is needed to avoid paradoxes such as the one by Russell.
(4) **Pairing.**
\[ \forall x, \forall y, \exists Z, x \in Z \land y \in Z. \]
Given any sets \( x \) and \( y \), there is a set \( Z = \{x, y\} \) which contains both.

(5*) **Union.**
\[ \forall F, \exists Z, \forall Y, \forall x, (x \in Y \land Y \in F) \Rightarrow x \in Z. \]
Given any collection of sets \( F \), there is a set that contains their union. When \( F \) contains two sets \( X \) and \( Y \), we write their union as
\[ Z = X \cup Y = \{x \mid x \in X \lor x \in Y\}. \]

(6) **Replacement.** For all statements \( f \),
\[ \forall X, \exists Y, \forall x, x \in X \Rightarrow (\exists y, y \in Y \land y = f(x)). \]
The set \( Y \) is denoted by \( \{y : \exists x \in X, y = f(x)\} \) and is called the "image of \( X \) under \( f \)" (the axiom states that "the image of a set under a function is a set").

(7*) **Axiom of Infinity.**
\[ \exists X, \emptyset \in X \land (\forall y, y \in X \Rightarrow y \cup \{y\} \in X) \]
There is a set \( X \) such that \( \emptyset \in X \) and whenever \( y \in X \), then \( y \cup \{y\} \in X \) (such a set is called a "successor set"). In other words, this axiom states that there exists a set whose elements have the recursive structure of the natural numbers. Recall \( \mathbb{N} = \{0\} \cup \{x + 1 \mid x \in \mathbb{N}\} \).

(8*) **Power set.**
\[ \forall X, \exists Y, \forall Z, Z \subseteq X \Rightarrow Z \in Y. \]
Given a set \( X \), there is a set denoted \( Y = \mathcal{P}(X) \) which contains all subsets of \( X \).

(9) **The Axiom of Choice.** If \( X \) is a set of non-empty pairwise disjoint sets, then there is a set \( Y \) which has exactly one element in common with each element of \( X \).

### 2.3 Common Sets

The convention of most mathematicians is to use "blackboard bold" characters for the most commonly used sets of numbers.
Definition 2.2.

- $\mathbb{N} = \{0, 1, 2, \ldots\}$, natural numbers.
- $\mathbb{Z} = \{\ldots, -2, 1, 0, 1, 2, \ldots\}$, integers.
- $\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z}, b \neq 0\}$, rational numbers.
- $\mathbb{R}$, real numbers.
- $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}, i = \sqrt{-1}\}$, complex numbers.

The following hold:

$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

2.4 Operations on Sets

You’ve probably been told that sets are unordered and don’t respect repetitions. For example $\{1, 2, 3\} = \{3, 2, 1\} = \{2, 3, 3, 1, 3\}$. The reason that sets are unordered, is that there is essentially only one question you can ask of a set $X$:

"Is $x \in X$?"

You can’t ask how many times $x$ is in $X$, nor whether it precedes or follows another element. Since the “three” sets above all give the same answer to the questions $1 \in X$, $2 \in X$, $3 \in X$ (all true) and $x \in X$ false for any other $x$, we see that they are all the same sets. We can’t distinguish between them using membership ($\in$).

We now define operations on sets. The only tools at are disposal are logical operators, membership ($\in$), and anything provided in the ZFC axioms (restriction to subsets, power set, union).

Definition 2.3. Let $X$ and $Y$ be sets.

- $X \cup Y = \{x \mid x \in X \lor x \in Y\}$, union.
- $X \cap Y = \{x \mid x \in X \land x \in Y\} = \{x \in X \mid x \in Y\}$, intersection.
- $X \setminus Y = \{x \in X \mid x \notin Y\}$, set difference.
• If \( Y \subseteq X \), then we sometimes write \( Y^c = X \setminus Y \) for the complement of \( Y \) in \( X \). Our textbook calls \( X \) the “universal set”. Often this notation is used without explicitly introducing \( X \). You should ask what set the complement is taken in if its not clear from context.

• \( X \triangle Y = (X \cup Y) \setminus (X \cap Y) \), symmetric difference.

• \( \mathcal{P}(X) = \{ Y \mid Y \subseteq X \} \), the power set of \( X \).

• \( X \times Y = \{(x,y) \mid x \in X \land y \in Y\} \), Cartesian product.

