MATH 8 LECTURE NOTES

XIN ZHOU

ABSTRACT. This is the set of lecture notes for Math 8 during Spring quarter of 2018 at UC Santa Barbara. The lectures follow closely [1][2].

CONTENTS

1. Sets 2
2. Proofs 8
3. Quantifiers 13
4. Number systems 17
5. Decimals 20
6. Inequalities 24
7. Mathematical Induction 25
  7.1. Guessing the answer 27
  7.2. The Σ notation 27
  7.3. Principle of Strong Mathematical Induction 27
8. Integers 29
9. Prime factorization 32
10. Equivalence relations 37
11. Functions 41
12. Infinity 45
References 50

Date: June 7, 2018.
We will cover Chapter 1-3, 8, 10-13, 17-19, 21 with additional topics as time permits. This includes the following topics:

- Sets
- Number systems, decimals
- Proofs
  - Direct proof
  - Induction
  - Proof by contradiction.
- Inequalities
- Prime number
  - Prime factorization
- Equivalence relations
- Functions
- Infinity

Remark 0.1. For this class, reading the textbook and doing the homework is more important than previous math classes you may have taken. Also, you are expected to write homework proofs neatly and clearly, using complete sentences.

This course covers set theory, logic, functions and equivalence relations, and techniques of proof, including induction. We will also cover topics in number theory relating to integers and primes. The goals of the course are to introduce fundamental mathematical concepts and definitions and to be able to use them to write clear, logically correct proofs.

1. Sets

A set is just a collection of objects, we call those objects the elements of the set.

Remark 1.1. Actually, this simple definition for sets is inadequate and gives rise to contradictions. “The set of all sets” cannot exist. But we will not talk about this. If you are really interested, you may read any set theory textbook.

How to describe a set?

- the first way is just to make a list of all the objects in the set and put curly brackets (braces) around the list. For example,
  - \{1\} is a set consisting of the objects 1.
  - \{\pi, UCSB, math8\} is a set consisting of three objects.
  - \{1, 2, 3\} is a set consisting of the objects 1, 2, and 3.
– Here is a tricky one. \( \{1, \{2\}\} \) is the set consisting of two objects, one is just the number 1, and the other one is the set \( \{2\} \). We can make a really complicated set,
\[
\{1, \{2\}, \{3, \{4, 5\}\}\}.
\]

• But the first way is not that convenient to describe a set in some cases. For instance, if we have infinite number of objects, like all positive integers, or if there are no explicit expressions for the objects, like the solution of the equation \( x^\pi + x^2 - 8 = 0 \). So we use the following form
\[
\{x | P(x)\},
\]
where \( P(x) \) is a condition or property of \( x \). This is to be read “the set of all \( x \) such that \( x \) satisfies the condition \( P(x) \)”. Sometimes, we also use \( \{x \in X | P(x)\} \) to denote a set, where \( X \) is the domain of \( x \). For example,
\[
\begin{align*}
\{x & | x \text{ is a positive integer}\}, \\
\{x & | x \text{ is a real number, } x^2 < 2\}, \\
\{x & | x \text{ is a real number and } x^\pi + x^2 - 8 = 0\}.
\end{align*}
\]

Let me introduce some notations.

• \( \mathbb{N} = \{x | x \text{ is a natural number}\} = \{1, 2, 3, \ldots\} \) (ellipsis indicates that the list continues in the obvious way). The set of all natural numbers
• \( \mathbb{Z} = \{x | x \text{ is an integer}\} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\} \)
• Note that in many other textbooks, 0 is also a natural number.
• \( \mathbb{R} = \{x | x \text{ is a real number}\} \). The set of all real numbers
• \( \mathbb{Q} = \{x | x = \frac{p}{q}, p, q \in \mathbb{Z}, \text{ and } q \neq 0\} \). The set of all rational numbers.

**Definition 1.2.** We also define the empty set to be the set consisting of no objects, and denote by the symbol \( \emptyset \).

**Example 1.3.** Consider the following sets
\[
\{x | x \in \mathbb{R}, x^2 + 1 = 0\} = \emptyset,
\]
\[
\{n | n \in \mathbb{N}, n \text{ is odd and } n = k(k + 1) \text{ for some } k \in \mathbb{N}\} = \emptyset,
\]

We often use capital letters to denote sets.

**Definition 1.4.** Let \( S \) be a set. If \( x \) is an element of \( S \), then we write
\[
x \in S,
\]
and say $x$ belongs to $S$. If some other object $y$ does not belong to $S$, we write $y \notin S$.

For example, consider the set

$$\{1, \{2\}, \{3, \{4,5\}\}\}.$$ 

We know that

$$1 \in S, \{2\} \in S, \{3, \{4,5\}\} \in S,$$

but $2 \notin S, \{4,5\} \notin S$.

**Definition 1.5.** We say that two sets are equal when they consist of exactly the same elements.

For instance $A = \{1, 2, 3\} = \{2, 1, 3\}$. But here is another set $B = \{1, 1, 2, 3\}$. Since this set contains the same elements with the set $A$, it is equal to the set $A$. Indeed, the definition of a set requires that it contains distinct objects, so you cannot express a set like that. Consider the following set

$$C = \{x|x \in \mathbb{R}, x^2 - 2x + 1 = 0\} = \{1\}.$$ 

**Definition 1.6.** Let $S$ be a set. We say that a set $T$ is a subset of $S$ if every element of $T$ also belongs to $S$. We write

$$T \subseteq S.$$

if $T$ is a subset of $S$, and

$$T \not\subseteq S$$

if not.

Be careful with the notations $\subseteq$ and $\in$.

Note that, empty set $\emptyset$ is a subset of every set, and any set is a subset of itself.

**Example 1.7.** Let $S = \{1, 2, 3\}$. Then

$$\{1\} \subseteq S, \{1, 2\} \subseteq S, \{1, 3\} \subseteq S, \{2, 3\} \subseteq S, \{1, 2, 3\} \subseteq S, \emptyset \subseteq S.$$ 

**Example 1.8.** Let $S = \{1, 2\}, UCSB$. Then

$$\{1\} \subseteq S, \{UCSB\} \subseteq S, \{2\} \not\subseteq S, \{\{2\}\} \subseteq S.$$
Example 1.9. Let $S = \{1, \{2\}, \{3, \{4, 5\}\}\}$. Then
\[
\{1\} \subseteq S, \emptyset \subseteq S,
\]
and
\[
\{2\} \not\subseteq S
\]
since 2 is not an element of $S$. But
\[
\{\{2\}\} \subseteq S,
\]
Similarly,
\[
\{3\} \not\subseteq S, \{4, 5\} \not\subseteq S, \{\{3, \{4, 5\}\}\} \subseteq S.
\]
How about $\{1, \{2\}\}$? Yes.

For numbers, we have operations like addition and multiplication. For sets, we also have set operations. Next, we will introduce some set operations.

Definition 1.10. Let $A$ and $B$ be two sets. The union of $A$ and $B$, written $A \cup B$, is the set consisting of all elements that lie in either $A$ or $B$ (or both). That is,
\[
A \cup B = \{x | x \in A \text{ or } x \in B\}.
\]
The intersection of $A$ and $B$, written $A \cap B$, is the set consisting of all elements that lie in both $A$ and $B$. That is,
\[
A \cap B = \{x | x \in A \text{ and } x \in B\}.
\]

Definition 1.11. Let $A$ and $B$ be two sets. Their difference is defined to be the set
\[
A - B = A \setminus B = \{x | x \in A \text{ and } x \notin B\},
\]
\[
B - A = B \setminus A = \{x | x \in B \text{ and } x \notin A\},
\]
where $\setminus$ is backslash.

We can also use Venn diagram to explain those definitions.

Lemma 1.12. Let $A$ and $B$ be two sets. Then by Venn diagram, it is easy to see that
\[
A \cup B = (A - B) \cup (A \cap B) \cup (B - A),
\]
and $A - B, B - A, A \cap B$ are mutually disjoint.
Example 1.13. Let \( A = \{1, 2, 3, 4\} \) and \( B = \{3, 4, 5\} \). We have
\[
A \cup B = \{1, 2, 3, 4, 5\}, \quad A \cap B = \{3, 4\},
\]
\[
A - B = \{1, 2\}, \quad B - A = \{5\}.
\]

Example 1.14. Let \( A = [0, 2] \) and \( B = [1, 3] \). Then
\[
A \cup B = [0, 3], \quad A \cap B = [1, 2], \quad A - B = [0, 1], \quad B - A = (2, 3].
\]

Remark 1.15. What about set operations involving more than two sets? We know that for numbers, there is a default order of operations. For example
\[
2 + 3 \times 5,
\]
we know that we should do the multiplication first.

But for set operations, union, intersection, and difference operations are all equal in the order.

For example, the expression
\[
A \cup B \cap C
\]
does not make any sense because we do not know which operation we should do first: should we take the union first, and then the difference, or should we take the difference first and then the union?

So if we have more than one of these at a time, we have to use parentheses to indicate which of these operations should be done first.

In order to make this clear, we need to either write
\[
(A \cup B) \cap C \quad \text{or} \quad A \cup (B \cap C).
\]

Example 1.16. Let \( C = \{1, 3, 5, 7\} \). Then \( B \cap C = \{3, 5\} \),
\[
(A \cup B) \cap C = \{1, 3, 5\}, \quad A \cup (B \cap C) = \{1, 2, 3, 4, 5\}.
\]

Lemma 1.17. Let \( A \) and \( B \) be two sets. If \( A \subseteq B \), then
\[
A \cup B = B, \quad A \cap B = A, \quad A - B = \emptyset.
\]

Definition 1.18. We say that \( A \) and \( B \) are disjoint sets if they have no elements in common, i.e., if \( A \cap B = \emptyset \).

Question 1.19. Let \( A \) and \( B \) be two sets. If \( A \) and \( B \) are disjoint, i.e., \( A \cap B = \emptyset \), then
\[
A - B = A, \quad B - A = B.
\]
We can extend the definitions of union and intersection to many sets. If \( A_1, A_2, \ldots, A_n \) are sets, their union and intersection are defined as

\[
A_1 \cup A_2 \cup \ldots \cup A_n = \{x | x \in A_i \text{ for some } i, 1 \leq i \leq n\},
\]

\[
A_1 \cap A_2 \cap \ldots \cap A_n = \{x | x \in A_i \text{ for all } i, 1 \leq i \leq n\},
\]

We also use the following more concise notation

\[
A_1 \cup A_2 \cup \ldots \cup A_n = \bigcup_{i=1}^{n} A_i,
\]

\[
A_1 \cap A_2 \cap \ldots \cap A_n = \bigcap_{i=1}^{n} A_i.
\]
2. PROOFS

Let us talk about mathematical proofs. One of the goals of this course is to learn how to write proofs.

Why we need proofs? There are many math questions and mathematicians are trying to find answers to those questions. They can do experiments; they can use computers to simulate; they can also guess and try. Sometimes they may use those methods to get the answer, but for mathematicians, they are not convinced unless they can prove it. For many other subjects, in order to explain something, people propose a theory and this theory can explain 99% of the cases, but fails the 1%. Later, a new theory will come up and replace the old one.

we only have two situations: we can prove it or we can’t. Once it is proved, then it will be always true and will not be replaced.

