Contents

Chapter 1 Introduction
Chapter 2 The Language and Literature of Mathematics
Chapter 3 Basic Set Theory
Chapter 4 The Real Numbers
Chapter 5 Infinite Sequences
Chapter 6 Continuous Functions of a Real Variable
Chapter 7 Derivatives
Chapter 8 Area and Integrals
Chapter 1

Introduction

1.1 Purpose of the Course

Most of your university courses focus on new material for you to learn, and test you on how well you have learned and understood it. Your Mathematics courses have emphasized learning algorithms to construct derivatives and integrals, to multiply and invert matrices, and to solve linear equations.

Your objective in this course is different: it is to solve theoretical problems in mathematics and to use rigorous mathematical language to prove your solutions are correct.

For this, you must

- Fully understand the problem, and the mathematical terms used. In particular, make sure you understand
  1. what you are given (the hypotheses);
  2. what you are asked to prove (the conclusion);
  3. the terminology and the notation;
- Figure out why the statement to prove is true.
- Express your idea in a formal mathematical proof.

You will learn techniques you can use to solve problems, but there are no algorithms for this - each problem requires its own idea for a solution.

Definition 1: A mathematical proof is a sequence of mathematical statements which together demonstrate the truth of an assertion.

1. Each assertion must follow logically from the preceding ones and the hypotheses, and must explain why!
2. The last statement (the conclusion) states that the assertion is true.

**Definition 2** A mathematical statement either defines notation or terminology, or else it is a clear, unambiguous assertion, usually using mathematical terms.

Here are some easy examples of mathematical proofs:

**Example 1** 1. Prove that if a student is in Math 310 (the hypothesis), s/he is registered at the University of Maryland (the conclusion).

**Proof:** To be in Math 310 you must be registered at the University. Therefore, since the student is in Math 310, s/he is registered at the University. **q.e.d.**

2. Prove that if \( x > 1 \) then \( x^2 > x \).

**Proof:** \( x^2 - x = x(x - 1) \).
Since \( x > 1 \), \( x \) and \( x - 1 \) are both positive.
Therefore \( x(x - 1) \) is positive.
Therefore \( x^2 > x \). **q.e.d.**

3. If \( n \) is a natural number and \( n > 1 \), prove that \( n^3 - 1 \) is divisible by \( n - 1 \).

**Proof:** Direct multiplication gives \( (n - 1)(n^2 + n + 1) = n^3 - 1 \).
Since \( n^2 + n + 1 \) is a natural number, \( n^3 - 1 \) is divisible by \( n - 1 \).
**q.e.d.**

4. Prove that if every student in Math 310 gets at least \( 7/10 \) on a homework then the class average is at least \( 7/10 \).

**Proof:** Let the number of students in the class be \( n \) (**Establishes notation**).
Let \( x_i \) be the grade of the \( i \)th student.
Then, by definition, the class average is
\[
\frac{\sum_{i=1}^{n} x_i}{n}.
\]

Since each \( x_i \geq 7 \), therefore
\[
\frac{\sum_{i=1}^{n} x_i}{n} - 7 = \frac{\sum_{i=1}^{n} x_i}{n} - \frac{7n}{n} = \frac{\sum_{i=1}^{n} (x_i - 7)}{n} \geq 0.
\]
Therefore the class average is at least \( 7/10 \).

5. **The Division Proposition:** If \( p < n \) are natural numbers then for some natural number \( m < n \) and for some integer \( r \) with \( 0 \leq r < p \),
\[
n = mp + r.
\]
We call $r$ the remainder.

**Proof:** If $k > n$ then for any $r \geq 0$

$$kp \geq k > n.$$  

Thus if $mp + r = n$ it must be true that $mp \leq n$ and so $m \leq n$. 

Since $1 \leq n$ we can choose $m$ to be the largest natural number such that $mp \leq n$.

Then

$$n - mp \geq 0.$$  

On the other hand, if $n - mp \geq p$ then $(n - mp) - p > 0$ and so

$$n - (m + 1)p = (n - mp) - p > 0.$$  

Thus $m$ was not the largest natural number which multiplied by $p$ gave an answer at most $n$, and we assumed it was. 