• \( X^n = \underbrace{X \times \cdots \times X}_{n\text{-times}} = \{(x_1, \ldots, x_n) \mid \forall i, 1 \leq i \leq n \Rightarrow x_i \in X\} \).

• \( Y^X = \{ f : X \to Y \} \).

Example 2.4. Let \( X = \{1, 2, 3, 4, 5\} \), \( Y = \{3, 4, 5\} \) and \( Z = \{2, 3, 6\} \) be subsets of \( \mathbb{N} \).

\[
\begin{align*}
Y \cup Z &= \{2, 3, 4, 5, 6\} \\
X \cup Y &= X \\
X \cap Z &= \{2, 3\} \\
X \cap Y &= Y \\
X \setminus Z &= \{1, 4, 5\} \\
Y \setminus X &= \emptyset \\
Y^c &= \{n \in \mathbb{N} \mid n \leq 2\} \cup \{n \in \mathbb{N} \mid n \geq 6\} \\
X \triangle Z &= \{1, 4, 5, 6\} \\
\mathcal{P}(Y) &= \{\emptyset, \{3\}, \{4\}, \{5\}, \{3, 4\}, \{3, 5\}, \{4, 5\}, \{3, 4, 5\}\} \\
Y \times Z &= \{(3, 2), (3, 3), (3, 6), (4, 2), (4, 3), (4, 6), (5, 2), (5, 3), (5, 6)\} \\
Y^3 &= \{(3, 3, 3), (3, 3, 4), (3, 3, 5), (3, 4, 3), (3, 4, 4), (3, 4, 5), (3, 5, 3), \ldots, (5, 5, 5)\} \\
Z^X &= Z^5 \quad \text{or, perhaps, these are only “in correspondence.”}
\end{align*}
\]

In the example above, \( X, Y \) and \( Z \) are finite sets. That is to say, the number of distinct elements in these sets is given by a natural number (rather than some “infinite cardinal”). When a set \( X \) is finite, we use \( |X| \) to denote its size. For the sets above \( |X| = 5 \) and \( |Y| = |Z| = 3 \). We will define finite and infinite more carefully when we talk about bijective functions.

Two sets \( X \) and \( Y \) are called disjoint if \( X \cap Y = \emptyset \). A collection of sets \( X_1, \ldots, X_n \) is pairwise disjoint if for each pair of indices \( i \) and \( j \) with \( i \neq j \),
we have $X_i \cap X_j = \emptyset$. Equivalently, using the contrapositive, $X_1, \ldots, X_n$ are pairwise disjoint if $X_i \cap X_j \neq \emptyset$ implies $i = j$. If $X_1, \ldots, X_n$ are pairwise disjoint then
\[ |X_1 \cup \cdots \cup X_n| = |X_1| + \cdots + |X_n|. \]

If your sets are not disjoint, you can use the following theorem relating the size of a union of sets to the sizes of their intersections.

**Theorem 2.5** (Inclusion-exclusion). Let $X, Y$ and $Z$ be sets.
\[ |X \cup Y| = |X| + |Y| - |X \cap Y| \]
\[ |X \cup Y \cup Z| = |X| + |Y| + |Z| - |X \cap Y| - |X \cap Z| - |Y \cap Z| + |X \cap Y \cap Z| \]

The following theorem lists the basic identities that these set operations satisfy.

**Theorem 2.6.** For any sets $X$, $Y$ and $Z$ (all contained in some “universal set” $U$) we have
\[
\begin{align*}
X \cap X &= X & \text{idempotence} \\
X \cup X &= X \\
X \cap X^c &= \emptyset \\
X \cup X^c &= U & \text{complementation} \\
X \cap Y &= Y \cap X & \text{commutativity} \\
X \cup Y &= Y \cup X \\
X \cap (Y \cap Z) &= (X \cap Y) \cap Z & \text{associativity} \\
X \cup (Y \cup Z) &= (X \cup Y) \cup Z \\
(X \cap Y)^c &= X^c \cup Y^c & \text{De Morgan laws} \\
(X \cup Y)^c &= X^c \cap Y^c \\
X \cap (Y \cup Z) &= (X \cap Y) \cup (X \cap Z) & \text{distributivity} \\
X \cup (Y \cap Z) &= (X \cup Y) \cap (X \cup Z) \\
(X^c)^c &= X & \text{double complement} \\
X \cap \emptyset &= \emptyset \\
X \cup \emptyset &= X & \text{properties of the empty set}
\end{align*}
\]
\[ X \cap U = X \]
\[ X \cup U = U \]  \hspace{1cm} \text{properties of the universal set}
\[ X \cap (X \cup Y) = X \]
\[ X \cup (X \cap Y) = X \]  \hspace{1cm} \text{absorption laws}

**Proof.** We will only prove one of the above. Specifically, we prove that \( X \cap U = X \) assuming \( X \subseteq U \).