**Example 2.1.** There is a very famous conjecture about prime numbers named Goldbach’s Conjecture: every even \( n > 2 \) is the sum of two primes. For example,

\[
4 = 2 + 2, \quad 6 = 3 + 3, \quad 8 = 3 + 5, \quad 10 = 5 + 5, \quad 12 = 5 + 7, \ldots
\]

People verified that this is true up to \( 4 \times 10^{17} \). Numerical results suggest that this is almost true and mathematicians also believe that it is true. But since there is no proof, we cannot use it as a theorem to prove other results.

Before we talk about proofs, we need to introduce some notations. Let \( P \) and \( Q \) are statements or mathematical statements. For example, (we always use \( x \) to denote some real number, and \( n \) to denote some natural number):

\[
P_1: \ x = 2; \\
P_2: \ 2 < 3; \\
P_3: \ x > 3 \text{ and } x < 2; \\
P_4: \ n \text{ is an even integer}; \\
P_5: \ n = m^2 \text{ for some integer } m.
\]

For \( P_1, P_4, P_5 \), we do not know whether it is true or false before we know the values of \( x \) or \( n \). \( P_2 \) is true, but \( P_3 \) is false.

**Definition 2.2.** A mathematical statement is a mathematical sentence that is either true or false (but not both).

**Definition 2.3.** We write

\[
P \Rightarrow Q
\]
to mean that the statement $P$ implies statement $Q$ (It is called implication).

For example,

$$x = 2 \Rightarrow x^2 < 6.$$ 

**Remark 2.4.** In natural language (and intuitively in mathematics), the statement

$$P \Rightarrow Q$$

suggests a relationship between the statement $P$ and $Q$; namely that the truth of $P$ somehow forces the truth of $Q$. However, as a propositional connective, this relationship between $P$ and $Q$ is not required for logical implication.

The only case that $P \Rightarrow Q$ is false is that $P$ is true and $Q$ is false.

Other ways of saying $P \Rightarrow Q$ are

- if $P$, then $Q$; If $P$ is true, then $Q$ is true.
- $Q$ if $P$;
- $P$ only if $Q$.

**Definition 2.5.** The **negation** of a statement $P$ is the opposite statement, “not $P$”, written as $\overline{P}$ (some books use $\neg P$).

For instance, let us take the negation of all the examples on the previous page:

- $\overline{P_1}$: $x \neq 2$;
- $\overline{P_2}$: $2 > 3$;
- $\overline{P_3}$: $x \leq 3$ or $x \geq 2$;
- $\overline{P_4}$: $n$ is not an even integer;
- $\overline{P_5}$: $n \neq m^2$ for all integer $m$; $n$ cannot be written as $n = m^2$ for some integer $m$.

Note that $P \Rightarrow Q$ does not mean that $Q \Rightarrow P$. This is easy to understand. For example, $x > 1 \Rightarrow x > 0$, but $x > 0 \nRightarrow x > 1$.

**Definition 2.6.** If for some statements $P$ and $Q$, if both $P \Rightarrow Q$ and $Q \Rightarrow P$ are true, then we write

$$P \Leftrightarrow Q,$$

and say that “$P$ if and only if $Q$” or simply $P$ and $Q$ are equivalent.

For example,

$$x = 1 \Leftrightarrow x^3 = 1.$$
Lemma 2.7. An implication and its contrapositive are propositionally equivalent. That is, if \( P \Rightarrow Q \) is true, then we also have \( \bar{Q} \Rightarrow \bar{P} \). This is because if \( Q \) is true, then \( P \) cannot be true, as \( P \Rightarrow Q \).

Example 2.8. Let \( P \) be the statement \( x = 1 \), and \( Q \) the statement \( x^2 < 6 \). We have

- \( \bar{P} : x \neq 1 \); \( \bar{Q} : x^2 \geq 6 \).
- \( P \Rightarrow Q \) is true, but \( Q \Rightarrow P \) is not true.
- \( \bar{Q} \Rightarrow \bar{P} \) is true, and \( \bar{P} \Rightarrow \bar{Q} \) is false.

Example 2.9. Suppose we are given the following facts:

I get A for math 8 only if I finish every homework.

What can be deduced in the following?

a. I get A. (I finish every homework.)

b. I finish every homework. (Nothing)

c. I get B. (Nothing)

d. I do not finish every homework. (I will not get A.)

Definition 2.10. An axiom is a statement in math that is assumed to be true, without requiring proof.

In general, a proof contains a series of implications, starting with any basic axioms or given assumptions, until the desired conclusion is reached. Suppose we want to prove \( Q \) is true, and our assumption is \( P \). The process is

\[
P \Rightarrow P_1 \Rightarrow P_2 \ldots \Rightarrow Q.
\]

Using the logic, a proof may have different forms. The most common method of proof is direct proofs.

Definition 2.11. Direct proofs: starting with any basic axioms or given assumption, use a series of implications to conclude that the desired result is true.

Example 2.12. The square of an odd integer is odd.

Proof. Let \( n \) be an arbitrary odd integer. Since \( n \) is odd, it is 1 more than an even integer; that is, \( n \) can be written as \( 2m + 1 \) for some integer \( m \).

Therefore, \( n^2 = (1 + 2m)^2 = 4m^2 + 4m + 1 = 4(m^2 + m) + 1 \). \( \square \)
Remark 2.13. We could have written as this proof as the following series of implications:

\[ n \text{ is odd} \rightarrow n = 2m + 1 \rightarrow n^2 = 1 + 4(m^2 + m) \rightarrow n^2 \text{ is odd}. \]

But, this is too terse and somewhat strange. We are writing proofs so that people can understand; we are human beings, not computers. We need to use some English words to make the proof smooth and readable. For example, we can use “then”, “therefore”, “hence” and so on. (Certainly, you guys know English better than I).

We also have indirect proofs, for example, proof by contradiction is a form of indirect proof. Sometimes, it is almost impossible to give a direct proof, for instance, if we want to prove the following statements:

- there exists infinitely many prime numbers. If we want to prove this directly, our only choice is to list infinitely many primes or give a general formula for infinitely many primes. But this is almost impossible or very hard. However, we can easily prove this using proof by contradiction.

Definition 2.14. Proof by contradiction: Suppose we want to prove a statement \( P \) is true. We first assume that \( P \) is false, that is, \( \bar{P} \) is true. We then start with \( \bar{P} \) to deduce a statement \( Q \) which is impossible or contradict our assumption. Hence, \( P \) must be true.

If \( \bar{P} \) is true, then \( Q \) is true. \( Q \) is false. Hence, \( P \) is true.

This uses \( \bar{P} \Rightarrow Q \) is equivalent to \( Q \Rightarrow P \). That is, an implication and its contrapositive are propositionally equivalent.

Example 2.15. There is no greatest even integer.

Proof. Suppose the conclusion is not true. (We take the negation of the theorem and suppose it to be true.) Suppose there is a greatest even integer \( N \). (We must deduce a contradiction.)

Then for every even integer \( n \), \( N \geq n \).

Now let \( M = N + 2 \). Then, \( M \) is an even integer. (Because it is a sum of even integers.) Also, \( M > N \) since \( M = N + 2 \). Therefore, \( M \) is an integer that is greater than the greatest integer.

This contradicts the assumption that \( N \geq n \) for every even integer \( n \). This completes the proof. \( \square \)

Example 2.16. Let \( n \) be an integer such that \( n^2 \) is a multiple of 3. Then \( n \) is also a multiple of 3.
Proof. Suppose \( n \) is not a multiple of 3. Then, if we divide \( n \) by 3, we get a remainder of either 1 or 2; that is, \( n \) is either 1 or 2 more than a multiple of 3.

If the remainder is 1, then \( n = 1 + 3k \) for some integer \( k \) and

\[
    n^2 = (1 + 3k)^2 = 9k^2 + 6k + 1 = 1 + 3(2k + 3k^2).
\]

But this means that \( n^2 \) is 1 more than a multiple of 3, which contradicts our assumption that \( n^2 \) is a multiple of 3.

Similarly, if \( n = 2 + 3k \) for some integer \( k \), then

\[
    n^2 = (3k + 2)^2 = 1 + 3(1 + 4k + 3k^2).
\]

Hence, we have shown that assuming \( n \) is not a multiple of 3 leads to a contradiction (false statement).

Sometimes, we need to prove a statement is false - in other words, disproving it. For some cases, we can use an example to prove that a statement is false; this example is called a counterexample and we call the method **disproof by counterexample**.

**Example 2.17.** Prove the following statements are false.

1. All men are Chinese.
2. Every positive integer is equal to the sum of two integer squares.

Proof. In order to prove a statement is false or disprove it, we only need to prove its negation is true.

For (a), its negation is “not all men are Chinese”, or equivalently, “there exists a man who is not Chinese”. How to prove this? We can just find a man who is not Chinese, and this man will be a counterexample to (a).

For (b), its negation is not every positive integer is equal to the sum of two integer squares, or “there exists a positive integer that is not equal to the sum of two squares. 3 will be the counterexample.
3. Quantifiers

Next, we introduce two symbols. In mathematics, we will often see two types of statements. They are so common and important that we will introduce some symbols to denote them.

Let us consider the following examples:

• There is an integer $x$ such that $x^3 = 27$.
• For some integer $x$, $x^2 = -1$.
• There exists a positive integer that is not equal to the sum of two integer squares.

You can see that all the statements have the form: there exists some integer with a certain property. This type of statements is everywhere in math. So we will introduce a symbol $\exists$ (the backward $E$) to denote “there exists”.

Definition 3.1. The existential quantifier is denoted by the symbol $\exists$, and is read “there exists”.

So we can rewrite the above statements as follows:

1. $\exists x \in \mathbb{Z}$ such that $x^3 = 27$.
2. $\exists x \in \mathbb{Z}$ such that $x^2 = -1$.
3. $\exists x \in \mathbb{Z}$ and $x > 0$ such that $x$ is not equal to the sum of two integer squares.

They all have the form:

$$\exists x \text{ such that } P(x),$$

or simply

$$\exists x (x \in X)\ P(x).$$

It means that there is at least one value of $x$ for which $P(x)$ is true.

To prove an existence statement is true, we only need to find just one object satisfying the required property. There might be many objects satisfying the property, but just one object is enough to conclude that the statement is true.

To prove an existence statement is false, we need to show that no such object satisfying the required property; that is, for all $x$, $P(x)$ is false or $\overline{P(x)}$ is true.

For (1), we know $x = 3$ satisfies the property. For (2), since there is no such $x$, the statement is false. For (3), $x = 3$ has the required property.

Let us consider another type of statements:

• For all integers $n$, $n^2 \geq 0$.
• The cube of any integer is positive.
• Every positive integer is equal to the sum of two integer squares.

Similarly, you can see that all the statements have the form: for all integer, a certain property is true. We will introduce a symbol $\forall$ (the upside down $A$) to denote “for all”.

**Definition 3.2.** The **universal quantifier** is denoted by the symbol $\forall$, and is read “for all”.

So we can rewrite the above statements as follows:

(4) $\forall n \in \mathbb{Z}, n^2 \geq 0$.
(5) $\forall x \in \mathbb{Z}, x^3 > 0$.
(6) $\forall x \in \mathbb{Z}$ and $x > 0$, $x$ is equal to the sum of two integer squares.

They all have the form:

$$\forall x (x \in X), P(x).$$

To show that a “for all” statement is true, we need to prove it is true for all objects; that is, we need to give a general argument. To show it is false, we only need to find a counterexample. (4) is obviously true. (5) is false since any negative integer is a counterexample. (6) is false since 3 is a counterexample.

Actually, many math statements are very complicated and they have more than one quantifiers. For instance, (6) can be rewritten as

(6) $\forall x \in \mathbb{Z}$ and $x > 0$, $\exists m, n \in \mathbb{Z}$ such that $x = m^2 + n^2$.