Thus $n - mp < p$ and we may set $r = n - mp$. q.e.d.

Everyone, even at an early age, learns what it means for a statement to follow logically from others, and sometimes it will be obvious that the statement is correct. So, writing proofs and solving problems does not require you to learn to think logically, since you already know how to do that.

However, this course is about Mathematics. Most of the problems are about mathematical objects, and so you must get to know and understand these objects well and you must acquire the ability to read and write in mathematical language, since that is the language in which the problems are posed and in which your proofs must be expressed.

**Tips for Using Mathematical Language:**

1. Mathematical assertions are always either true or false. There is no 'middle ground'.

2. Mathematical language is about mathematical objects such as sets, numbers, functions, ...).

3. When writing mathematics it is almost always useful to label the objects under consideration. This is called "establishing notation".

4. In Mathematics words and symbols must be carefully and precisely defined, with no ambiguity.

5. **When you write a mathematical statement it must say exactly what you mean:**

   (a) Even if you feel that the reader will understand what you wrote although you have not said it exactly, in this course that is not acceptable.
(b) Consequently you will need to always write in clear, unambiguous complete sentences, which say precisely what you mean.

Homework requirements

1. Your answers must be a sequence of mathematical statements as described above, and you will get a low grade if they do not meet those requirements, even if they give the impression that you understand the material.

2. In particular your answers must be:
   (a) typed or written legibly in ink;
   (b) clear, precise, and unambiguous;
   (c) without excess verbiage;
   (d) written in complete, grammatical sentences.

3. Each statement must follow logically from the previous ones, with an explanation as to why.

4. Your homework answers may rely on facts as outlined below, but you must in each case indicate which fact you are using and give the explicit reference:
   (a) any statement already proved in class;
   (b) any statement proved in an earlier part of the text;
   (c) any statement in any earlier exercise.

Tips for Writing Proofs from Elizabeth (a former Math 310 student):

1. Learning to write proofs is not like learning how to multiply matrices, and it can take mental effort and time.

2. Remain calm; if you put enough time into this class, you will figure out how to write proofs.

3. Give yourself plenty of time to work on a proof.

4. Before you start the proof, look back at the given definitions and lemmas that you think might be useful for this proof. Make sure you fully understand the definitions and lemmas and see if you can find a way to connect them back to what you are trying to prove.

5. Restate definitions in your proofs, so you are less likely to use a definition incorrectly.

6. Working in groups is fine, but make sure you try each exercise thoroughly on your own first. You will not have your group to help when it comes time to take the test.
7. Never hesitate to ask Professor Halperin a question when you are stuck: either in class, in his office, or by email. He will welcome your interest: I know!

8. Most importantly: This class is designed to help you learn how to write proofs. While it is important you understand the concepts, you must be able to use what you learn to complete a correctly written proof in order to succeed.

**In Summary**, to **succeed** in this course you will need to:

1. **Absolutely understand the definitions** of the words and symbols you run into.

2. Be **totally comfortable** with the meaning and properties of the objects the words refer to.

3. **Understand** and be able to explain to others the precise meaning of each mathematical statement you encounter.

4. Write so that each sentence has a **single** meaning which will be clear to any reader

5. Hand in each homework and **master** the assigned readings in the text.

6. Get help from me in class, by email, or in office hours, when you do not understand something.

7. Join a study group, but never as a freeloader. If you freeload your homework score may be OK, but your tests and exams will NOT be.

**Students who invest this time and effort usually do well. Students who do not make this effort usually do not.**
Chapter 2

The Language and Literature of Mathematics

2.1 The Language of Mathematics

Definitions and notation are essential tools in keeping our language unambiguous and terse and, as with any new language, it is essential that you internalize the definitions and notation rather than simply memorize them. Indeed, when learning to speak or write a new language you need to be able to use the words spontaneously without having to call up each corresponding English word and then translate it. In the same way, you need to be able to speak/write mathematics, not just remember the dictionary of definitions.