In order to show that the two sets \( X \cap U \) and \( X \) are equal, we proceed by double inclusion. That is, we first show \( X \cap U \subseteq X \) and later show \( X \subseteq X \cap U \). This suffices to prove \( X \cap U = X \).

For our first containment, \( X \cap U \subseteq X \), take an arbitrary element \( x \in X \cap U \).

From the definition of intersection, we know \( x \in X \) and \( x \in U \). Since we have \( x \in X \) and \( x \) was an arbitrary element of \( X \cap U \), we then have that \( X \cap U \subseteq X \).

For our second containment, \( X \subseteq X \cap U \), take an arbitrary element \( x \in X \).

We need to show that \( x \in X \cap U \). Since \( X \subseteq U \), we know that any element of \( X \) is an element of \( U \). In particular, \( x \in X \) so \( x \in U \). Thus, \( x \in X \) and \( x \in U \). Therefore \( x \in X \cap U \). We conclude that \( X \subseteq X \cap U \).

Since we have shown both \( X \cap U \subseteq X \) and \( X \subseteq X \cap U \), we must have \( X \cap U = X \). \( \Box \)

Using double inclusion is the most common way to show that two sets are equal. When you write such a proof, you’ll want to clearly mention both inclusions and why each holds. It is very common for one direction to be significantly harder than the other. The previous proof was meant to introduce you to this technique, and so, was excessively verbose. Here’s how one can write the proof more concisely:

**Theorem 2.7.** If \( X \subseteq U \) then \( X \cap U = X \).

**Proof.** Show that \( X \cap U = X \) by double inclusion. Take \( x \in X \cap U \), arbitrarily. Since \( x \in X \cap U \), \( x \in X \). Therefore \( X \cap U \subseteq X \). For the opposite inclusion, take \( x \in X \). Since \( X \subseteq U \), we have \( x \in U \) as well. Therefore \( x \in X \cap U \) and consequently \( X \subseteq X \cap U \). Since we have proven both containments, we know \( X \cap U = X \). \( \Box \)

We’ve already introduced Cartesian products of sets \( X^n \) whose elements are \( n \)-tuples (pairs, triples, quadruples, quintuples, etc., depending on \( n \)). An \( n \)-tuple \( (x_1, x_2, \ldots, x_n) \in X^n \) is an ordered sequence of elements \( x_i \in X \). You can
also think of it as a function \( \{1, \ldots, n\} \rightarrow X \) which assigns \( i \mapsto x_i \). (There’s really no other information in the function other than a choice of output \( x_i \) for each input \( i \).

We generalize this idea: a **family** of elements of \( X \) is an indexed collection \((x_i)_{i \in A}\) where \( A \) is our index set and each \( x_i \in X \). If \( A = \{1, \ldots, n\} \) then our family is simple an \( n \)-tuple \((x_1, \ldots, x_n)\). If \( A = \mathbb{N} \) then our family \((x_i)_{i \in \mathbb{N}}\) is a sequence \( x_0, x_1, x_2, \ldots \) and so on. Our index set may be more exotic as well.
Section 3: Equivalence Relations and Functions

Definition 3.1. If $X$ and $Y$ are sets, then a **binary relation** from $X$ to $Y$ is a subset $R \subseteq X \times Y$. Whenever $(x, y) \in R$, we write $xRy$ and say that “$x$ is related to $y$ under $R$.”

Quite often $X$ and $Y$ will be the same set. In this case, we simply say that $R$ is a relation on $X$. Relations are used to mathematically represent orderings by size, divisibility, or containment. They are also used to group objects together and form equivalences between objects. Finally, functions are a specific type of relation.

Example 3.2.

(i) Let $X = \{1, 2, 3, 4\}$. The “strictly less than” relation $L$ on $X$ is the subset $L \subseteq X \times X$ given by $L = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$.

(ii) Let $X = \{2, 3, 4, 5, 6\}$. The divisibility relation $D$ on $X$ is the subset $D \subseteq X \times X$ given by $D = \{(2, 2), (2, 4), (2, 6), (3, 3), (3, 6), (4, 4), (5, 5), (6, 6)\}$.