**Example 3.3.** Consider another example:

• for any integer $a$, there is an integer $b$ such that $a + b = 0$.
• $\forall a \in \mathbb{Z}, \exists b \in \mathbb{Z}$ such that $a + b = 0$.
• There exists an integer $b \in \mathbb{Z}$ such that for all integer $a \in \mathbb{Z}$, $a + b = 0$.
• $\exists b \in \mathbb{Z}$ such that $\forall a \in \mathbb{Z}$, $a + b = 0$.

These two statements look very similar, but they have completely different meanings - in fact, one is true and one is false. These examples tell you that the **order of quantifiers really matters**.

Next, we will study the negation of a statement involving quantifiers. First, let’s see how to negate the existence statements.

**Example 3.4.** Consider the following statements:

(1) $\exists x \in \mathbb{Z}$ such that $x^3 = 27$. 
The negation of this statement is “there does not exist an integer \( x \) such that \( x^3 = 27 \). In other words, every integer has cube not equal to 27 or for all integers \( x \), \( x^3 \neq 27 \).
\[
\forall x \in \mathbb{Z}, \; x^3 \neq 27.
\]
(2) \( \exists x \in \mathbb{Z} \) such that \( x^2 = -1 \).
\[
\forall x \in \mathbb{Z}, \; x^2 \neq -1.
\]
(3) \( \exists x \in \mathbb{Z} \) and \( x > 0 \) such that \( x \) is not equal to the sum of two integer squares.

There exists a positive integer that is not equal to the sum of two integer squares.

Its negation is there does not exist a positive integer that is not equal to the sum of two integer squares. That is, every positive integer is equal to the sum of two integer squares.
\[
\forall x \in \mathbb{Z} \text{ and } x > 0, \; \exists m, n \in \mathbb{Z} \text{ such that } x = m^2 + n^2.
\]

We can conclude that for the existence statement
\[
\exists x, \; P(x),
\]
its negation is
\[
\forall x, \; \overline{P(x)}.
\]

This means that the negation of an existence statement is just “for all” statement.

Second, let’s see how to negate the “for all” statements.

**Example 3.5.** Consider the following:

(4) \( \forall n \in \mathbb{Z}, \; n^2 \geq 0 \).
\[
\exists n \in \mathbb{Z} \text{ such that } n^2 < 0.
\]
(5) \( \forall x \in \mathbb{Z}, \; x^3 > 0 \).

We can conclude that for the “for all” statement
\[
\forall x, \; P(x),
\]
its negation is
\[
\exists x, \; \overline{P(x)}.
\]

**Remark 3.6.** To summarize: when forming the negation of a statement involving quantifiers, we change \( \exists \) to \( \forall \), change \( \forall \) to \( \exists \) and negate the conclusion.
Example 3.7. For the statements containing more than one quantifiers, we just do negations step by step as follows:

(7) For any integer $x$ and $y$, there is an integer $z$, such that $x^2 + y^2 = z^2$.

\[ \Leftrightarrow \forall x \in \mathbb{Z}, \forall y \in \mathbb{Z}, (\exists z \in \mathbb{Z}, \text{ such that } x^2 + y^2 = z^2); \]

Its negation is:

\[ \exists x \in \mathbb{Z}, \exists y \in \mathbb{Z}, (\exists z \in \mathbb{Z}, \text{ such that } x^2 + y^2 = z^2); \]

\[ \Leftrightarrow \exists x \in \mathbb{Z}, \exists y \in \mathbb{Z}, (\forall z \in \mathbb{Z}, x^2 + y^2 \neq z^2); \]

\[ \Leftrightarrow \exists x \in \mathbb{Z}, \exists y \in \mathbb{Z}, \text{ such that } \forall z \in \mathbb{Z}, x^2 + y^2 \neq z^2. \]

We finally record the following result for set operations.

Proposition 3.8. (Proposition 17.1) Let $A, B, C$ be sets. Then

\[ A \cap (B \cup C) = (A \cap B) \cup (A \cap C). \]

Proof.

\[ x \in A \cap (B \cup C) \Leftrightarrow x \in A \text{ and } x \in B \cup C \]

\[ \Leftrightarrow x \in A \text{ and } x \in (B \text{ or } C) \]

\[ \Leftrightarrow (x \in A \text{ and } x \in B) \text{ or } (x \in A \text{ and } x \in C) \]

\[ \Leftrightarrow x \in (A \cap B) \cup (A \cap C). \]

We can also prove $A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$ and

\[ (A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C). \]
4. Number systems

We will introduce three number systems: the real numbers, the integers and the rational numbers.

In the past, maybe 2000 or 3000 years ago, people were using integers and they were also familiar with the fractions. Fractions are just rational numbers. But it then turned out that the rational numbers are not enough to describe every length. A simple example is $\sqrt{2}$. Suppose we have a right triangle with sides $a, b, c$ Pythagorean theorem implies that $a^2 + b^2 = c^2$. So people discovered the irrational numbers.

What are the real numbers? They are just all the rational numbers and irrational numbers. We can also view the real numbers as follows: If we draw an infinite straight line and choose a point on this line as the origin. We also choose a unit of length and label the whole numbers. Real numbers can be thought of as points on an infinitely long number line.

Obviously, the real numbers have a natural ordering, that is, we can compare the values of any two real numbers. For instance, suppose $x$ and $y$ are two real numbers. If $x$ is to the left of $y$ on the real line, then $x < y$.

**Definition 4.1.** The integers are just whole numbers, denoted by $\mathbb{Z}$.

**Definition 4.2.** A real number of the form $\frac{m}{n}$ (where $m, n \in \mathbb{Z}, n \neq 0$) is called a rational number, denoted by $\mathbb{Q}$.

An irrational number is a real number that is not a rational number.

Note that different fractions can represent the same rational numbers. We say that the rational $\frac{m}{n}$ is in lowest terms if no canceling is possible - that is, $m$ and $n$ have no common factor except 1 or $m$ and $n$ are coprime to each other.

**Definition 4.3.** If two integers $a$ and $b$ are said to be relatively prime, mutually prime, or coprime if the only positive integer that divides both of them is 1. That is, the only common positive factor of the two numbers is 1. This is equivalent to their greatest common divisor being 1.

For real numbers, we can do addition and multiplication.

**Fact 4.4.** (Rule 2.1) For all $a, b, c \in \mathbb{R}$,

1. Commutative law for addition and multiplication: $a + b = b + a, \ a b = b a$.
2. Associative law for addition and multiplication: $a + (b + c) = (a + b) + c, \ a (bc) = (ab)c$. 
(3) Distributive law: \( a(b + c) = ab + ac \).

These rules are axioms for real numbers and they cannot be proved or deduced from other facts.

Next, we will prove some facts about real numbers.

The first fact is that rational numbers are relatively dense on the real line.

**Proposition 4.5.** (Proposition 2.1) Between any two distinct rationals there is another rational.

**Proof.** Let \( r \) and \( s \) be two rationals. We can assume \( r < s \), otherwise the proof processes the same way by flipping \( r \) and \( s \). Let \( t = \frac{1}{2}(r + s) \). We now prove that \( t \) is a rational, and \( r < t < s \) (this is the mathematical meaning of saying \( t \) is between \( r \) and \( s \)).

We first prove that \( t \) is a rational. Since \( r, s \) are rationals, \( r = \frac{m}{n}, s = \frac{p}{q} \) for some \( m, n, p, q \in \mathbb{Z} \) and \( n \neq 0, q \neq 0 \). Then

\[
\frac{1}{2}(r + s) = \frac{1}{2} \left( \frac{m}{n} + \frac{p}{q} \right) = \frac{mq + np}{2nq}.
\]

Since \( mq + np, 2nq \in \mathbb{Z} \) and \( 2nq \neq 0 \), we know \( t \) is a rational.

Next we prove \( r < t < s \). Since by assumption \( r < s \), we have \( \frac{r}{2} < \frac{s}{2} \). Therefore

\[
r = \frac{r}{2} + \frac{r}{2} < \frac{r}{2} + \frac{s}{2} = t;
\]

similarly one can prove \( t < s \).

So we finish the proof. \( \square \)

**Proposition 4.6.** (Proposition 2.3) \( \sqrt{2} \) is not rational.

**Proof.** Let us prove by contradiction. Assume by contradiction that \( \sqrt{2} \) is a rational. Then

\[
\sqrt{2} = \frac{m}{n}, \text{ for some } m, n \in \mathbb{Z} \text{ and } n \neq 0.
\]

We can assume that \( \frac{m}{n} \) is in its lowest terms.

Take the squares, we have:

\[
2 = \frac{m^2}{n^2} \implies m^2 = 2n^2.
\]

Therefore \( m^2 \) is an even integer. Next we prove the statement:

(P): if \( m \in \mathbb{Z} \) and \( m^2 \) is even, then \( m \) is even.
Proof. We prove by contradiction. If \( m \) is not even, then \( m = 2k + 1 \) for some \( k \in \mathbb{Z} \). Therefore

\[
m^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,
\]

is odd. This is a contradiction to the assumption that \( m^2 \) is even. So \( m \) is even, and we finish the proof.

Now go back to the proof. Since \( m^2 \) is an even integer, by the above statement we know that \( m \) is also an even integer. Write \( m = 2k \) for some \( k \in \mathbb{Z} \), then

\[
2n^2 = m^2 = (2k)^2 = 4k^2 \implies n^2 = 2k^2.
\]

Hence \( n^2 \) is an even integer. Using the above statement again, \( n \) is an even integer. Therefore 2 is a common factor of \( m \) and \( n \). This is a contradiction to the assumption that \( \frac{m}{n} \) is in its lowest terms. As a conclusion, \( \sqrt{2} \) is not rational.

**Proposition 4.7.** (Proposition 2.4) Let \( a \) be a rational number and \( b \) an irrational.

1. Then \( a + b \) is irrational.
2. If \( a \neq 0 \), then \( ab \) is also irrational.

Proof. We will prove this by contradiction.

**Question 4.8.** Consider the following.

1. a rational + a rational = rational.
2. a rational + an irrational = irrational.
3. an irrational + an irrational =?
4. a rational \times a rational = rational.
5. a rational \times an irrational =?
6. an irrational \times an irrational=?

**Proposition 4.9.** (Proposition 2.5) Between any two real numbers there is an irrational.

Proof. The idea is to use the fact that the sum of a rational number and an irrational number is irrational.

Let \( a \) and \( b \) be any two real numbers. Suppose \( a < b \). We can choose a large positive integer such that

\[
\frac{\sqrt{2}}{n} < b - a.
\]

If \( a \) is a rational number, then \( a + \frac{\sqrt{2}}{n} \) is an irrational between \( a \) and \( b \). If \( a \) is an irrational, then \( a + \frac{1}{n} \) is an irrational that lies between \( a \) and \( b \).
5. Decimals

We introduce the decimal notation. Decimal notation is the writing of numbers in a base 10 numeral system.

We know that every point on the real line represents a real number and we can write out some of them, for example, the whole numbers, the rational numbers, and maybe some irrational numbers, like $\sqrt{2}$. But how about other irrational numbers. We need to find a way to express all real numbers. So we use decimal notation which is the writing of numbers in a base 10 numeral system.