You learn a second language most easily by speaking it with others to whom it comes naturally. You learn to drive a car by driving it and to walk by walking. You learn to write/speak mathematics by writing it and presenting it and getting feedback when you get it right and how to correct it when you don’t. The golden rule when writing: never write anything whose meaning is unclear to yourself! You can also use this text to find many detailed examples of how to write a proof correctly.

The language of mathematics consists of assertions about mathematical objects. Mathematical objects include the natural numbers, the integers, the rationals, the real numbers, sets, maps, functions and many other things. Mathematical language has its own vocabulary: for example, two words which appear throughout mathematics are the words set and sequence:

Definition 3

1. A set is a specified collection of distinct objects, abstract or concrete, called its elements. An element $x$ in a set $S$ is said to belong to $S$, and
we denote this by \( x \in S \). We often write

\[ S = \{ x, y, z, \cdots \}, \]

where \( x, y, z, \cdots \) are the elements of \( S \).

2. A sequence starting at \( k \) is a list \( (x_n)_{n \geq k} \) of objects, possibly with repetitions, and indexed by the integers \( n, n \geq k \). If \( (x_n)_{n \geq k} \) is a sequence, then \( x_n \) is the \( n \)th term in the sequence.

Note: The objects listed in a sequence form a set, as seen in the following example.

Example 2

1. The list \( 0, 1, 0, , 1, 0, 1, \cdots \) is a sequence. The corresponding set is the collection \( \{0, 1\} \) which only has two elements.

2. The list \( 2, 4, 2, 4, 6, 2, 4, 6, 8, 2, 4, 6, 8, 10, \cdots \) is a sequence. The corresponding set is the collection of even natural numbers.

Mathematical language also relies heavily on symbols to express assertions and proofs in a short and clear manner, but in any specific proof each symbol must have a precise meaning and the same symbol may never be used with two different meanings. In particular, we fix the following symbols for the entire course.

Example 3 Basic notation

1. \( N \) will denote the set of natural numbers, and we write \( N = \{1, 2, 3, \ldots\} \).

2. \( Z \) will denote the set of integers: \( Z = \{0, \pm 1, \pm 2, \pm 3 \ldots\} \).

3. \( Q \) will denote the set of rational numbers: these are the real numbers which can be written \( p/q \) with \( p \in Z \) and \( q \in N \). Note that \( p/q = r/s \) if and only if \( ps = rq \).

4. \( R \) will denote the set of real numbers.

5. \( C \) will denote the set of complex numbers.

6. The absolute value of a real number \( x \) is denoted by \( |x| \) and is defined by:

\[
|x| = \begin{cases} 
  x, & x \geq 0, \\
  -x, & x < 0 
\end{cases}
\]

Note that \( |x| \geq 0 \) for all \( x \in R \).

7. q.e.d. stands for quid erat demonstrandum, which is Latin for "what was to be proved". It is used at the end of a proof to signal that the proof is complete.
Definition 4  
1. An integer \( n \) is **even** if it is divisible by 2.

2. An integer \( n \) is **odd** if it is not even.

3. A **prime number** is a natural number \( n \) which is larger than 1 and is not divisible by any natural number except 1 and \( n \).

Example 4 "Sigma" and "Pi" notation

1. The "**Sigma**" notation is used for sums:
   If \( v_i \) are numbers or vectors (or anything else we know how to add) then
   \[
   \sum_{i=1}^{n} v_i = v_1 + v_2 + \cdots + v_n.
   \]
   More generally,
   \[
   \sum_{i=1}^{n} v_{k_i} = v_{k_1} + v_{k_2} + \cdots + v_{k_n}.
   \]
   In this notation the "\( i \)" just indexes the terms being added or multiplied, and we could use any letter instead without changing the meaning. Thus
   \[
   \sum_{i=1}^{n} v_i = \sum_{i=1}^{n} v_i = \sum_{q=1}^{n} v_q = v_1 + v_2 + \cdots + v_n.
   \]
   In other words, the "\( i \)" is a dummy variable just as in calculus, where we have
   \[
   \int f(x)dx = \int f(u)du
   \]

