Notes on Real Analysis

Lee Larson

1 Sets and Functions

1.1 Basic Ideas

Set theory is a large and complicated subject in its own right. There is no time in this course to touch any but the simplest parts of it. Instead, we'll just look at a few topics from what is often called "naive set theory."

We begin with a few definitions.

A *set* is a collection of objects called *elements.* Usually, sets are denoted by the capital letters $A, B, \dots, Z$. A set can consist of any type of elements. Even other sets can be elements of a set. The sets we typically deal with here have real numbers as their elements.

If a is an element of the set A, we write $a \in A$. If a is not an element of the set A, we write $a \notin A$.

If all the elements of A are also elements of B, then A is a *subset* of B. In this case, we write $A \subset B$ or $B \supset A$.

Two sets A and B are *equal*, if they have the same elements. In this case we write $A = B$. It is easy to see that $A = B$ iff $A \subset B$ and $B \subset A$. Establishing that both of these containments are true is a standard way to show that two sets are equal.

There are several ways to describe a set.

A set can be described in words such as "P is the set of all presidents of the United States." This is cumbersome for complicated sets.

All the elements of the set could be listed in curly braces as $S = \{2, 0, a\}$. If the set is large, this is impractical, or impossible.

More common in mathematics is set builder notation. Some examples are

$$\begin{aligned} P &= \{p : p \text{ is a president of the United states}\} \\ &= \{\text{Washington, Adams, Jefferson,} \dots, \text{Clinton}\} \end{aligned}$$

and

$$S = \{n : n \text{ is a prime number}\} = \{2, 3, 5, 7, 11, \dots\}.$$

In general, the set builder notation defines a set in the form

$$\{\text{formula for a typical element} : \text{object to plug into the formula}\}.$$

A more complicated example is the set of perfect squares:

$$S = \{n^2 : n \text{ is an integer}\} = \{0, 1, 4, 9, \dots\}.$$

The existence of several sets will be assumed. The simplest of these is the *empty set*, which is the set with no elements. It is denoted as $\emptyset$. The *natural numbers* is the set $\mathbb{N} = \{1, 2, 3, \dots\}$ consisting of the positive integers. The set $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ is the set of all integers. Clearly, $\emptyset \subset A$, for any set A and

$$\emptyset \subset \mathbb{N} \subset \mathbb{Z}.$$

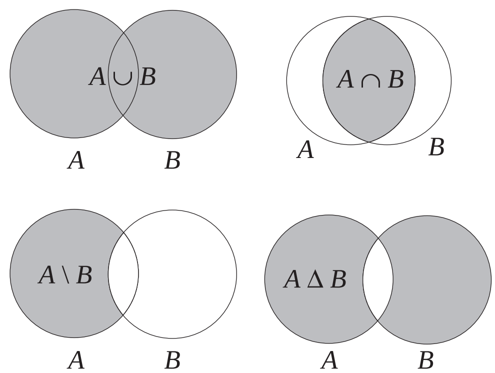

Figure 1: These are Venn diagrams showing the four standard binary operations on sets. In this figure, the set which results from the operation is shaded.

Definition 1.1. Given any set A, the *power set* of A, written $\mathcal{P}(A)$, is the set consisting of all subsets of A; i. e.,

$$\mathcal{P}(A) = \{B : B \subset A\}.$$

Problem 1. If a set S has n elements for $n \in \mathbb{Z}$ and $n \geq 0$, how many elements are in $\mathcal{P}(S)$?

1.2 Algebra of Sets

Let A and B be sets. There are four common binary operations used on sets.[1]

The *union* of A and B is the set

$$A \cup B = \{x : x \in A \vee x \in B\}.$$

The *intersection* of A and B is the set

$$A \cap B = \{x : x \in A \wedge x \in B\}.$$

The *difference* of A and B is the set

$$A \setminus B = \{x : x \in A \wedge x \notin B\}.$$

The *symmetric difference* of A and B is the set

$$A \Delta B = (A \cup B) \setminus (A \cap B).$$

[1] In the following, some standard logical notation is used. The symbol $\vee$ is the logical nonexclusive "or." The symbol $\wedge$ is the logical "and." Their truth tables are as follows:

$\wedge$	T	F
T	T	F
F	F	F

$\vee$	T	F
T	T	T
F	T	F

Another common set operation is *complementation*. The complement of a set A is usually thought of as the set consisting of all elements which are not in A. But, a short reflection will convince the reader that this is not a well-stated definition because the collection of elements not in A is not a precisely understood collection. To make sense of the complement of a set, there must be a well-defined *universal set* U which contains all the sets in question. Then the *complement* of a set $A \subset U$ is $A^c = U \setminus A$.