(iii) Let $X = \{1, 2, 3, 4\}$. The equality relation $E$ on $X$ is the subset $E \subseteq X \times X$ given by $E = \{(1, 1), (2, 2), (3, 3), (4, 4)\}$.

(iv) Let $f : \{3, 4, 5\} \to \{0, 1\}$ be the function with $f(3) = f(5) = 0$ and $f(4) = 1$. This function is given by the relation $f \subseteq \{3, 4, 5\} \times \{0, 1\}$ defined as $f = \{(3, 0), (4, 1), (5, 0)\}$. This set is often called the **graph of $f$**, rather than $f$ itself (we will thoroughly examine functions in Section 3.3).

3.1 Orderings

A set $X$ can be ordered with either a partial order or a total order.

Definition 3.3. A partial order on $X$ is a binary relation $\leq$ on $X$ that is **reflexive**, **antisymmetric** and **transitive**.

- **reflexive**: $x \leq x$ for all $x \in X$.
- **anti-symmetric**: $x \leq y$ and $y \leq x$ implies $x = y$ for all $x, y \in X$.
- **transitive**: $x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in X$. 
A **total order** is a partial order where every pair $x, y \in X$ satisfies either $x \leq y$ or $y \leq x$.

**Example 3.4.**

(i) The sets $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are totally ordered by $\leq$, using the usual definition.

(ii) The set $\mathbb{N}$ is partially ordered by divisibility. For $a, b \in \mathbb{N}$, we write $a \mid b$ if there exists $c \in \mathbb{N}$ with $ca = b$ and say that $a$ divides $b$. This is a partial order on $\mathbb{N}$, but not a total order since $5 \nmid 7$ and $7 \nmid 5$.

(iii) The set $\mathbb{Z}$ is not partially ordered by divisibility since $-3 \mid 3$ and $3 \mid -3$ but $3 \neq -3$ (this breaks antisymmetry).

(iv) The powerset $\mathcal{P}(X)$ of a set $X$ (i.e., the set of all subsets of $X$), is ordered by containment; $Y \subseteq Z$.

### 3.2 Equivalence Relations

**Definition 3.5.** A relation $E$ on a set $X$ is an equivalence relation if it is reflexive, symmetric and transitive.

- **reflexive**: $xEx$ for all $x \in X$.
- **symmetric**: $xEy$ implies $yEx$ for all $x, y \in X$.
- **transitive**: $xEy$ and $yEz$ implies $xEz$ for all $x, y, z \in X$.

Whenever two elements $x, y \in X$ satisfy $xEy$ we say they are **equivalent**.

**Example 3.6.**

(i) Equality is an equivalence relation on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and any other set you can think of.

(ii) The relation $\leq$ is not an equivalence relation since it is not symmetric; $2 \leq 3$, but $3 \not\leq 2$.

(iii) Let $A = \{\text{all students in Mthe217}\}$. Let $a \sim b$ if $a$ and $b$ have the same age. Then $\sim$ is an equivalence relation on $A$.

(iv) Congruence modulo $n$: Fix $n \in \mathbb{Z}$. We say $a, b \in \mathbb{Z}$ are **congruent modulo** $n$ and write

$$a \equiv b \mod n$$
if $n$ divides $a - b$ (i.e., $a - b = qn$ for some $q \in \mathbb{Z}$). We show in Proposition 3.7 that congruence modulo $n$ is an equivalence relation on $\mathbb{Z}$; this equivalence relation will be later studied more closely.

**Proposition 3.7.** Congruence is an equivalence relation on $\mathbb{Z}$.

**Proof.** We first show that congruence is reflexive. If we take $a \in \mathbb{Z}$, then $a - a = 0 = 0 \cdot n$. Therefore $a \equiv a \mod n$.

Now we show that congruence is symmetric. If we have any $a, b \in \mathbb{Z}$ with $a \equiv b \mod n$ then $a - b = qn$ for some $q \in \mathbb{Z}$. Thus, $b - a = (-q)n$ and hence $b \equiv a \mod n$. This shows the symmetry of congruence.

Finally, for transitivity, take $a, b, c \in \mathbb{Z}$ with $a \equiv b \mod n$ and $b \equiv c \mod n$. We want to show that $a \equiv c \mod n$. First, $a - b = qn$ and $b - c = rn$ for some $q, r \in \mathbb{Z}$. Adding these two expressions gives $a - c = qn + rn = (q + r)n$. Thus, $a \equiv c \mod n$. 