2. The "**Pi**" notation is used for products:
   If \( v_i \) are numbers then
   \[
   \prod_{i=1}^{n} v_i = v_1 \cdot v_2 \cdots v_n.
   \]
   More generally,
   \[
   \prod_{i=1}^{n} v_{k_i} = v_{k_1} \cdot v_{k_2} \cdots v_{k_n}.
   \]

3. In particular, if \( C \) is a **constant**, then
   \[
   \sum_{i=1}^{n} C = C + \cdots + C \text{ (n times)} = nC
   \]
   and
   \[
   \prod_{i=1}^{n} C = C \cdots C = C^n.
   \]
Proposition 1  
(Difference theorem) For any real numbers \( a, b \in \mathbb{R} \) and any \( n \in \mathbb{N} \),
\[
b^n - a^n = (b - a) \sum_{i=0}^{n-1} b^i a^{n-1-i}.
\]

Proof: Note that
\[
b \left( \sum_{i=0}^{n-1} b^i a^{n-1-i} \right) = \sum_{i=0}^{n-1} b^{i+1} a^{n-1-i} = \sum_{j=0}^{n-1} b^{j+1} a^{n-1-j}.
\]
In the second expression set \( j + 1 = i \). Then
\[
\sum_{j=0}^{n-1} b^{j+1} a^{n-1-j} = \sum_{i=1}^{n} b^i a^{n-i},
\]
and so
\[
b \left( \sum_{i=0}^{n-1} b^i a^{n-1-i} \right) = \sum_{i=1}^{n} b^i a^{n-i} = b^n + \sum_{i=1}^{n-1} b^i a^{n-i}.
\]
On the other hand,
\[
(-a) \left( \sum_{i=0}^{n-1} b^i a^{n-1-i} \right) = - \sum_{i=0}^{n-1} b^i a^{n-i} = -a^n - \sum_{i=1}^{n-1} b^i a^{n-i}.
\]
Adding these two lines gives the equation of the Proposition.
q.e.d.

Example 5  
Words and expressions which frequently appear in mathematical statements.

1. A positive real number means a number \( x \) such that \( x > 0 \). Thus zero is not a positive number!

2. A negative real number means a number \( x \) such that \( x < 0 \). Thus zero is not a negative number!

3. Let....: This is a statement which defines some terminology or establishes notation. As an example: Let "\( n \)" be a natural number greater than 5.

4. Theorem, proposition, lemma: These are names for mathematical statements that are going to be proved.
2.2 Logical statements

In Mathematics, you will frequently encounter statements about statements! Moreover, just as we use symbols to represent unspecified numbers, functions, and matrices, so also we may use symbols to represent statements!

Example 6 Here are four statements:

1. Statement A: John and Mary are students in Math 310.
2. Statement B: John and Mary are registered as students in UMD.
3. A implies B.
4. B implies A.

In this example, the statement A implies B is true, and the statement B implies A is false.

Here are some frequently used logical statements and expressions:

1. "If A then B ." This assertion states that if A is true then B is true. Here A is the hypothesis or assumption (both words mean the same thing), and B is the conclusion. This is often denoted by

   \[ A \Rightarrow B. \]

   It may also be phrased as

   (a) A implies B, or as
   (b) B if A.

2. A only if B. This means that A is not true unless B is true. In other words, if B is not true then A is not true.

3. "A if and only if B" This means that if B is true then so is A, and that if B is not true then A is not true. In other words, A is true if and only B is true. This is often denoted by

   \[ A \leftrightarrow B. \]

4. "For each \( x \in S \) there exists a \( y \) such that A is true" means that for each choice of \( x \) there exists a \( y \) (\( y \) will usually depend on \( x \)) with the properties prescribed by A.

5. The "converse" to the statement "If A is true then B is true" is the statement "if B is true then A is true".

6. The "contrapositive" to the statement "A is true if B is true" is the statement "B is not true if A is not true".
The next Lemma shows that a statement and its contrapositive are equivalent! The proof relies on the fact that in mathematics, every statement is either true or false.