Theorem 1.1. *Let A, B and C be sets.*
(a) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
(b) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Proof. (a) This is proved as a sequence of equivalences.

$$\begin{aligned} x \in A \setminus (B \cup C) &\iff x \in A \wedge x \notin (B \cup C) \\ &\iff x \in A \wedge x \notin B \wedge x \notin C \\ &\iff (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \notin C) \\ &\iff x \in (A \setminus B) \cap (A \setminus C) \end{aligned}$$

(b) This is also proved as a sequence of equivalences.

$$\begin{aligned} x \in A \setminus (B \cap C) &\iff x \in A \wedge x \notin (B \cap C) \\ &\iff x \in A \wedge (x \notin B \vee x \notin C) \\ &\iff (x \in A \wedge x \notin B) \vee (x \in A \wedge x \notin C) \\ &\iff x \in (A \setminus B) \cup (A \setminus C) \end{aligned}$$

□

Theorem 1.1 is a version of a group of set equations which are often called DeMorgan's Laws. The more usual statement of DeMorgan's Laws are in Corollary 1.2. Corollary 1.2 is an obvious consequence of Theorem 1.1 when there is a universal set to make the complementation well-defined.

Corollary 1.2 (DeMorgan's Laws). *Let A and B be sets.*

(a) $(A \cup B)^c = A^c \cup B^c$

(b) $(A \cap B)^c = A^c \cup B^c$

Problem 2. Prove that for any sets A and B,

(a) $A = (A \cap B) \cup (A \setminus B)$

(b) $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$ and that the sets $A \setminus B$, $B \setminus A$ and $A \cap B$ are pairwise disjoint.

We often have occasion to work with large collections of sets. For example, we could have a sequence of sets $A_1, A_2, A_3, \dots$, where there is a set A_n associated with each $n \in \mathbb{N}$. In general, let Λ be a set and suppose for each $\lambda \in \Lambda$ there is a set A_λ. The collection $\{A_\lambda : \lambda \in \Lambda\}$ is called a *collection of sets indexed by* Λ. In this case, Λ is called the *indexing set* for the collection.

Example 1.1. For each $n \in \mathbb{N}$, let $A_n = \{k \in \mathbb{Z} : k^2 \le n\}$. Then

$$A_1 = A_2 = A_3 = \{-1, 0, 1\},\ A_4 = \{-2, -1, 0, 1, 2\},\ \dots,\\ A_{50} = \{-7, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, 7\}, \dots$$

is a collection of sets indexed by $\mathbb{N}$.

Several of the binary operations can be extended to work with indexed collections. In particular, using the indexed collection from the previous paragraph, we define

$$\bigcup_{\lambda \in \Lambda} A_\lambda = \{x : x \in A_\lambda \text{ for some } \lambda \in \Lambda\}$$

and

$$\bigcap_{\lambda \in \Lambda} A_\lambda = \{x : x \in A_\lambda \text{ for all } \lambda \in \Lambda\}.$$

DeMorgan's Laws can be generalized to indexed collections.

Theorem 1.3. *If $\{B_\lambda : \lambda \in \Lambda\}$ is an indexed collection of sets and A is a set, then*

$$A \setminus \bigcup_{\lambda \in \Lambda} B_\lambda = \bigcap_{\lambda \in \Lambda} (A \setminus B_\lambda)$$

and

$$A \setminus \bigcap_{\lambda \in \Lambda} B_\lambda = \bigcup_{\lambda \in \Lambda} (A \setminus B_\lambda).$$

Problem 3. Prove Theorem 1.3.

1.3 Functions and Relations

When listing the elements of a set, the order in which they are listed is unimportant; e. g., $\{a, b\} = \{b, a\}$. If the order in which n items are listed is important, the list is called an *n-tuple*. (Strictly speaking, an n-tuple is not a set.) We denote an n-tuple by enclosing the ordered list in parentheses. For example, if x_1, x_2, x_3, x_4 are 4 items, the 4-tuple (x_1, x_2, x_3, x_4) is different from the n-tuple (x_2, x_1, x_3, x_4).