**Example 5.1.** Suppose a real number has decimal expression:

$$a_0.a_1a_2a_3,$$

where $a_0$ is an integer and $a_1, a_2, a_3$ are integers between 0 and 9. We know that this number equals to

$$a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3}.$$

For example,

$$100, 1.5, \frac{1}{2} = 0.5,$$

$$\frac{1}{3} \text{ (one third)} = 0.333\ldots, \pi = 3.14159265358979323846264338\ldots.$$

123.456: the 3 is in the **Ones** position, meaning 3 ones (which is 3); the 2 is in the **Tens** position meaning 2 tens (which is twenty); and the 1 is in the **Hundreds** position, meaning 3 hundreds; the 4 is in the **Tenths** position; the 5 is in the **Hundredths** position; the 6 is in the **Thousandths** position.

What if a decimal has an infinite (non-terminating) fractional part? What does that meaning? In order to make it a little bit more precise, we need the following lemma. Given

$$a_0.a_1a_2\ldots,$$

this equals to the sum of the series

$$a_0 + \sum_{k=1}^{\infty} \frac{a_k}{10^k}.$$

**Proposition 5.2.** Let $x$ be a real number.

1. If $x \neq 1$, then

$$x + x^2 + x^3 + \ldots + x^n = \frac{x(1 - x^n)}{1 - x}.$$
2. If \(-1 < x < 1\), then the sum to infinity 
\[ x + x^2 + x^3 + \ldots = \frac{x}{1-x}. \]

For example, if we let \(x = \frac{1}{10}\) (one tenth), then 
\[ \frac{1}{10} + \frac{1}{10^2} + \frac{1}{10^3} + \ldots = \frac{1}{9}. \]

Next, we will show that every real number has a decimal expression and the expression is almost unique. First, we prove the following:

**Proposition 5.3.** (Proposition 3.2) Every real number \(x\) has a decimal expression 
\[ x = a_0.a_1a_2a_3\ldots, \]
where \(a_0\) is an integer and \(a_1, a_2, a_3, \ldots\) are integers between 0 and 9.

**Proof.** \(\square\)

We now know that every real number has a decimal expression, a natural question is that is the decimal expression unique, that is: can the same real number have two different decimal expression? The answer is Yes. Let us see an example

\[ 0.99999\ldots = 1. \]

We will next show that this is the only case that a real number has two different decimal expressions. With this proposition, we can say that the decimal expression is almost unique.

**Proposition 5.4.** Suppose that \(a_0.a_1a_2a_3\ldots\) and \(b_0.b_1b_2b_3\ldots\) are two different decimal expressions for the same real number. Then one of these expressions ends in 9999\ldots and the other ends in 000\ldots.

**Proof.** We may assume that \(a_0 = b_0 = 0\) (Since we can divide by 10 or 100,\ldots). Let \(x\) be the real number that has two expressions, that is, 
\[ x = 0.a_1a_2a_3\ldots = 0.b_1b_2b_3. \]

Let the first place where the two expressions disagree be the \(k\)-th place (\(k\) could be 1 of course). Thus, 
\[ x = 0.a_1a_2\ldots a_{k-1}a_k\ldots = 0.a_1a_2\ldots a_{k-1}b_k\ldots, \]
where \(a_k \neq b_k\). We may assume that \(a_k > b_k\), hence \(a_k \geq b_k + 1\). Note that we have 
\[ x \geq 0.a_1a_2\ldots a_{k-1}a_k0000 \]
and 

\[ x \leq 0.a_1a_2 \ldots a_{k-1}b_k9999 = 0.a_1a_2 \ldots a_{k-1}(b_k + 1)000. \]

Hence, all the inequalities are indeed equalities. It follows that \( a_k = b_k + 1 \) and

\[ x = 0.a_1 \ldots a_k000 = 0.a_1 \ldots (a_k - 1)999 \ldots. \]

□

We know that real numbers include rational numbers and irrational numbers. So we want to know what are the decimal expressions for rationals and irrationals? Certainly, they should be different. So what is the difference?

**Example 5.5.** Let us first consider some rational numbers.

\[ \frac{1}{2} = 0.5, \quad \frac{1}{3} = 0.333 \ldots, \quad \frac{7}{6} = 1.16666 \ldots, \quad \frac{8}{7} = 1.142857142857 \ldots. \]

You can see that they are either finite or infinite, and if it is infinite, then there is a sequence of digits that eventually repeats forever. We call such a decimal expression **periodic**.

If a periodic decimal has the form (repeating decimal with \( b_1 \ldots b_l \) repeating)

\[ a_0.a_1a_2 \ldots a_k b_1b_2 \ldots b_l b_1b_2 \ldots b_l \ldots, \]

then we write it as

\[ a_0.a_1 \ldots a_k \overline{b_1 \ldots b_l}. \]

That a rational number must have a finite or recurring decimal expansion can be seen to be a consequence of the long division algorithm, in that there are at most \( q - 1 \) possible nonzero remainders on division by \( q \), so that the recurring pattern will have a period less than \( q \).

Rational numbers have finite or infinite repeating decimal expressions while irrational numbers have infinite non-repeating decimal representations.

**Proposition 5.6.** The decimal expression for any rational number is (finite or) periodic.

**Proof.** Consider a rational number \( \frac{m}{n} \) (\( m, n \in \mathbb{Z} \)). Think about how we get the decimal expression for a fraction.

\[ m = a_0n + b_0, \quad 0 \leq b_0 \leq n - 1. \]

\[ 10b_0 = a_1n + b_1, \ldots \]

If some \( b_i = 0 \), then we get a finite decimal. Suppose for all \( b_i \neq 0 \). Note that \( b_i \) is an integer between 0 and \( n - 1 \), so at most after \( n \) steps, \( b_i \) will repeat. □
Example 5.7. Consider $9/8$.

\[ 9 = 1 \times 8 + 1; \quad 10 \times 1 = 1 \times 8 + 2; \quad 10 \times 2 = 2 \times 8 + 4; \quad 10 \times 4 = 5 \times 8. \]

So we have $9/8 = 1.125$.

Consider $8/7$.

\[ 8 = 1 \times 7 + 1; \quad 10 \times 1 = 1 \times 7 + 3; \quad 10 \times 3 = 4 \times 7 + 2; \quad 10 \times 2 = 2 \times 7 + 6; \]
\[ 10 \times 6 = 8 \times 7 + 4; \quad 10 \times 4 = 5 \times 7 + 5; \quad 10 \times 5 = 7 \times 7 + 1; \quad 10 \times 1 = 1 \times 7 + 3; \ldots \]

So we have $8/7 = 1.142857142857\ldots$.

Proposition 5.8. Every periodic decimal is rational.

Proof. The idea is that if we can do some operations that make the decimal to be a finite decimal, then we are done since very finite decimal must be a rational number.

Let $x = a_0.a_1a_2\ldots a_kb_1b_2\ldots b_l$. We may assume $a_0 = 0$. Then

\[ 10^k x = a_0a_1a_2\ldots a_kb_1b_2\ldots b_l, \quad 10^{l+k} x = a_0a_1a_2\ldots a_kb_1b_2\ldots b_l.b_1b_2\ldots b_l, \]

and

\[ 10^{l+k} x - 10^k x = a_0a_1a_2\ldots a_kb_1b_2\ldots b_l - a_0a_1\ldots a_k. \]

\[ \square \]

Example 5.9. Consider $x = 0.\overline{123}$. Then

\[ 10 \times x = 1.\overline{23}, \quad 100 \times 10x = 123.\overline{23}. \]

This gives

\[ 990x = 122, \quad x = \frac{122}{990} = \frac{61}{495}. \]
6. Inequalities

An inequality is a statement about real numbers involving one of the symbols “>”, “\(\geq\)”, “<” or “\(\leq\)”; for example, \(x > 2\) or \(x^2 - 4y \leq 2x + 2\). In this chapter we shall present some elementary notions concerning manipulation of inequalities.

Here are the rules concerning the ordering of the real numbers.

**Rules:** given \(x, y \in \mathbb{R}\):

1. either \(x > 0\), \(x < 0\) or \(x = 0\);
2. if \(x > y\), then \(-x < -y\);
3. if \(x > y\), \(c \in \mathbb{R}\), then \(x + c > y + c\);
4. if \(x > 0\), \(y > 0\), then \(xy > 0\);
5. if \(x > y\), \(y > z\), then \(x > z\).

Note that (3) \(\Rightarrow\) (2), since if \(x > y\), then \(x + (-x - y) > y + (-x - y)\), and this implies \(-y > -x\).

In what follows (in class) we discussed Example 5.1 to Example 5.15 in textbook. (We omit them here since we followed exactly the same as in the textbook.)
7. Mathematical Induction

Mathematical induction is a mathematical proof technique. It is a very powerful tool to prove a statement involving positive integers.

Let us first consider the following statements:

(1) The sum of the first \( n \) positive odd integers is equal to \( n^2 \). That is

\[ \forall n \in \mathbb{N}, P(n). \]

Here, \( P(n) \) is “the sum of the first \( n \) positive odd integers is equal to \( n^2 \)."

(2) If \( p > -1 \) then \((1 + p)^n \geq 1 + np\).

We can easily check that statement (1) is true for many positive integers.

\[
1 = 1^2; 1 + 3 = 2^2; 1 + 3 + 5 = 3^2; 1 + 3 + 5 + 7 = 4^2.
\]

In order to prove this “for all” statement, we need to show it is true for all positive integer \( n \). There are infinitely many positive integers, and it is impossible to verify all of them. So how can we prove this statement for all \( n \)?

The answer is just the principle of mathematical induction. It is the following.

**Theorem 7.1.** (Principle of mathematical induction) Suppose that for each positive integer \( n \) we have a statement \( P(n) \). If we prove the following two things

(1) \( P(1) \) is true.

(2) for all \( n \in \mathbb{N}, \) if \( P(n) \) is true then \( P(n+1) \) is also true.

then \( P(n) \) is true for all positive integers \( n \).

**Remark 7.2.** The first step is called the base step and the second step is called the inductive step.

**Example 7.3.** The sum of the first \( n \) positive odd integers is equal to \( n^2 \).

**Proof.** The base step: this is trivial, since \( 1 = 1 \).

For the inductive step, we first assume that \( P(n) \) is true, where \( n \) is any fixed natural number, we need to prove that \( P(n + 1) \) is true. Since \( P(n) \) is true, we have

\[
1 + 3 + 5 + \ldots + (2n - 1) = n^2.
\]

This implies that

\[
1 + 3 + 5 + \ldots + (2n - 1) + (2n + 1) = n^2 + 2n + 1 = (n + 1)^2,
\]
which implies that $P(n+1)$ holds true.

By the principle of mathematical induction, we know that $P(n)$ is true for all positive integers $n$. □

The induction is like the **Domino effect**. If you have infinitely many dominoes on a very long table and you want to let them all fall, then how to do this? We know that as long as we can achieve the following two things

1. let the first domino fall;
2. make sure that for each domino, if it falls, it will hit the domino next to it and that domino falls.

So all dominoes will fall.

**Example 7.4.** If $p > -1$ then $(1 + p)^n \geq 1 + np$.

*Proof.* The base step is true, since $(1 + p)^1 = 1 + p \geq 1 + 1 \cdot p = 1 + p$.

Assume $(1 + p)^n \geq 1 + np$ is true, then since $(1 + p) > 0$, we have

$$(1 + p)^{n+1} \geq (1 + np)(1 + p) = 1 + (n + 1)p + np^2 \geq 1 + (n + 1)p.$$

So we proved the induction step, and hence finish the proof. □

Note that in the base step, the first case might not be starting from 1. Actually, it can start with 0 or any integers. It depends on the statement that we prove. For instance, we need to prove that “$P(n)$ is true for all integers $n \geq 3$”. This is because $P(1)$ and $P(2)$ make no sense or $P(1)$ and $P(2)$ are not true.