**Lemma 1**

1. The statement "$A$ implies $B$" is true if and only if its contrapositive is true.

2. 

$"A \Rightarrow B \text{ and } B \Rightarrow A" \iff "A \iff B".$

**Proof:**

1. Suppose that $A$ implies $B$. Then if $B$ is not true, $A$ cannot be true! Thus if the statement $A$ implies $B$ is true, so is its contrapositive. On the other hand, suppose the contrapositive is true. Then if $A$ is true it cannot be that $B$ is not true. Thus $B$ is true and so $A$ implies $B$.

2. By definition "$A$ implies $B$" and "$B$ implies $A$" together are the same as "$A$ is true if and only if $B$ is true".

q.e.d.

### 2.3 Examples of mathematical writing

1. **Statement:** Let $(x_n)$ be the sequence of rational numbers defined by $x_n = 1 - 1/n$. Then for each $k \in \mathbb{N}$ there is some natural number $N$ such that $|1 - x_n| < 1/2k$ if $n \geq N$.

   In this example the natural number $N$ depends on the choice of $k$! In fact,
   
   $$|1 - x_n| = |1 - (1 - 1/n)| = |1/n| = 1/n.$$

   Thus if $n \geq N$ then $|1 - x_n| = 1/n \leq 1/N.$

   Therefore if if $N = 3k$ then for $n \geq N, |1 - x_n| \leq 1/3k < 1/2k$.

2. **This example illustrates the importance of using words carefully and correctly.**

   Let $(x_n)$ be the sequence of integers defined by $x_n = (-1)^n$. Consider the following two assertions

   (a) **Statement:** There is an $\varepsilon \in \mathbb{R}$, and there is an $N \in \mathbb{N}$ such that $|x_n - 0| < \varepsilon$ for $n \geq N$.  

12
(b) **Statement:** For each $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that $|x_n - 0| < \varepsilon$ for $n \geq N$.

**The first is correct and the second is false.**

To see that the first is true pick $\varepsilon = 3$ and $N = 1$. Then for $n \geq N$,

$$|x_n - 0| = |(-1)^n - 0| = 1 < 3 = \varepsilon.$$

To see that the second assertion is false pick $\varepsilon = 1/2$. Then for all $n$,

$$|x_n - 0| = 1$$

which is not less than $1/2$.

3. **Statement:** For each real number $a$ and for each $\varepsilon > 0$ there is a $\delta > 0$ such that if $|x - a| < \delta$ then $|x^2 - a^2| < \varepsilon$.

In this example $\delta$ will depend on both $a$ and $\varepsilon$: a larger $a$ and a smaller $\varepsilon$ will require a smaller $\delta$.

4. **Statement:** If $x = 2$ and $y = 4$ then $x + y = 6$. This assertion is true.

However, the converse assertion: "if $x + y = 6$ then $x = 2$ and $y = 4$" is false, because if $x = 1$ and $y = 5$ then $x + y = 6$. (This is called a counterexample)

5. **Statement:** Suppose $x$ and $y$ are natural numbers. Then $xy = 2$ only if either $x \geq 2$ or $y \geq 2$.

**Proof:** The statement says that if $xy = 2$ then one of $x$ and $y$ must be at least 2.

But if neither $x$ nor $y$ is at least 2 then both must be 1.

In this case $xy = 1$, contrary to our hypothesis.

Thus one of $x, y$ must be at least 2. q.e.d.

6. By contrast, the statement $xy = 2$ if and only if one of $x \geq 2$ and $y \geq 2$ is false, since it means two things:

   (a) If $xy = 2$, then $x \geq 2$ or $y \geq 2$, AND
   (b) If $x \geq 2$ or $y \geq 2$ then $xy = 2$.

But if $x = y = 2$ then $xy = 4$ and so the second statement is not true.

7. This example highlights the importance of "getting the order right", with two similar statements:

   (a) For each student in this class there is a date on which that student was born.
(b) For each date there is a student in the class who was born on that date.

The first statement is true and the second is false.