**Definition 3.8.** Given an equivalence relation $\sim$ on $X$, the **equivalence class** of $a \in X$ is the set $[a] = \{b \in X \mid b \sim a\}$.

If our equivalence relation is congruence modulo $n$ on $\mathbb{Z}$, then equivalence classes of integers are called **congruence classes**.

**Proposition 3.9.** Suppose $\sim$ is an equivalence relation on $X$. For $a, b \in X$, $a \sim b$ if and only if $[a] = [b]$.

**Proof.** In order to show the forward direction, we assume that $a \sim b$; we aim to show $[a] = [b]$ by double inclusion. Take an arbitrary element $c \in [b]$. Since $c \in [b]$ we have $b \sim c$. As $a \sim b$, we know $a \sim c$ by transitivity. Therefore $c \in [a]$. We have shown $[b] \subseteq [a]$. The reverse containment $[a] \subseteq [b]$ holds by symmetry. Thus, $[a] = [b]$.

For the opposite direction, assume that $[a] = [b]$. Since $b \sim b$, we have $b \in [b] = [a]$ and hence $a \sim b$. Therefore $[a] = [b]$ implies $a \sim b$. 

**Example 3.10.** There are 5 different congruence classes of integers modulo 5.

$[0] = \{\ldots, -15, -10, -5, 0, 5, 10, 15, \ldots\}$
$[1] = \{\ldots, -14, -9, -4, 1, 6, 11, 16, \ldots\}$
$[2] = \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\}$
$[3] = \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\}$
$[4] = \{\ldots, -11, -6, -1, 4, 9, 14, 19, \ldots\}$
Note that $[5] = [0]$ since $5 \equiv 0 \mod 5$, $[6] = [1]$ since $6 \equiv 1 \mod 5$ and so on. Note also that the above congruence classes are pairwise disjoint and their union yields the entire set $\mathbb{Z}$.

**Definition 3.11.** Suppose $\sim$ is an equivalence relation on $X$. The set of all equivalence classes of elements in $X$ is called the **quotient set** and is denoted by

$$S/\sim = \{[a] | a \in S\}.$$  

If $X = \mathbb{Z}$ and our equivalence relation $\sim$ is congruence modulo $n$, then we write $\mathbb{Z}_n = \mathbb{Z}/\sim$.

From the previous example $|\mathbb{Z}_5| = 5$ and, more generally, $|\mathbb{Z}_n| = n$.

**Proposition 3.12.** Let $\sim$ be an equivalence relation on $X$. The sets in $X/\sim$ are pairwise disjoint and their union is $X$.

**Proof.** In order to show that distinct equivalence classes are disjoint (i.e., $[a] \neq [b] \Rightarrow [a] \cap [b] = \emptyset$), we prove the contrapositive (which is logically equivalent). In other words, we need to show that if $[a], [b] \in X/\sim$ have $[a] \cap [b] \neq \emptyset$ then $[a] = [b]$.

If we assume $[a] \cap [b] \neq \emptyset$, then there exists $c \in [a] \cap [b]$. As $c \in [a]$ and $c \in [b]$, we have $a \sim c$ and $b \sim c$. By transitivity $a \sim b$ and therefore $[a] = [b]$ by our previous proposition. This shows that the equivalence classes in $X/\sim$ are pairwise disjoint.

As every $[a] \in X/\sim$ is a subset of $X$, the union of all these sets is still a subset of $X$; $\bigcup_{[a] \in X/\sim} [a] \subseteq X$. If we take $b \in X$ arbitrarily, then $b \in [b] \in X/\sim$. Therefore $b \in \bigcup_{[a] \in X/\sim} [a]$. This shows that $X \subseteq \bigcup_{[a] \in X/\sim} [a]$. Therefore these sets are equal. \(\Box\)

**Definition 3.13.** Let $X$ be a set. A set $Y \subseteq \mathcal{P}(X)$ of subsets of $X$ is a **partition** of $X$ if the sets in $Y$ are pairwise disjoint and their union is $X$.