Because they are used so often, a 2-tuple is called an *ordered pair* and a 3-tuple is called an *ordered triple.*

Definition 1.2. The *Cartesian product* of A and B is the set of all ordered pairs

$$A \times B = \{(a, b) : a \in A \wedge b \in B\}.$$

Example 1.2. If $A = \{a, b, c\}$ and $B = \{1, 2\}$, then

$$A \times B = \{(a, 1), (a, 2), (b, 1), (b, 2), (c, 1), (c, 2)\}.$$

A useful way to visualize the Cartesian product of two sets is as a table. The Cartesian product from Example 1.2 is contained in the entries of the following table.

	1	2
a	$(a, 1)$	$(a, 2)$
b	$(b, 1)$	$(b, 2)$
c	$(c, 1)$	$(c, 2)$

Of course, the common Cartesian plane from analytic geometry is nothing more than a variation of this idea of listing the elements of a Cartesian product as a table.

By induction, the definition of Cartesian product can be extended to the case of more than two sets. If $\{A_1, A_2, \dots, A_n\}$ are sets, then

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \dots, a_n) : a_k \in A_k \text{ for } 1 \leq k \leq n\}$$

is a set of n-tuples.

Definition 1.3. If A and B are sets, then any $R \subset A \times B$ is a *relation from A to B*. If $(a, b) \in R$, we write aRb.

In this case,

$$\text{dom}\,(R) = \{a : (a, b) \in R\}$$

is the *domain* of R and

$$\text{ran}\,(R) = \{b : (a, b) \in R\}$$

is the *range* of R.

Suppose $R \subset A \times A$.

The relation R is *symmetric*, if $aRb \iff bRa$.

The relation R is *reflexive*, if aRa whenever $a \in A$.

The relation R is *transitive*, if $aRb \wedge bRc \implies aRc$.

The relation R is an *equivalence relation* on A, if it is symmetric, reflexive and transitive.

Example 1.3. Let R be the relation on $\mathbb{Z} \times \mathbb{Z}$ defined by $aRb \iff a \leq b$. Then R is reflexive and transitive, but not symmetric.

Example 1.4. Let R be the relation on $\mathbb{Z} \times \mathbb{Z}$ defined by $aRb \iff a^2 = b^2$. In this case, R is an equivalence relation. It is evident that aRb iff $b = a$ or $b = -a$.

Problem 4. Suppose R is an equivalence relation on A. For each $x \in A$ define $C_x = \{y \in A : xRy\}$. Prove that if $x, y \in A$, then either $C_x = C_y$ or $C_x \cap C_y = \emptyset$. (The collection $\{C_x : x \in A\}$ is the set of *equivalence classes* induced by R.)

Definition 1.4. A relation $R \subset A \times B$ is a *function* if $aRb_1 \wedge aRb_2 \implies b_1 = b_2$.

If $f \subset A \times B$ is a function and $\text{dom}(f) = A$, then we usually write $f : A \to B$ and $f(a) = b$ instead of afb.

If $f : A \to B$ is a function, then the usual intuitive interpretation is to regard f as a rule that assigns each element of A to a unique element of B. It's not necessarily the case that each element of B is assigned something from A.

Example 1.5. Define $f : \mathbb{N} \to \mathbb{Z}$ by $f(n) = n^2$ and $g : \mathbb{Z} \to \mathbb{Z}$ by $g(n) = n^2$. In this case $\text{ran}(f) = \{n^2 : n \in \mathbb{N}\}$ and $\text{ran}(g) = \text{ran}(f) \cup \{0\}$.

Definition 1.5. If $f : A \to B$ and $g : B \to C$, then the *composition* of g with f is the function $g \circ f : A \to C$ defined by $g \circ f(a) = g(f(a))$.

In Example 1.5, $g \circ f(n) = g(f(n)) = g(n^2) = (n^2)^2 = n^4$ makes sense for all $n \in \mathbb{N}$, but $f \circ g$ is undefined at $n = 0$.

There are several important types of functions.

Definition 1.6. A function $f : A \to B$ is a *constant function*, if $\text{ran}(f)$ has a single element; i. e., there is a $b \in B$ such that $f(a) = b$ for all $a \in A$.

Definition 1.7. A function $f : A \to B$ is *surjective (or onto B)*, if $\text{ran}(f) = B$.

In a sense, constant and surjective functions are the opposite extremes. A constant function has the smallest possible range and a surjective function has the largest possible range. Of course, a function $f : A \to B$ can be both constant and surjective, if B has only one element.