**Theorem 7.5.** (Principle of mathematical induction II) Let $k$ be any integer. Suppose that for each integer $n \geq k$ we have a statement $P(n)$. If

1. $P(k)$ is true.
2. for all integer $n \geq k$, $P(n) \Rightarrow P(n+1)$.

then $P(n)$ is true for all integers $n \geq k$.

**Example 7.6.** For every integer $n \geq 4$, $2^n < n!$.

*Proof.* Note that $2^4 = 8 > 3! = 6$. The base step: $P(4)$ is true. This follows easily since $2^4 = 16 < 4! = 24$.

The inductive step: we assume for $n \geq 4$, $P(n)$ is true, that is $2^n < n!$. Then we have

$$2^{n+1} = 2^n \times 2 < 2 \times n! < (n + 1)!.$$

□
7.1. **Guessing the answer.** You can see that the induction only tells you how to prove some statements involving positive integers, such as some identities. However, it doesn’t tell us how to find those identities. For example, consider the following identity
\[
\frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \cdots + \frac{1}{n \times (n+1)} = \frac{n}{n+1}
\]
Certainly, this is a statement involving positive integers and you can prove this by induction. But the question is if I don’t tell you the sum, how can you find it?

The most naive way is that you just guess the answer. Since it is a statement about positive integers, you can always try some simple cases, like \(n = 1, n = 2, n = 3\), as long as you try enough cases, you will notice a pattern. So you can guess the answer. How to verify your answer, you just prove if by induction.

So, induction does not tell you the answer, but it can help you to verify your guessing.

7.2. **The \(\Sigma\) notation.**

**Remark 7.7.** *(The Sigma notation)* Suppose we have a sequence of numbers \(a_1, a_2, \ldots, a_n\), we write the sum of all of them as
\[
a_1 + a_2 + \ldots + a_n = \sum_{k=1}^{n} a_k.
\]
Here, the summation begins with \(k = 1\) and stops at \(k = n\). For instance, if \(a_k = \frac{1}{k^2}\), then
\[
\sum_{k=1}^{n} \frac{1}{k^2} = \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2}.
\]
Actually, we can start and end with any integer.

7.3. **Principle of Strong Mathematical Induction.**

**Theorem 7.8.** *(Principle of Strong Mathematical Induction)* Let \(k\) be an integer. Suppose that for each integer \(n \geq k\) we have a statement \(P(n)\). If we prove the following two things:

(1) \(P(k)\) is true;

(2) for all \(n \geq k\), if \(P(i)\) are all true for \(k \leq i \leq n\), then \(P(n + 1)\) is also true;

then \(P(n)\) is true for all \(n \geq k\).
As an application, we prove the existence of prime factorization for integers.

**Definition 7.9.** A prime number is a positive integer \( p \) such that \( p \geq 2 \) and the only positive integers dividing \( p \) are 1 and \( p \).

**Proposition 7.10.** (Prime Factorization) Every positive integer greater than 1 is equal to a product of prime numbers.

We will prove in Chapter 11 that this factorization is unique.

**Proof.** Let \( P(n) \) be the statement that “\( n \) is equal to a product of prime numbers”.

The base step: \( P(2) \) is true. For the inductive step, suppose \( n \) is any positive integer greater than 1, and assume that \( P(2), \ldots, P(n) \) are all true. We need to show that \( P(n + 1) \) is true.

If \( n + 1 \) is a prime number, then we are done. If \( n + 1 \) is not a prime, by the definition of prime numbers, there is a positive integer \( a \) dividing \( n + 1 \) such that \( a \neq 1 \) and \( a \neq n + 1 \). Suppose \( n + 1 = ab \), where \( a, b \in \{2, 3, \ldots, n\} \).

By our assumption, \( P(a) \) and \( P(b) \) are both true, that is, \( a \) and \( b \) are equal to a product of prime numbers. Since \( n + 1 = ab \), we conclude that \( n + 1 \) is also equal to a product of prime numbers. This finishes the inductive step.

By the strong induction, we know \( P(n) \) is true for all \( n \geq 2 \). \( \square \)
8. Integers

Next, we discuss the integers. From this chapter, we will learn some basic stuff of number theory. For this chapter, our goal is to study how to find the common factors of two integers.

Let us start with some basic definitions.

**Definition 8.1.** Let \( a, b \in \mathbb{Z} \). We say that \( a \) (\( a \neq 0 \)) divides \( b \) (or \( a \) is a factor of \( b \)) if \( b = ac \) for some integer \( c \). When \( a \) divides \( b \), we write \( a \mid b \).

For instance, \( 2 \mid 4 \), \( -3 \mid -6 \), \( -5 \mid 10 \), \( 12 \mid 12 \). Note that 1 divides any integer and any integer divides itself.

Given two integers \( a \) and \( b \), in most cases, \( a \) will not divide \( b \). But we can still talk about “divide \( a \) into \( b \)” and get a quotient and a remainder.

**Proposition 8.2.** Let \( a \) be a positive integer. Then for any integer \( b \in \mathbb{Z} \), there are integers \( q, r \) such that

\[
b = qa + r \quad \text{and} \quad 0 \leq r < a.
\]

The integer \( q \) is called the quotient, and \( r \) is the remainder.

**Proof.** Let us consider the rational number \( \frac{b}{a} \). Since any number must lie between two consecutive integers, we assume that

\[
q \leq \frac{b}{a} < q + 1.
\]

Multiplying by the positive integer \( a \) (note that \( a \) is positive), we obtain

\[
qa \leq b < (q + 1)a.
\]

This implies that

\[
0 \leq b - qa < a.
\]

So we set \( r = b - qa \). \( \square \)

**Proposition 8.3.** Let \( a, b, d \in \mathbb{Z} \), and suppose that \( d \mid a \) and \( d \mid b \). Then \( d \mid (ma + nb) \) for any \( m, n \in \mathbb{Z} \).

Given two integers \( a \) and \( b \), we know that in most cases, \( a \) will not divide \( b \), but it is very likely that they share common factors.

Next, we will talk about the Euclidean Algorithm, which tells us how to compute the common factors of two integers.
Definition 8.4. Let $a, b \in \mathbb{Z}$. A common factor of $a$ and $b$ is an integer that divides both $a$ and $b$.

The highest common factor (greatest common divisor, gcd) of $a$ and $b$, when at least one of them is not zero, written $hcf(a, b)$, is the largest positive integer that divides both $a$ and $b$.

Note that $hcf(a, b) = hcf(-a, -b) = hcf(-a, b)$. Every time when we write $hcf(a, b)$, this implies that at least one of them is not zero.

For example, $hcf(2, 3) = 1$, $hcf(4, 8) = 4$, $hcf(0, 7) = 7$.

Let us do an example to illustrate the process.

Example 8.5. Find $hcf(5817, 1428)$. We let $d = hcf(5817, 1428)$.

Step 1: Divide 1428 into 5817 (let the smaller number divide the larger number)

$5817 = 4 \times 1428 + 105$

We get the quotient 4 and the remainder 105. Since $d|5817$ and $d|1428$, we have $d|105$.

Step 2: Divide 105 into 1428

$1428 = 13 \times 105 + 63$

Since $d|105$ and $d|1428$, we have $d|63$.

Step 3: Divide 63 into 105

$105 = 1 \times 63 + 42$

Since $d|105$ and $d|63$, we have $d|42$.

Step 4: Divide 42 into 63

$63 = 1 \times 42 + 21$

Since $d|63$ and $d|42$, we have $d|21$.

Step 5: Divide 21 into 42

$42 = 2 \times 21 + 0$

Now, we claim that $d = 21$. First, we have $d|21$. Next, we go through from the last step to the first step and we will get that 21 is a common factor of 1428 and 5817. So $d \geq 21$. As $d|21$, we have $d = 21$.

The general steps are as follows (assume that $a > 0$)

$b = q_1a + r_1, 0 \leq r_1 < a$

$a = q_2r_1 + r_2, 0 \leq r_2 < r_1$

$r_1 = q_3r_2 + r_3, 0 \leq r_3 < r_2$

$r_{n-2} = q_nr_{n-1} + r_n, 0 \leq r_n < r_{n-1}$
\[ r_{n-1} = q_{n+1}r_n + 0. \]

Therefore, the highest common factor \( \text{hcf}(a, b) = r_n \), the last non-zero remainder.

The Euclidean algorithm process is quite simple and you just do it step by step, but it has many important applications.

**Proposition 8.6.** If \( a, b \in \mathbb{Z} \) and \( \text{hcf}(a, b) = d \), then there are integers \( s \) and \( t \) such that \( d = sa + tb \).

This Proposition is extremely important, it contains almost all the information of highest common factor. Every time you see a statement involving hcf, you should first write this down and see whether you can use this to prove the statement.

**Proposition 8.7.** If \( a, b \in \mathbb{Z} \), then any common factor of \( a \) and \( b \) also divides \( \text{hcf}(a, b) \).

**Definition 8.8.** If \( a, b \in \mathbb{Z} \) and \( \text{hcf}(a, b) = 1 \), we say that \( a \) and \( b \) are **coprime** to each other.

**Proposition 8.9.** Let \( a, b \in \mathbb{Z} \).

1. Suppose \( c \) is an integer such that \( a, c \) are coprime to each other, and \( c|ab \). Then \( c|b \).
2. Suppose \( p \) is a prime number and \( p|ab \). Then either \( p|a \) or \( p|b \).

**Proof.** For (1), since \( \text{hcf}(a, c) = 1 \), we have

\[ 1 = sa + tc \]

for some integers \( s \) and \( t \). Multiplying by \( b \) gives

\[ b = sab + tcb. \]

For (2), we show that if \( p \) does not divide \( a \), then \( p|b \). Since \( p \) is a prime, the only positive factors of \( p \) are 1 and \( p \). This implies \( \text{hcf}(a, p) = 1 \). Then part (1) gives the result.

**Fact 8.10.** Let \( p \) be a prime number and \( a \) be an integer. Then either \( \text{hcf}(a, p) = 1 \) or \( \text{hcf}(a, p) = p \).

**Proposition 8.11.** Let \( a_1, a_2, \ldots, a_n \in \mathbb{Z} \), and let \( p \) be a prime number. If \( p|a_1a_2\cdots a_n \), then \( p|a_i \) for some \( i \).

This is proved by induction.
9. PRIME FACTORIZATION

We have seen that any integer greater than 1 is equal to a product of prime numbers; that is, it has a prime factorization.

In this Chapter, we will show that this factorization is unique.

This result is so important and fundamental, so we call it “The fundamental Theorem of Arithmetic”. It is the following

Theorem 9.1. Let \( n \) be an integer with \( n \geq 2 \).

1. Then \( n \) is equal to a product of prime numbers, that is

\[
  n = p_1 p_2 \ldots p_k,
\]

where \( p_1, \ldots, p_k \) are primes and \( p_1 \leq p_2 \leq \ldots \leq p_k \).

2. This prime factorization is unique: in other words, if

\[
  n = p_1 \ldots p_k = q_1 \ldots q_l
\]

where \( p_i \) and \( q_i \) are all primes and \( p_1 \leq p_2 \leq \ldots \leq p_k \) and \( q_1 \leq q_2 \leq \ldots \leq q_l \), then

\[
  k = l, \; \text{and} \; p_i = q_i, \; 1 \leq i \leq k.
\]

This theorem tells us if you write down the primes in the factorization in an increasing order, then the writing or expression is unique.