8. **Statement:** A natural number $n$ is odd if and only if it has the form $n = 2k - 1$ for some natural number $k$.

**Proof:** If $n$ has the form $n = 2k - 1$ for some natural number $k$, then it is not divisible by 2. Therefore $n$ is odd.

On the other hand, any natural number, when divided by 2 has a remainder of either 0 or 1. If the remainder is 0 then the number is divisible by 2, and so it is even.

Thus if $n$ is odd, when divided by 2 it has a remainder of 1. In this case $n = 2m + 1$ where $m = 0$ or $m \in \mathbb{N}$. Set $k = m + 1$. Then $k \in \mathbb{N}$ and $n = 2k - 1$. q.e.d.

9. **Statement:** For every rational number there is a natural number which is larger.

**Proof:** The rational number must have the form $p/q$ in which $p \in \mathbb{Z}$ and $q \in \mathbb{N}$. (Establishes notation)

If $p \leq 0$ then $p/q < 1$.

Otherwise $p \in \mathbb{N}$ and $p/q \leq p < p + 1$. q.e.d.

10. **Statement:** For every positive rational number, $a$, there is a natural number $n$ such that $1/n < a$.

**Proof:** Write $a = p/q$ with $p, q \in \mathbb{N}$.

Then $1/(q + 1) < 1/q \leq p/q$. Set $n = q + 1$. q.e.d.

**Exercise 1** 1. For each statement below decide if it is true or false and either provide a proof that it is true or give an example to show that it is false.

(a) If $x \in \mathbb{Z}$ and $x \geq 0$ then $x \in \mathbb{N}$.

(b) If $p, q \in \mathbb{Z}$ and $p + q \in \mathbb{N}$ then $p \in \mathbb{N}$.

(c) If $x \in \mathbb{N}$, $y \in \mathbb{Z}$, and $z \in \mathbb{C}$, then $x, y, z \in \mathbb{R}$.

(d) If $p, q \in \mathbb{N}$, and $p \geq 2$ and $q \geq 6$, then $p + q$ is even.
2. Describe what is wrong with each of the following statements:

(a) Laptops are popular among college students. Therefore at least one student at Maryland has a laptop.

(b) It is true that $u = 6$ for every $x, y \in \mathbb{Q}$ such that $x + y < u$.

3. Does (a) or (b) correctly add "such that" to the statement: "every student has a height $h$"?

(a) Every student has a height $h$ such that $h \leq a$ where $a$ is the height of the tallest student.

(b) Every student has a height $h$ such that every student is in a class.

4. Is the following statement true or false? Provide a proof for your answer.

"Suppose $A$ and $B$ are natural numbers, and that $A = 5$ if and only if $B = 2$. If $B \neq 2$, then $A = 7$.

5. What is the contrapositive of the following statements?

(a) $A \neq 8$ if $B = 9$.

(b) $A = 8$ if $B \neq 9$.

6. What is the converse of the following statements?

(a) $B = 9$ if $A = 8$.

(b) $A \neq 8$ if $B = 9$.

7. What is the hypothesis and what is the conclusion in the following assertions?

(a) The cube of an odd natural number is odd.

(b) For every $\varepsilon > 0$ there is some $p \in \mathbb{N}$ such that $1/p \leq \varepsilon$.

8. What is the contrapositive of the statement: For every $\varepsilon > 0$ there is a point $(x, y)$ in the plane whose distance from the origin is between $\varepsilon/2$ and $\varepsilon$.

9. What is the converse of the following statement about natural numbers: $n, m$?

If $n > m$ then for some $k \in \mathbb{N}, km > n$.

Is the original statement true? Is the converse true? Provide proofs that show your answers are correct.
10. What is the converse of the statement "A implies B"? Construct an example where a statement is false but the converse is true.

11. What is the converse of the statement "A is true if and only if B is true."? What is the difference in meaning between the original statement and the converse?

12. Is (a) or (b) a proof that the following statement is wrong? **Statement:**
   All ducks are red.
   
   (a) There exists a yellow duck, A. Therefore not all ducks are red.
   