Definition 1.8. A function $f : A \to B$ is *injective (or one-to-one)*, if $f(a) = f(b)$ implies $a = b$.

The terminology "one-to-one" is very descriptive in this case. An illustration of this definition is in Figure 2. In Example 1.5, f is injective while g is not.

Definition 1.9. A function $f : A \to B$ is *bijective*, if it is both surjective and injective.

A bijective function can be visualized as pairing up all the elements of A and B. In a sense, A and B must have the same number of elements for this to happen. This idea will be explored further in the next section.

Definition 1.10. If $f : A \to B$, $C \subset A$ and $D \subset B$, then the *image* of C is the set $f(C) = \{f(a) : a \in C\}$. The *inverse image* of D is the set $f^{-1}(D) = \{a : f(a) \in D\}$.

Definitions 1.9 and 1.10 work together in the following way. Suppose $f : A \to B$ is bijective and $b \in B$. The fact that f is surjective guarantees that $f^{-1}(b) \neq \emptyset$. Since f is injective, $f^{-1}(b)$ contains exactly one element, say a, where $f(a) = b$. In this way, it is seen that f^{-1} is a rule that assigns each element of B to exactly one element of A; i. e., f^{-1} is a function with domain B and range A.

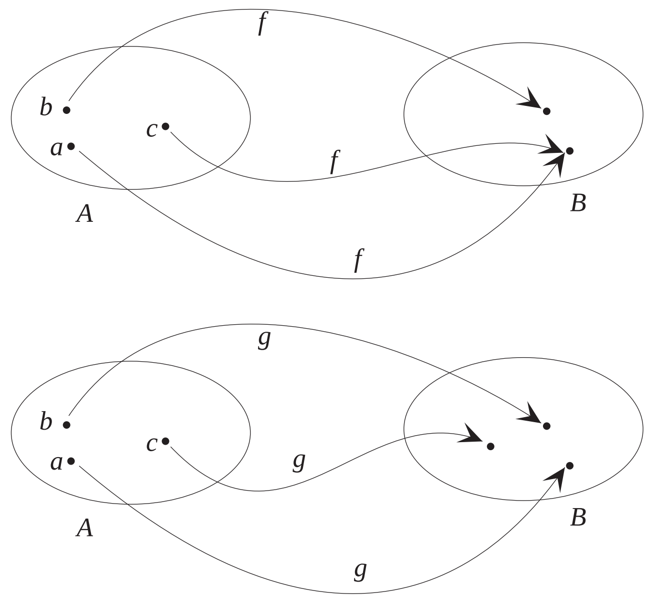

Figure 2: These diagrams show two functions, $f : A \to B$ and $g : A \to B$. The function g is injective and f is not because $f(a) = f(c)$.

Definition 1.11. If $f : A \to B$ is bijective, we define the *inverse* of f to be a function $f^{-1}B \to A$ with the property that $f^{-1} \circ f(a) = a$ for all $a \in A$ and $f \circ f^{-1}(b) = b$ for all $b \in B$.

Example 1.6. Let $A = \mathbb{N}$ and B be the even natural numbers. If $f : A \to B$ is $f(n) = 2n$ and $g : B \to A$ is $g(n) = n/2$, it is clear f is bijective. Since $f \circ g(n) = f(n/2) = 2n/2 = n$ and $g \circ f(n) = g(2n) = 2n/2 = n$, we see $g = f^{-1}$.

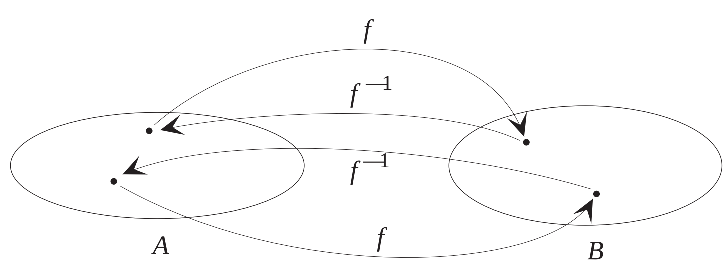

Figure 3: This is one way to visualize a general invertible function. First f does something to a and then f^{-1} undoes it.