Proof. We have already proved part (1).

For part (2), we prove by contradiction. Suppose there is an integer \( n \) does not satisfy the property. Then in the equation

\[
  p_1 \ldots p_k = q_1 \ldots q_l,
\]

we cancel any primes that are common on both sides. Since we assume that the two factorizations are different, not all primes cancel, and we end up with an equation

\[
  r_1 \ldots r_a = s_1 \ldots s_b.
\]

and

\[
  r_i \neq s_j.
\]

Since \( r_1|s_1 \ldots s_b \), by the previous proposition, we know that \( r_1|s_j \) for some \( j \). But \( s_j \) is prime, its positive factors are only 1 and itself. Hence, \( r_1 = s_j \). But this is a contradiction. \( \square \)
By the above theorem, since some of the primes may be equal, we can write the factorization as the following form (we assume that \( n \) is positive)

\[
n = p_1^{a_1} p_2^{a_2} \cdots p_m^{a_m},
\]

where \( p_1 < p_2 < \ldots < p_m \) and \( a_i \) are positive integers.

For instance,

\[
28 = 2^2 \times 7, \quad 36 = 2^2 \times 3^2.
\]

Next, we will show some consequences of the Fundamental Theorem of Arithmetic”.

The first one is about the factors of a given integer.

**Proposition 9.2.** Let \( n \) be a positive integer with \( n \geq 2 \) and

\[
n = p_1^{a_1} p_2^{a_2} \cdots p_m^{a_m},
\]

where \( p_1 < p_2 < \ldots < p_m \) and \( a_i \in \mathbb{N} \). If \( s \) divides \( n \), then \( s \) can be written as

\[
s = p_1^{b_1} \cdots p_m^{b_m}
\]

with \( 0 \leq b_i \leq a_i \).

**Proof.** We assume the prime factorization of \( s \) is the following

\[
s = q_1^{c_1} \cdots q_k^{c_k}.
\]

We only need to prove that every \( q_i \) must be equal to some \( p_j \) and the corresponding powers satisfy \( c_i \leq a_j \). Since \( s \mid n \), there exists an integer \( t \) such that \( n = st \). \( \square \)

**Definition 9.3.** Let \( a \) and \( b \) be two positive integers. Define the **least common multiple** of \( a \) and \( b \), denoted by \( \text{lcm}(a,b) \), to be the smallest positive integer that is divisible by both \( a \) and \( b \).

For instance, \( \text{lcm}(15, 21) = 105 \), \( \text{lcm}(1, 20) = 20 \), \( \text{lcm}(17, 34) = 34 \).

The next proposition tells us how to determine the hcf and lcm of tow given integers.

**Proposition 9.4.** Let \( a, b \geq 2 \) be two integers with prime factorizations

\[
a = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}, \quad b = p_1^{s_1} p_2^{s_2} \cdots p_m^{s_m},
\]

where \( p_1 < p_2 < \ldots < p_m \) and all \( r_i, s_i \geq 0 \) (note that we allow some of the \( r_i \) and \( s_i \) to be \( 0 \)). Let

\[
c_i = \min\{r_i, s_i\}, \quad d_i = \max\{r_i, s_i\}.
\]

Then
Proof. For part (1), we may assume that
\[ \text{hcf}(a, b) = p_1^{e_1} \cdots p_m^{e_m}. \]
By the previous Proposition, \( e_i \leq r_i \) and \( e_i \leq s_i \). Hence, \( e_i \leq c_i \).
Note that the RHS is a common factor of \( a \) and \( b \), so RHS divides \( \text{hcf}(a, b) \) (since every common factor divides the \( \text{hcf} \)) and \( c_i \leq e_i \).
For part (2), we may assume that
\[ \text{lcm}(a, b) = p_1^{l_1} \cdots p_m^{l_m}. \]
Since \( a \mid \text{lcm}(a, b) \) and \( b \mid \text{lcm}(a, b) \), by the previous proposition, we have \( l_i \geq d_i \). Hence, \( \text{lcm}(a, b) \geq p_1^{d_1} \cdots p_m^{d_m} \). As the RHS is a common multiple of \( a \) and \( b \), we have \( \text{RHS} = \text{lcm}(a, b) \).
For part (3), it follows easily from the fact that
\[ c_i + d_i = r_i + s_i. \]
\( \square \)

**Example 9.5.** Consider \( 90 = 2 \times 3^2 \times 5 \) and \( 84 = 2^2 \times 3 \times 7 \). We can rewrite them as
\[ 90 = 2^1 \times 3^2 \times 5^1 \times 7^0, \quad 84 = 2^2 \times 3^1 \times 5^0 \times 7^1. \]
Hence, we have
\[ \text{hcf} = 2^1 \times 3^1 \times 5^0 \times 7^0 = 6, \]
\[ \text{lcm} = 2^2 \times 3^2 \times 5^1 \times 7^1 = 1260, \]
and
\[ 6 \times 1260 = 90 \times 84. \]

**Proposition 9.6.** Let \( n \) be a positive integer. Then \( \sqrt{n} \) is rational if and only if \( n \) is a perfect square (i.e., \( n = m^2 \) for some integer \( m \)).
Proof. The if part is trivial. We prove the only if part. So we assume that $\sqrt{n}$ is rational, that is, $$\sqrt{n} = \frac{r}{s},$$ where $r, s \in \mathbb{N}$ and $\frac{r}{s}$ is in lowest terms. Taking square, we get $$ns^2 = r^2.$$ Note that each prime in the factorization of $r^2$ has even power. This also holds for $s^2$. Hence, each prime factor of $n$ also have even power. Otherwise, suppose $q | n$ and $q$ has odd power. On the RHS, $q$ occurs to an even power, but on the LHS, $q$ occurs to an odd power, which contradicts the fundamental theorem of arithmetic. $\square$

**Proposition 9.7.** Let $a$ and $b$ be positive integers that are coprime to each other.

(1) If $ab$ is a square, then both $a$ and $b$ are also squares.

(2) More generally, if $ab$ is an $n$-th power (for some positive integer $n$), then both $a$ and $b$ are also $n$-th powers.

*Proof.** We only prove part (1). Let the prime factorizations of $a$ and $b$ be $$a = p_1^{d_1} \cdots p_k^{d_k}, \quad b = q_1^{e_1} \cdots q_l^{e_l},$$ where $p_1 < p_2 < \ldots < p_k$ and $q_1 < \ldots < q_l$.

Since $a$ and $b$ are coprime to each other, that means $hcf(a, b) = 1$, this implies that none of the $p_i$ are equal to the $q_i$, i.e., $p_i \neq q_j$. As $ab$ is a square, then $ab = c^2$ for some positive integer $c$. Assume that $c$ has prime factorization $$c = r_1^{f_1} \cdots r_m^{f_m}.$$ Then $$p_1^{d_1} \cdots p_k^{d_k} q_1^{e_1} \cdots q_l^{e_l} = r_1^{2f_1} \cdots r_m^{2f_m}.$$ By the fundamental theorem of arithmetic, each $p_i$ is equal to some $r_j$ and $d_i = 2f_j$, and similarly for $q_i$ and their powers. $\square$

**Theorem 9.8.** There are infinitely many prime numbers.

Just as we talked before, if you want to prove this directly, you need to list infinitely many prime numbers or you may write down a general formula that produces infinitely many primes. But that is possible and indeed this theorem can be proved easily by proof by contradiction.
Proof. We prove by contradiction. Assume the statement is false, so there are only finitely many primes. Assume they are

\[ p_1, p_2, \ldots, p_n. \]

Let us consider the following integer

\[ N = 1 + p_1 p_2 \ldots p_n. \]

By the prime factorization, \( N \) is equal to a product of primes, say \( N = q_1 \ldots q_l \), where \( q_i \) is prime. As \( q_1 \) is prime, it must belong to the above list, so \( q_1 = p_i \) for some \( i \).

Now, \( q_1 | N \) and \( q_1 | p_1 p_2 \ldots p_n \). This implies \( q_1 | 1 \) which is impossible. This is a contradiction. \( \square \)
10. Equivalence relations

Before we talk about the equivalence relations, we need to introduce a definition.

**Definition 10.1.** Let $A$ and $B$ be two sets. The **Cartesian product** of $A$ and $B$, written $A \times B$, is the set consisting of all symbols of the form $(a, b)$ with $a \in A$ and $b \in B$.

Such a symbol $(a, b)$ is called an **ordered** pair of elements of $A$ and $B$. Two ordered pairs $(a, b), (a', b')$ are deemed to be equal if and only if both $a = a'$ and $b = b'$.

**Example 10.2.** Let $A = \{1, 2\}$ and $B = \{1, 2, 3\}$. Then $A \times B$ has six ordered pairs

$$(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3).$$

Note that $A \times B$ and $B \times A$ are different.

**Definition 10.3.** Let $S$ be a set. A **relation** on $S$ is defined as follows. We choose a subset $R$ of the Cartesian product $S \times S$; in other words, $R$ consists of some of the ordered pairs $(s, t)$ with $s, t \in S$. For those ordered pairs $(s, t) \in R$, we write $s \sim t$ and say that $s$ is related to $t$.

As for $(s, t) \notin R$, we write $s \not\sim t$. Thus, the symbol $\sim$ relates various pairs of elements of $S$. It is called a **relation** on $S$.

**Example 10.4.** Consider the following examples.

(1) Let $S$ be a set and $a \sim b \iff a = b$. Hence, $R = \{(a, a) \in S \times S \mid a \in S\}$. This is an equivalence relation.

(2) Let $S = \mathbb{R}$ and define $a \sim b \iff a < b$. Then

$$R = \{(a, b) \in \mathbb{R} \times \mathbb{R} \mid a, b \in \mathbb{R}, a < b\}.$$  

This is Not reflexive and symmetric.

(3) Let $S = \{1, 2\}$. Set $R = \{(1, 1), (1, 2)\} \subseteq S \times S$. Then

$$1 \sim 1, 1 \sim 2, 2 \sim 1, 2 \sim 2.$$  

This is Not reflexive.

(4) Let $S$ be all the students at UCSB, and define $a \sim b$ if and only if $a$ is the boyfriend of $b$. So $a$ is male and $b$ is female. Note that $a \sim a$. If $a \sim b$, then $b \sim a$.

This is Not.
(5) Let $S = \mathbb{Z}$ and let $m$ be a positive integer. Define $a \sim b$ if and only if $a \equiv b \mod m$. Then $R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a \equiv b \mod m\}$.

This is an equivalence relation.

(6) Let $S$ be all the people in US. Define $a \sim b$ if and only if $a$ and $b$ have the same father.

This is an equivalence relation.

(7) Let $S$ be all the students in UCSB, and define $a \sim b$ if and only if $a$ and $b$ are classmates.

This is Not transitive.

(8) Let $S = \mathbb{R}$. Define $a \sim b \iff a - b \in \mathbb{Z}$.

This is an equivalence relation.

In general, a relation on a set $S$ is just a subset of $S \times S$, and there is nothing much to say about them. But we will see that a certain type of relations satisfying some interesting properties. This type of relation is called the **equivalence relations** and they are the most important type of relations.

**Definition 10.5.** Let $S$ be a set, and let $\sim$ be a relation on $S$. Then $\sim$ is an **equivalence relation** if the following three properties hold for all $a, b, c \in S$:

1. $a \sim a$ ($\sim$ is reflexive);
2. If $a \sim b$, then $b \sim a$ ($\sim$ is symmetric);
3. If $a \sim b$ and $b \sim c$, then $a \sim c$ ($\sim$ is transitive).