From the previous proposition, $X/\sim$ is a partition of $X$ for any equivalence relation $\sim$. The reverse also holds:

**Proposition 3.14.** Let $Y$ be a partition of $X$. Define a relation on $X$ by $a \sim b$ if there exists $Z \in Y$ with $a \in Z$ and $b \in Z$. The relation $\sim$ is an equivalence relation.
3.3 Functions

Definition 3.15. A function \( f : X \to Y \) is a relation \( \text{Gr}(f) \subseteq X \times Y \) which satisfies the following condition: for all \( x \in X \) there exists a unique (exactly one) \( y \in Y \) with \((x, y) \in \text{Gr}(f)\).

For \( x \in X \), the unique element \( y \in Y \) such that \((x, y) \in \text{Gr}(f)\) is denoted \( y = f(x) \) and called the image of \( x \) under \( f \). For function \( f : X \to Y \), the set \( X \) is called the domain of \( f \) while \( Y \) is called the range or codomain of \( f \). \( \text{Gr}(f) \) is also called the graph of \( f \).

Remark 3.16. A less formal definition of a function \( f : X \to Y \) is that it is a “rule” that assigns to every element \( x \in X \) exactly one element \( y \in Y \) called the image of \( x \) under \( f \) and denoted by \( y = f(x) \).

Definition 3.17. Let \( f : X \to Y \) be a function and let \( A \subseteq X \) and \( B \subseteq Y \) be sets. The image of \( A \) under \( f \) is the set
\[
  f(A) = \{ f(a) \mid a \in A \} = \{ y \in Y \mid \exists a \in A, f(a) = y \}.
\]
The image of the whole domain is simply called the image of \( f \): \( \text{Im } f = f(X) \). The pre-image of \( B \) is the set \( f^{-1}(B) = \{ x \in X \mid f(x) \in B \} \). The pre-image of an element \( y \in Y \) is the set \( f^{-1}(y) = \{ x \in X \mid f(x) = y \} \).

Observe that \( f(A) \subseteq Y \) while \( f^{-1}(B) \subseteq X \).

Example 3.18. Let \( f : \mathbb{R} \to \mathbb{R} \) be given by \( f(x) = x^2 \). Using interval notation,
- \( f([2, 3]) = [4, 9] \),
- \( f((-3, 2)) = [0, 9] \),
- \( f^{-1}([4, 9]) = [-3, -2] \cup [2, 3] \), and
- \( f^{-1}(2) = \{ \sqrt{2}, -\sqrt{2} \} \).

Exercise 3.19. Let \( f : X \to Y \). Show that the relation on \( X \) defined by \( a \sim b \) if \( f(a) = f(b) \) is an equivalence relation. Such equivalence relation is called the kernel equivalence of the function \( f \).

Definition 3.20. Let \( \sim \) be an equivalence relation on set \( X \), and let \( f : X \to Y \) be a function. The function \( \overline{f} : X/\sim \to Y \) given by \( \overline{f}([x]) = f(x) \) is well-defined on the quotient set \( X/\sim \) if \( f \) is constant on the equivalence classes of \( X \) (i.e., for all \( a, b \in X \) with \( a \sim b \) we have \( f(a) = f(b) \)).
Note that if \( f : X/\sim \to Y \) is not well-defined on \( X/\sim \), then \( f \) is not a function.

**Example 3.21.** Let \( f : \mathbb{Z} \to \{0, 1\} \) be given by

\[
f(a) = \begin{cases} 
0 & \text{a is even,} \\
1 & \text{a is odd.}
\end{cases}
\]

The function \( \overline{f} : \mathbb{Z}_5 \to \{0, 1\} \) given by \( \overline{f}([a]) = f(a) \) is not well-defined. For example, \( 2 \equiv 7 \mod 5 \) and 2 is even while 7 is odd. Therefore \( f(a) \) is not constant on \( a \in [2] = \{\ldots, -8, -3, 2, 7, \ldots, \} \) since it outputs both 0 and 1. In other words, since \( [2] = [7] \), we can’t define a function with \( 0 = f([2]) = f([7]) = 1 \).

If instead we define \( \overline{f} : \mathbb{Z}_4 \to \{0, 1\} \) by \( \overline{f}([a]) = f(a) \), then \( \overline{f} \) is well-defined: if \( a \equiv b \mod 4 \) then \( a - b = 4q \) for some \( q \in \mathbb{Z} \), so either \( a \) and \( b \) are both odd, or both even since they differ by a multiple of 4.

Common practice is to skip the definition of \( f \). For example, \( g : \mathbb{Z}_4 \to \{0, 1\} \) given by

\[
g([a]) = \begin{cases} 
0 & \text{a is even,} \\
1 & \text{a is odd}
\end{cases}
\]

is a well-defined function and equal to \( \overline{f} \) above.