Example 1.7. Let $f : \mathbb{N} \to \mathbb{Z}$ be defined by

$$f(n) = \begin{cases} (n-1)/2, & n \text{ odd}, \\ -n/2, & n \text{ even} \end{cases}$$

It's quite easy to see that f is bijective and

$$f^{-1}(n) = \begin{cases} 2n+1, & n \geq 0, \\ -2n, & n < 0 \end{cases}$$

The following theorem will be used in Section 1.4.

Theorem 1.4 (Schröder-Bernstein). *Let A and B be sets. If there are injective functions $f : A \to B$ and $g : B \to A$, then there is a bijective function $h : A \to B$.*

Proof. Let $B_1 = B \setminus f(A)$. If $B_k \subset B$ is defined for some $k \in \mathbb{N}$, let $A_k = g(B_k)$ and $B_{k+1} = f(A_k)$. This inductively defines A_k and B_k for all $k \in \mathbb{N}$. Use these sets to define $\tilde{A} = \bigcup_{k \in \mathbb{N}} A_k$ and $h : A \to B$ as

$$h(x) = \begin{cases} g^{-1}(x), & x \in \tilde{A} \\ f(x), & x \in A \setminus \tilde{A} \end{cases}.$$

It must be shown that h is well-defined, injective and surjective.

To show h is well-defined, let $x \in A$. If $x \in A \setminus \tilde{A}$, then it is clear $h(x) = f(x)$ is defined. On the other hand, if $x \in \tilde{A}$, then $x \in A_k$ for some k. Since $x \in A_k = g(B_k)$, we see $h(x) = g^{-1}(x)$ is defined. Therefore, h is well-defined.

To show h is injective, let $x, y \in A$ with $x \neq y$. If both $x, y \in \tilde{A}$ or $x, y \in A \setminus \tilde{A}$, then the assumptions that g and f are injective, respectively, imply $h(x) \neq h(y)$. The remaining case is when $x \in \tilde{A}$ and $y \in A \setminus \tilde{A}$. Suppose $x \in A_k$ and $h(x) = h(y)$. Then there is an $x_1 \in B_1$ such that

$$x = \underbrace{g \circ f \circ g \circ f \circ \cdots \circ f \circ g}_{k-1 \ f\text{'s and } k \ g\text{'s}}(x_1).$$

This implies

$$h(x) = g^{-1}(x) = \underbrace{f \circ g \circ f \circ \cdots \circ f \circ g}_{k-1 \ f\text{'s and } k-1 \ g\text{'s}}(x_1) = f(y)$$

so that

$$y = \underbrace{g \circ f \circ g \circ f \circ \cdots \circ f \circ g}_{k-2 \ f\text{'s and } k-1 \ g\text{'s}}(x_1) \in A_{k-1} \subset \tilde{A}.$$

This contradiction shows that $h(x) \neq h(y)$. We conclude h is injective.

To show h is surjective, let $y \in B$. If $y \in B_k$ for some k, then $h(A_k) = g^{-1}(A_k) = B_k$ shows $y \in h(A)$. If $y \notin B_k$ for any k, $y \in f(A)$ because $B_1 = B \setminus f(A)$, and $g(y) \notin \tilde{A}$, so $y = h(x) = f(x)$ for some $x \in A$. This shows h is surjective. □

Problem 5. If $f : A \to B$ is bijective, then f^{-1} is unique.

1.4 Cardinality

There is a way to use sets to formalize and generalize how we count. For example, suppose we want to count how many elements are in the set $\{a, b, c\}$. The natural way to do this would be to point at each element in succession and say "one, two, three." What is really happening is that we're defining a bijective function between $\{a, b, c\}$ and the set $\{1, 2, 3\}$. This idea can be generalized.

Definition 1.12. Given $n \in \mathbb{N}$, an *initial segment* of $\mathbb{N}$ is the set $\overline{n} = \{1, 2, \ldots, n\}$. The trivial initial segment is $\overline{0} = \emptyset$. A set S has *cardinality* n, if there is a bijective function $f : S \to \overline{n}$. In this case, we write $\operatorname{card}(S) = n$.

The cardinalities defined in Definition 1.12 are called the *finite* cardinal numbers. They correspond to the everyday counting numbers we usually use. The idea can be generalized still further.

Definition 1.13. Let A and B be two sets. If there is an injective function $f : A \to B$, we say $\operatorname{card}(A) \le \operatorname{card}(B)$.