**Definition 10.6.** Let $S$ be a set and $\sim$ an equivalence relation on $S$. For $a \in S$, define

$$\text{cl}(a) = \{s \in S \mid s \sim a\}.$$  

Thus, $\text{cl}(a)$ is the set of objects that are related to $a$. The subset $\text{cl}(a)$ is called an **equivalence class** of $\sim$. The equivalence classes of $\sim$ are the subsets $\text{cl}(a)$ as $a$ ranges over the elements of $S$.

**Remark 10.7.** We have the following remarks for the equivalence classes:

1. Every element of $S$ belongs to exactly one equivalence.
2. Even if $a \neq b$, $\text{cl}(a)$ and $\text{cl}(b)$ can be equal.
3. $\text{cl}(a) = \text{cl}(b)$ if and only if $a \sim b$.

**Example 10.8.** Let $m$ be a positive integer, and let $\sim$ be the equivalence relation on $\mathbb{Z}$ defined as

$$a \sim b \iff a \equiv b \mod m.$$
Then the equivalence classes are
\[ cl(0) = \{ n \in \mathbb{Z} : n \equiv 0 \mod m \}, \]
\[ cl(1) = \{ n \in \mathbb{Z} : n \equiv 1 \mod m \}, \ldots, \]
\[ cl(m-1) = \{ n \in \mathbb{Z} : n \equiv m - 1 \mod m \}. \]

We claim that these are all the equivalence classes.

For any integer \( n \), then there are integers \( q, r \) such that \( n = qm + r \) with \( 0 \leq r \leq m - 1 \). We know that
\[ n \equiv r \mod m, \]
then \( n \in cl(r) \) and \( cl(n) = cl(r) \). Hence, any equivalence class \( cl(n) \) is equal to one of those listed above.

In particular, let \( m = 2 \). Then the equivalence classes
\[ cl(0) = \{ n \in \mathbb{Z} : n \equiv 0 \mod 2 \}, \]
\[ cl(1) = \{ n \in \mathbb{Z} : n \equiv 1 \mod 2 \}. \]

Definition 10.9. (Partition) A partition of a set \( S \) is a collection of subsets \( S_1, \ldots, S_k \) such that each element of \( S \) lies exactly one of these subsets.

Another way of putting this is that the subsets \( S_1, \ldots, S_k \) have the properties that their union is \( S \) and any two of them are disjoint, i.e., \( S_i \cap S_j = \emptyset \) for any \( i \neq j \).

Example 10.10. Let \( S = \{1, 2, 3, 4\} \). Then the subsets
\[ \{1\}, \{2\}, \{3, 4\} \]
forms a partition of \( S \), while the subsets
\[ \{1\}, \{1, 2\}, \{3, 4\} \]
or the subsets
\[ \{1\}, \{3\}, \{4\} \]
do not.

Example 10.11. Consider the following examples:
(1) Let \( S \) be a set and \( a \sim b \iff a = b \). Then \( cl(a) = \{a\} \).
(2) Let \( S = \mathbb{R} \). Define \( a \sim b \iff a - b \in \mathbb{Z} \). Then
\[ cl(a) = \{a + n \mid n \in \mathbb{Z}\}. \]
And \( cl(x) \) is a partition of \( \mathbb{R} \) for \( 0 \leq x < 1 \).
**Example 10.12.** Fix $m \in \mathbb{N}$. Then 
\[ \text{cl}(0), \text{cl}(1), \text{cl}(2), \ldots, \text{cl}(m-1) \]
is a partition of $\mathbb{Z}$.

**Proposition 10.13.** Let $S$ be a set and $\sim$ an equivalence relation on $S$. Then the equivalence classes of $\sim$ form a partition of $S$.

**Proof.** We need to show two things: first, every element of $S$ belongs to an equivalence classes; second, every element lies in only one equivalence class.

Since $a \sim a$, then $a \in \text{cl}(a)$. Next, we show that if $a \in \text{cl}(s)$ and $a \in \text{cl}(t)$, then $\text{cl}(s) = \text{cl}(t)$.

For any $x \in \text{cl}(s)$, we have $x \sim a$. Note that $a \sim t$, then $x \sim t$. This implies $x \in \text{cl}(t)$ and hence $\text{cl}(s) \subseteq \text{cl}(t)$. Similarly, we have $\text{cl}(t) \subseteq \text{cl}(s)$. \qed

Indeed, the converse of the above Proposition is also true, that is, if $S$ is a set and $S_1, \ldots, S_k$ is a partition of $S$, then there is a unique equivalence relation on $S$ which has the $S_i$ as its equivalence classes. This relation is defined as follows: for $x, y \in S$,
\[ x \sim y \iff \exists i \text{ such that } x, y \in S_i. \]
11. Functions

Definition 11.1. Let $S$ and $T$ be two sets. A function from $S$ to $T$, denoted by $f : S \to T$, is a rule that assigns to each $s \in S$ a single element of $T$, denoted by $f(s)$.

The set of inputs is called the domain. The set $T$ is called the codomain of $f$. The image of $f$ is the set of all elements of $T$ that are equal to $f(s)$ for some $s \in S$, i.e.,

$$f(S) = \{ f(s) \mid s \in S \} \subseteq T.$$ 

Remark 11.2. We have the following remark.

1. The image of $f$ may not be equal to the whole set $T$.
2. For each element $s \in S$, the function $f$ associates or selects a unique element in $T$. The uniqueness condition does not allow $s$ to be assigned to distinct elements of $T$.
3. It does allow different elements of $S$ to be assigned to the same element of $T$.

Remark 11.3. A function $f : S \to T$ is well-defined if

1. We know what $f(s)$ is for every $s \in S$.
2. It never maps $s \in S$ to two (or more) different elements of $T$.
3. $f(s) \in T$ for every $s \in S$.

Example 11.4. Consider the following examples.

1. Let $X = \{1, 2, 3\}$ and $Y = \{0, 1, \pi\}$. Let $f(1) = f(2) = 0$ and $f(3) = \pi$. $f$ is not one-to-one and onto.
2. Let $X = \mathbb{R} = Y$ and $f(x) = x^2$. The image of $f$ is $f(\mathbb{R}) = \{ y \in \mathbb{R} \mid y \geq 0 \}$. $f$ is not one-to-one and onto.
3. Define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ by $f(m, n) = m - n$. The image is $\mathbb{Z}$. $f$ is not one-to-one, but $f$ is onto.
4. Fix an integer $m \in \mathbb{N}$. Let $S = \mathbb{Z}$ and $T = \{0, 1, 2, \ldots, m - 1\}$. Define

$$f(n) = r, \text{ where } n \equiv r \mod m, r \in T.$$ 

$f$ is onto.

Definition 11.5. Let $f : S \to T$ be a function. We say that $f$ is onto if the image $f(S) = T$, i.e., for every element $t \in T$, there exists $s \in S$ such that $f(s) = t$. In this case, we say that $f$ is surjective, or $f$ is a surjection.
We say that $f$ is one-to-one if whenever $x, y \in S$ with $s \neq t$, then $f(x) \neq f(y)$, that is, $f$ sends different elements of $S$ to different elements of $T$. Another way of saying this is that for any $x, y \in S$, if $f(x) = f(y)$ then $x = y$. In this case, we say that $f$ is injective, or $f$ is a surjection.

We say $f$ is a bijection if $f$ is both onto and one-to-one.

**Example 11.6.** Let $A$ be an $m \times n$ matrix. Consider the linear transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m, \quad T(x) = Ax.$$  

Then

1. $T$ is onto $\iff$ The equation $T(x) = b$ has a solution for every $b \in \mathbb{R}^m$. $\iff$ The columns of $A$ span $\mathbb{R}^m$. That is: every $b \in \mathbb{R}^m$ is a linear combination of the columns of $A$. $\iff$ $\text{rank}(A) = m$.

2. $T$ is one-to-one $\iff$ The equation $T(x) = 0$ has only the trivial solution $x = 0$. $\iff$ The columns of $A$ are linearly independent. $\iff$ $\text{rank}(A) = n$.

**Proposition 11.7.** Let $S$ and $T$ be finite sets. Let $f : S \to T$ be a function.

1. If $f$ is onto, then $|S| \geq |T|$.
2. If $f$ is one-to-one, then $|S| \leq |T|$.
3. If $f$ is a bijection, then $|S| = |T|$.

**Example 11.8.** Part (2) of the above proposition implies that if $|S| > |T|$, then there is no $1-1$ function from $S$ to $T$. This gives the **Pigeonhole Principle**:

If we put $n + 1$ or more pigeons into $n$ pigeonholes, then there must be a pigeonhole containing more than one pigeon.

This can be proved easily by contradiction. If you use the Pigeonhole Principle in a right way, then it will be a very powerful tool and will give us some surprising results. The key is to choose the appropriate pigeons and pigeonholes.

**Example 11.9.** For any 6 integers, there must be two integers whose difference is divisible by 5. View the 6 integers as the pigeons, and their remainders on division by 5 as the pigeonholes. Then there is a pigeonhole containing at least two pigeon (integers), and they have the same remainder, i.e., their difference is a multiple of 5.

Given a function $f : S \to T$, for $t \in T$, we may want to figure out $s$ such that $f(s) = t$. In order to make this “inverse” of $f$ to be a function, we need the following two things:
(1) For every element \( t \in T \), we can find \( s \in S \) such that \( f(s) = t \). In other words, \( f \) is onto.

(2) For every element \( t \in T \), we can only find exactly one element \( s \) such that \( f(s) = t \), that is, \( f \) is one-to-one.

Therefore, in order to define an inverse function of \( f \), we need \( f \) to be a bijection.

**Definition 11.10.** Let \( f \) be a function from \( S \) to \( T \). If \( f \) is a bijection, then there exists a function \( f^{-1} : T \to S \), called the inverse function of \( f \) satisfying

\[
f^{-1}(t) = s \iff f(s) = t.
\]

Hence, \( f^{-1}(f(s)) = s \) and \( f(f^{-1}(t)) = t \).

**Example 11.11.** Consider the following examples:

(1) Define \( f : \mathbb{R} \to \mathbb{R} \) by \( f(x) = 8 - 2x \). Then \( f \) is a bijection and 
\[
f^{-1}(t) = \frac{1}{2}(8 - t).
\]

(2) Define \( f : \mathbb{R} \to \mathbb{R} \) by \( f(x) = x^3 \). Then \( f \) is a bijection and 
\[
f^{-1}(t) = \sqrt[3]{t}.
\]

(3) Define \( f : \mathbb{R}_+ \to \mathbb{R}_+ \) by \( f(x) = x^2 \). Then \( f \) is a bijection and 
\[
f^{-1}(t) = \sqrt{t}.
\]

**Definition 11.12.** Let \( S, T, U \) be sets and let \( f : S \to T \) and \( g : T \to U \) be functions. The **composition** of \( f \) and \( g \) is the function \( g \circ f : S \to U \) ("\( g \) composed with \( f \)") which is defined by

\[
(g \circ f)(s) = g(f(s)), \forall s \in S.
\]

**Remark 11.13.** We have

- In order to make \( g \circ f \) meaningful, the image of \( f \) must be contained in the domain of \( g \).
- \( f \circ g \) and \( g \circ f \) are totally different. Sometimes, they may not be well-defined.