**Definition 3.22.** A function \( f : X \to Y \) is injective (one-to-one) if for every \( a, b \in X \) with \( a \neq b \) we have \( f(b) \neq f(a) \). A function \( f : X \to Y \) is surjective (onto) if for every \( c \in Y \), there exists some \( a \in X \) with \( f(a) = c \). A function which is both injective and surjective is called *bijective*.

Note that we could have used the contrapositive to define injectivity: \( f \) is injective if for all \( a, b \in X \), \( f(a) = f(b) \) implies \( a = b \). We also could have said that \( f : X \to Y \) is surjective if \( \text{Im}(f) = Y \).

If there is a bijection \( f : X \to Y \), some authors will say that \( f \) gives a one-to-one correspondence between the elements of \( X \) and the elements of \( Y \). (We do not say there is a correspondence if \( f \) is only injective.)

**Example 3.23.**

(i) Let \( f : \mathbb{N} \to \mathbb{N} \) where \( f(n) = 2n + 1 \). This function is injective since \( f(m) = f(n) \Rightarrow 2m + 1 = 2n + 1 \Rightarrow 2m = 2n \Rightarrow m = n \). This function is not surjective since \( 4 \notin \text{Im} f \).
(ii) Let \( g : \mathbb{R} \to \mathbb{R} \) where \( g(x) = 2x + 1 \). This function is both injective (same proof as above) and surjective. (Proof: Given \( y \in \mathbb{R} \), \( g\left(\frac{y-1}{2}\right) = 2\frac{y-1}{2} + 1 = y \).) Thus, \( g \) is a bijection.

(iii) The function \( h : \mathbb{R} \to \mathbb{R} \) given by \( h(x) = x^2 \) is neither injective nor surjective. However, the function \( k : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) with \( k(x) = x^2 \) is bijective, where \( \mathbb{R}_{\geq 0} \) denotes the set of the non-negative reals.

(iv) The identity function \( \text{id}_X : X \to X \) defined by \( \text{id}_X(x) = x \) is a bijection.

**Definition 3.24.** The composition \( g \circ f : X \to Z \) of two functions \( f : X \to Y \) and \( g : Y \to Z \) is the function defined by \( (g \circ f)(x) = g(f(x)) \).

An endomorphism is a function \( f : X \to X \). Endomorphisms can be composed with themselves repeatedly: \( f^n = f \circ \cdots \circ f \), \( n \) times.

**Example 3.25.** Let \( f, g : \mathbb{R} \to \mathbb{R} \) be given by \( f(x) = x + 1 \) and \( g(x) = x^2 \). Then \( f \circ g(x) = f(x^2) = x^2 + 1 \) and \( g \circ f(x) = g(x + 1) = (x + 1)^2 \). So, \( f \circ g \neq g \circ f \).

**Lemma 3.26.** Let \( f : A \to B \), \( g : B \to C \) and \( h : C \to D \) be functions. Then:

(i) \( f \circ \text{id}_A = f = \text{id}_B \circ f \).

(ii) \( h \circ (g \circ f) = (h \circ g) \circ f \) (composition is associative).

(iii) If \( f \) and \( g \) are injective then \( g \circ f \) is injective.

(iv) If \( f \) and \( g \) are surjective then \( g \circ f \) is surjective.

(v) If \( f \) and \( g \) are bijective then \( g \circ f \) is bijective.

*Proof of iii.* Suppose \( f \) and \( g \) are injective and \( x, y \in A \) are distinct elements (i.e., \( x \neq y \)). As \( f \) is injective, \( f(x) \neq f(y) \). Since \( g \) is injective, \( g(f(x)) \neq g(f(y)) \). Therefore, \( (g \circ f)(x) \neq (g \circ f)(y) \), proving that \( g \circ f \) is injective.

**Definition 3.27.** Suppose \( f : X \to Y \) and \( g : Y \to X \) are functions. The function \( g \) is called a **compositional inverse** (or **inverse**) of \( f \) if both \( f \circ g = \text{id}_Y \) and \( g \circ f = \text{id}_X \).

Clearly, if \( g \) is an inverse of \( f \) then \( f \) is an inverse of \( g \).
**Lemma 3.28.** If there is a compositional inverse of \( f : X \to Y \) then that compositional inverse is unique.