According to Theorem 1.4, the Schröder-Bernstein Theorem, if $\operatorname{card}(A) \le \operatorname{card}(B)$ and $\operatorname{card}(B) \le \operatorname{card}(A)$, then there is a bijective function $f : A \to B$. As expected, in this case we write $\operatorname{card}(A) = \operatorname{card}(B)$. When $\operatorname{card}(A) \le \operatorname{card}(B)$, but no such bijection exists, we write $\operatorname{card}(A) < \operatorname{card}(B)$.

In particular, a set A is *countably infinite*, if $\operatorname{card}(A) = \operatorname{card}(\mathbb{N})$. In this case, it is common to write $\operatorname{card}(\mathbb{N}) = \aleph_0$.[2]

This leaves open the question whether all sets either have finite cardinality, or are countably infinite. This is answered by letting $S = \mathbb{N}$ in the following theorem.

Theorem 1.5. *If S is a set,* $\operatorname{card}(S) < \operatorname{card}(\mathcal{P}(S))$.

Proof. It is easy to see $\operatorname{card}(S) \le \operatorname{card}(\mathcal{P}(S))$, so it suffices to prove there is no surjective function $f : S \to \mathcal{P}(S)$.

To see this, assume there is such a function f and let $T = \{x \in S : x \notin f(x)\}$. Since f is surjective, there is a $t \in T$ such that $f(t) = T$. Either $t \in T$ or $t \notin T$.

If $t \in T = f(T)$, then the definition of T implies $t \notin T$, a contradiction. On the other hand, if $t \notin T = f(T)$, then the definition of T implies $t \in T$. These contradictions lead to the conclusion that no such function f can exist. $\square$

A set S is said to be *uncountably infinite*, or just *uncountable*, if $\aleph_0 < \operatorname{card}(S)$. Theorem 1.5 implies $\aleph_0 < \operatorname{card}(\mathcal{P}(\mathbb{N}))$, so $\mathcal{P}(\mathbb{N})$ is uncountable. In fact, the same argument implies

$$\aleph_0 = \operatorname{card}(\mathbb{N}) < \operatorname{card}(\mathcal{P}(\mathbb{N})) < \operatorname{card}(\mathcal{P}(\mathcal{P}(\mathbb{N}))) < \ldots$$

So, there are an infinite number of distinct infinite cardinalities.

[2] The symbol $\aleph$ is the Hebrew letter "aleph" and $\aleph_0$ is usually pronounced "aleph nought."

Extra Credit 1. Prove that if a set S is countably or uncountably infinite, then there is a proper subset $T \subsetneq S$ and a bijection $f : S \to T$. This property is often used as the definition of when a set is infinite.

Notice that Theorem 1.5 does not imply there are no sets B such that $\aleph_0 < \operatorname{card}(B) < \operatorname{card}(\mathcal{P}(\mathbb{N}))$. In fact, for many years the question of whether such sets exist was one of the most important open questions in mathematics. The assumption that no such sets exist is called the *continuum hypothesis.*

The continuum hypothesis was first stated as a conjecture by Georg Cantor in 1878. Kurt Gödel proved in 1938 that the continuum hypothesis does not contradict anything in normal set theory, but he did not prove it was true. In 1963 it was proved by Paul Cohen that the continuum hypothesis is actually unprovable as a theorem in standard set theory.

So, the continuum hypothesis is a statement which is neither true nor false within the framework of ordinary set theory. This means that in an axiomatic development of set theory, the continuum hypothesis, or a suitable negation of it, can be taken as an axiom.

The proofs of these theorems are quite complicated. A well-written introduction to to many of these ideas is contained in the book by Ciesielski [1].

Problem 6. Suppose that A_k is a set for each positive integer k.

(a) Show that $x \in \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k\right)$ iff $x \in A_k$ for infinitely many sets A_k.

(b) Show that $x \in \bigcup_{n=1}^{\infty} \left(\bigcap_{k=n}^{\infty} A_k\right)$ iff $x \in A_k$ for all but finitely many of the sets A_k.

The set $x \in \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k\right)$ from (a) is often called the *superior limit* of the sets A_k and $x \in \bigcup_{n=1}^{\infty} \left(\bigcap_{k=n}^{\infty} A_k\right)$ is often called the *inferior limit* of the sets A_k.

Problem 7. Given two sets A and B, it is common to let A^B be the set of all functions $f : B \to A$. Prove that for any set A, $\operatorname{card}\left(\overline{2}^A\right) = \operatorname{card}(\mathcal{P}(A))$.