**Example 11.14.** Consider the following examples:

(1) Let \( f(x) = \sin x \), \( g = x^2 + 1 \). Then
\[
g(f(x)) = \sin^2 x + 1, \ f(g(x)) = \sin(x^2 + 1).
\]

(2) Let \( f(x) = x^2 \) and \( g(x) = \sqrt{x} \). Then \( g(f(x)) = |x| \). But \( f(g(x)) = x \).
\[
f : \mathbb{R} \to \mathbb{R}_+, \ g : \mathbb{R}_+ \to \mathbb{R}.
\]

(3) Let \( f(x) = x + 1 \) and \( g(x) = \sqrt{x} \). Then \( f(g(x)) = \sqrt{x} + 1 \), but \( g \circ f \) does not exist.
(4) Let \( f : S \to T \) be a bijection, then the inverse function of \( f \) exists and we have
\[
    f^{-1} \circ f(s) = s, \quad f \circ f^{-1}(t) = t.
\]
That is, their composition are identity functions.

**Proposition 11.15.** Let \( S, T, U \) be sets and let \( f : S \to T \) and \( g : T \to U \) be functions.

1. If \( f \) and \( g \) are both one-to-one, so is \( g \circ f \).
2. If \( f \) and \( g \) are both onto, so is \( g \circ f \).
3. If \( f \) and \( g \) are both bijections, so is \( g \circ f \).
Next, we will talk about infinity. If we have two finite sets, then it is very easy to compare their sizes, i.e., the number of elements of the sets.

But what if these two sets are infinite sets, that is, they have infinite number of elements, such as $\mathbb{N}$, $\mathbb{R}$, $\mathbb{Q}$. Our old definition of sizes only applies to finite sets, it does not work for infinite sets. Since the integers are only part of the real numbers $\mathbb{R}$, we should feel that the set of real numbers should be larger than the set of integers. But they are all infinite, and how can we compare their sizes? So we really need to find a new way to define the size of a set, or we need to extend our definition of the size of a set.

We will use functions to define the size of sets. Let us start with finite set. Suppose we have a set $S$ of size $n$, that is

$$S = \{a_1, a_2, \ldots, a_n\}.$$ 

This means that we can find a function $f$ from $S$ to the set $\{1, 2, 3, \ldots, n\}$ such that

$$f(a_1) = 1, f(a_2) = 2, f(a_3) = 3, \ldots, f(a_n) = n,$$

and the function $f$ is a bijection. In this way, $S$ and $\{1, 2, 3, \ldots, n\}$ have the same size. Hence, we can re-define the meaning of the size of a set, that is,

$S$ has size $n \iff$ there is a bijection from $S$ to $\{1, 2, 3, \ldots, n\}.$

And we can extend this definition to arbitrary sets.

**Definition 12.1.** Two sets $A$ and $B$ are said to be equivalent to each other if there is a bijection from $A$ to $B$. We write $A \sim B$ if $A$ and $B$ are equivalent to each other.

**Remark 12.2.** We may think of two sets that are equivalent to each other as “having the same size”. So the above definition tells us when two sets are of the same size.

**Proposition 12.3.** The relation $\sim$ defined above is an equivalence relation.

**Example 12.4.** Let $A = \mathbb{N}$ and $B = \{2n \mid n \in \mathbb{N}\}$. It is obvious that $B$ is a subset of $A$ and $A$ contains more elements than $B$. If we use our old definition of size, then it seems that $A$ is larger than $B$. But our old definition of size only applies to finite sets, and we cannot apply it to infinite sets. Let us use the extended definition of size. Consider the following bijection

$$f : A \to B, \quad f(n) = 2n.$$
Thus, $A \sim B$, that is, they have the same size. Now, we have a set which has the same size with its subset. Of course, this cannot happen for finite sets, but it is possible for infinite sets having this property.

A related example is Hilbert’s hotel. It is demonstrated that a fully occupied hotel with infinitely many rooms may still accommodate additional guests, even infinitely many of them.

Next, we will say that if a set $A$ is equivalent to $\mathbb{N}$, i.e., $A$ has the same size with $\mathbb{N}$, then $A$ is countable.

**Definition 12.5.** A set $A$ is said to be countable if $A$ is equivalent to $\mathbb{N}$, that is, there is a bijection $f$ from $\mathbb{N}$ to $A$. Let $f(n) = a_n$. Since $f$ is onto, we then have

$$A = \{a_1, a_2, a_3, \ldots\}.$$ 

In other words, $A$ is countable if it is an infinite set and we can list all elements of $A$ as $\{a_1, a_2, \ldots\}$.

Note that, in some textbooks, finite sets can also be viewed as countable sets.

**Example 12.6.** Consider the following examples.

1. Obvious, $\mathbb{N}$ is countable.
2. The set $\{2n : n \in \mathbb{N}\}$ is countable.
3. The set $\{2n - 1 : n \in \mathbb{N}\}$ is countable.
4. The set $\{n^2 : n \in \mathbb{N}\}$ is countable.
5. A less obvious example is $\mathbb{Z}$. We can list all elements of $\mathbb{Z}$ as 0, 1, -1, 2, -2, 3, -3, .... The corresponding bijection $f : \mathbb{N} \rightarrow \mathbb{Z}$ is

$$f(2n) = n, f(2n - 1) = -(n - 1).$$

**Remark 12.7.** There are two ways to show that a set $A$ is countable:

1. Find a bijection $f$ from $\mathbb{N}$ to $A$.
2. We can list all elements of $A$ as $\{a_1, a_2, \ldots\}$.

**Fact 12.8.** If $A$ is a countable set and $B$ is a finite set, then $A \cup B$ is countable.

**Fact 12.9.** The union of any two countable sets is countable. That is, if $A$ and $B$ are countable sets, then $A \cup B$ is countable.

**Proof.** Let $A = \{a_1, a_2, \ldots\}$ and $B = \{b_1, b_2, \ldots\}$. Then we can $A \cup B$ as

$$a_1, b_1, a_2, b_2, a_3, b_3, \ldots$$
Let $f : \mathbb{N} \to A$ and $g : \mathbb{N} \to B$. Then the new function $h : \mathbb{N} \to A \cup B$ is given by
\[
h(2n - 1) = f(n), \quad h(2n) = g(n).
\]

**Proposition 12.10.** Every infinite subset of $\mathbb{N}$ is countable.

**Proof.** Let $S$ be an infinite subset of $\mathbb{N}$. Take $s_1$ to be the smallest integer in $S$; then take $s_2$ to be the smallest integer in $S - \{s_1\}$. In this way, we can list $S$ as
\[
s_1, s_2, s_3, \ldots.
\]

**Remark 12.11.** Indeed, every infinite subset of $\mathbb{Z}$ is also countable.

Let us consider the question of whether $\mathbb{Q}$ is countable. We cannot list the positive rationals in an increasing order, since no matter what rational we started the list with, there would be a smaller one.

**Proposition 12.12.** The set of rationals $\mathbb{Q}$ is countable.

Note that by the definition, in order to show that a set is countable, we need to find a bijection between this set and $\mathbb{N}$. But the following proposition gives a weaker condition that we only need to find a one-to-one function instead of a bijection.

**Proposition 12.13.** Let $S$ be an infinite set. If there is a one-to-one function $f : S \to \mathbb{N}$, then $S$ is countable.

**Proof.** Recall the image of $f$ is the set
\[
f(S) = \{f(s) \mid s \in S\} \subseteq \mathbb{N}.
\]
Since $f$ is one-to-one, $f(S)$ is an infinite set. Therefore, $f(S)$ is countable. There is a bijection $g : \mathbb{N} \to f(S)$.

We can regard $f$ as a function from $S$ to $f(S)$. Then, $f$ is onto and one-to-one, hence $f$ is a bijection from $S$ to $f(S)$. Then the composition
\[
f^{-1} \circ g : \mathbb{N} \to S
\]
is a bijection.

**Remark 12.14.** Note that the function is from $S$ to $\mathbb{N}$. This proposition is very powerful in proving that a set is countable.
Example 12.15. Consider the following examples

1) Define \( f : \mathbb{Q}^+ \to \mathbb{N} \) by
\[
f\left(\frac{m}{n}\right) = 2^m3^n,
\]
where \( m, n \in \mathbb{N} \) and \( \frac{m}{n} \) is in lowest terms. Here, we can choose any two positive integers that are co-prime to each other.

2) We will show the Cartesian product \( \mathbb{N} \times \mathbb{N} \) is countable. Define \( f : \mathbb{N} \times \mathbb{N} \to \mathbb{N} \) by
\[
f(m, n) = 2^m3^n.
\]
Moreover, \( \mathbb{N} \times \mathbb{N} \times \mathbb{N} \) is also countable.

3) In general, if \( S \) and \( T \) are countable sets, then the Cartesian product \( S \times T \) is also countable.

Not every infinite set is countable, so we say that a set is **uncountable** if it is an infinite set and it is not countable. There are many examples of uncountable sets and the most important one is the real numbers.

Theorem 12.16. The set \( \mathbb{R} \) of all real numbers is uncountable.

Proof. See [2, Theorem 21.1]. \( \Box \)

A direct consequence of the above theorem is that the set of irrational numbers \( \mathbb{R} - \mathbb{Q} \) is uncountable.

Now we give a definition of the size for all sets.

Definition 12.17. Let \( A \) and \( B \) be two sets. If \( A \) and \( B \) are equivalent to each other, i.e., there is a bijection from \( A \) to \( B \), we say that they have the same **cardinality**, and write \( |A| = |B| \).

If there is a 1-1 function from \( A \) to \( B \), we write \( |A| \leq |B| \).

If there is a 1-1 function from \( A \) to \( B \), but no bijection from \( A \) to \( B \), we write \( |A| < |B| \), and say that \( A \) has smaller cardinality than \( B \).

Remark 12.18. We have the following:

1) If a set \( S \) is finite with \( n \) elements, then the cardinality of \( S \) is just the number of elements of \( S \), i.e., \( |S| = n \).

2) we have
\[
|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{Q}| = \aleph_0,
\]
i.e., all the countable sets have the same cardinality.
(3) we have
\[ |\mathbb{N}| < |\mathbb{R}|. \]

Note that the real number has the same cardinality with the irrational numbers, i.e., \(|\mathbb{R}| = |\mathbb{R} - \mathbb{Q}|.|\)

(4) Any set \(S\) with cardinality greater than that of the natural numbers, or \(|X| > |\mathbb{N}|\), is said to be uncountable.

(5) If \(X\) is an infinite set, then \(|\mathbb{N}| \leq |X|.|\)

(6) There is no cardinal number between the cardinality of the real numbers and the cardinality of the natural numbers.

Now we know that all countable sets have the same cardinality with the natural numbers \(\mathbb{N}\), but for uncountable sets, they have different cardinality, that is, there is an uncountable set with cardinality greater than that of the real numbers \(\mathbb{R}\). In general, suppose we have a set. Then there is a way that we can always generate a set with larger cardinality. This is the following.

**Theorem 12.19.** Let \(S\) be a set. Then there is no bijection from \(S\) to \(\mathcal{P}(S)\). Consequently, \(|S| < |\mathcal{P}(S)|.|\)

**Proof.** See [2, Proposition 21.5]. \(\square\)

**Remark 12.20.** We have the following

1. If \(S\) is a finite set, then the above theorem is obvious. If \(|S| = n\), then \(\mathcal{P}(S) = 2^n\).

**Definition 12.21.** Let \(S\) be a set. Then the power set of \(S\), \(\mathcal{P}(S)\) is defined as
\[ \mathcal{P}(S) = \{X \mid X \subseteq S\}. \]
It is the set of all subsets of \(S\).
REFERENCES