**Proof.** Assume that both \( g : Y \to X \) and \( h : Y \to X \) are compositional inverses of \( f \). We have
\[
g = g \circ \text{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = \text{id}_X \circ h = h,
\]
proving \( g = h \). Since any two compositional inverses of \( f \) must be equal, if \( f \) has a compositional inverse, then it is unique. \( \square \)

The above proof shows something more specific: if \( f \) has a left inverse \( g \) and a right inverse \( h \) (i.e., \( g \circ f = \text{id}_X \) and \( f \circ h = \text{id}_Y \)) then these single-sided inverses are equal and \( f \) has a compositional inverse.

**Definition 3.29.** A function \( f : X \to Y \) is **invertible** if it has a compositional inverse. The unique compositional inverse of \( f \) is denoted \( f^{-1} : Y \to X \).

**Example 3.30.**

(i) The function \( f : \{1, 2, 3, 4, 5\} \to \{1, 2, 3\} \) and its compositional inverse are given below.

\[
\begin{align*}
  f(1) &= 3 & f^{-1}(1) &= 2 \\
  f(2) &= 1 & f^{-1}(2) &= 3 \\
  f(3) &= 2 & f^{-1}(3) &= 1 \\
  f(4) &= 5 & f^{-1}(4) &= 5 \\
  f(5) &= 4 & f^{-1}(5) &= 4
\end{align*}
\]

(ii) \( \exp : \mathbb{R} \to \mathbb{R}_{>0} \) has compositional inverse \( \ln : \mathbb{R}_{>0} \to \mathbb{R} \).

(iii) \( \sin : [-\pi/2, \pi/2] \to [-1, 1] \) has compositional inverse \( \arcsin : [-1, 1] \to [-\pi/2, \pi/2] \).

(iv) \( \text{id}_X : X \to X \) is its own compositional inverse.
**Theorem 3.31.** A function is invertible if and only if it is a bijection.

*Proof.* Assume \( f : X \to Y \) is invertible. Take \( a, b \in X \) with \( a \neq b \). As \( f^{-1}(f(a)) = a \neq b = f^{-1}(f(b)) \) we see that \( f(a) \neq f(b) \). Thus, \( f \) is injective.

If we pick \( c \in Y \) then \( f(f^{-1}(c)) = c \) and therefore \( c \in \text{Im} \ f \). Thus \( f \) is surjective and hence bijective.

For the other direction, assume that \( f : X \to Y \) is a bijection. Given \( c \in Y \), we know there exists \( a \in X \) with \( f(a) = c \) as \( f \) is surjective. If \( b \in X \) has \( f(b) = c = f(a) \) then \( a = b \) as \( f \) is injective. Thus, there is a unique element \( a \in X \) which maps to \( c \). So, define a map \( g : Y \to X \) by \( g(c) = a \) where \( a \) is the unique element with \( f(a) = c \). Since \( f(g(c)) = f(a) = c \) and \( g(f(a)) = g(c) = a \), we see that \( f \) is invertible with inverse \( g \). □

**Corollary 3.32.** The composition of two invertible functions is invertible.

*Proof.* This can be proven directly, or one can use the fact that a composition of two bijections is again a bijection. □

**Definition 3.33.** Two sets \( X \) and \( Y \) have the same **cardinality** if there exists a bijection between \( X \) and \( Y \). If \( X \) and \( Y \) have the same cardinality, we write \( |X| = |Y| \).

**Example 3.34.**

(i) There is a bijection between \( \mathbb{N} \) and \( \mathbb{Z} \) and there is a bijection between \( \mathbb{Z} \) and \( \mathbb{Q} \). Therefore,

\[
|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|.
\]

Any set that has the same cardinality as \( \mathbb{N} \) is called **countable**.

(ii) **Cantor’s diagonal argument** shows that there is no bijection between \( \mathbb{N} \) and \( \mathbb{R} \). Thus, \( \mathbb{R} \) is not countable. Since the inclusion map \( \iota : \mathbb{N} \to \mathbb{R} \), \( \iota(n) = n \), is an injection, we write \( |\mathbb{N}| < |\mathbb{R}| \). There can be no proof that there exists a set \( X \) with \( |\mathbb{N}| < |X| < |\mathbb{R}| \), nor can there be a proof that no such set exists. Therefore, one cannot use mathematics to decide whether there are infinite sets smaller than \( \mathbb{R} \) that are not countable. (See the wikipedia article on the **Continuum Hypothesis**.)