   (b) We have not seen all ducks. Therefore some may not be red.

13. Prove that if \( k \in \mathbb{N} \) and \( k^2 \) is even then \( k^2 \) is divisible by 4.

14. Suppose \( x, y \) are positive real numbers and \( n \in \mathbb{N} \).
   
   (a) Show that \( x < y \) if and only if \( x^n < y^n \).
   
   (b) If \( x < y \) show that if \( n > 1 \) then \( y^n - x^n < n(y - x)y^{n-1} \).

15. In each of the following statements, replace the blanks by words/expressions from the list:

   let, therefore, if, only if, only, since, therefore, some, then, by, even, implies, because, each, not, true, false, hypothesis, zero, one, theorem, follows

   so that the statements are true. Then supply a proof.

   (a) ___ \( n \) is a positive natural number ___ \( n + 1 \) is divisible ___ a prime number.
   
   (b) ___ ___ natural number divides a fixed integer then that integer is ___.
   
   (c) ___ \( S \) be a set with a single element. ___ there is ___ ___ sequence starting at ___ listing elements of \( S \).
   
   (d) The statements "A implies B" and "B is ___ true ___ A is ___ true" are equivalent.
   
   (e) ___ \( x < y \) be real numbers. ___ \( y < 0 \) ___ \( |y| < |x| \).

16. In each of the following, replace the blanks by words/expressions from the list in the previous problem to provide a proof of the given statement.

   (a) **Statement:** The sum of an even number of odd integers is even.

   **Proof:** ___ \( k \) be the number of integers and ___ \( n_i \) be the \( i^{th} \) integer.
   
   ___ \( k \) is even, \( k = 2m \), where \( m \) is an integer.
n_i is odd, n_i = 2m_i - 1, where m_i is an integer.

\[ \sum_{i=1}^{k} n_i = \sum_{i=1}^{2m} (2m_i - 1) = 2 \sum_{i=1}^{2m} (m_i) - \sum_{i=1}^{2m} 1 = 2 \sum_{i=1}^{2m} (m_i) - 2m. \]

\[ \sum_{i=1}^{k} n_i \text{ is } \frac{1}{2}. \]

q.e.d.

(b) **Statement**: If \( k \in \mathbb{N} \) and \( k^2 \) is divisible by 3, then \( k^2 \) is divisible by 9.

**Proof**: Let \( r \) be the remainder when \( k \) is divided by 3.
- \( k = 3p + r \) for \( p \in \mathbb{Z} \).
- \( k^2 = 9p^2 + 6pr + r^2 \).

Since by assumption, \( k^2 \) is divisible by 3, it follows that \( r^2 \) is divisible by 3. But \( r \) is one of 0, 1, 2 and 1 = 1^2 and 4 = 2^2 are not divisible by 3.
- \( r = 0 \) and \( k = 3p \) is divisible by 3.

q.e.d.

(c) **Statement**: For any natural numbers \( p \) and \( n \), \( (p + 1)^n - p^n \geq n \).

**Proof**: By the Difference Theorem,

\[ (p + 1)^n - p^n = (p + 1 - p) \sum_{i=0}^{n-1} (p + 1)^i p^{n-i-1} = \sum_{i=0}^{n-1} (p + 1)^i p^{n-i-1}. \]

\[ (p + 1)^i \geq 1 \quad \text{and} \quad p^{n-i-1} \geq 1, \]

it follows that

\[ \sum_{i=0}^{n-1} (p + 1)^i p^{n-i-1} \geq \sum_{i=0}^{n-1} 1 \geq n. \]

\[ (p + 1)^n - p^n \geq 1. \]

q.e.d.

### 2.4 Proof by Contradiction

This method to prove a statement, \( A \), is true uses the following steps:

1. Assume that \( A \) is false.
2. Deduce that statement, \( B \), known to be true must then also be false.
3. Since \( B \) is known to be true, conclude that \( A \) cannot be false.
4. Therefore \( A \) must be true.

**Example 7** Proof by contradiction
1. **Prove that there are infinitely many prime numbers.**

   **Proof:** Suppose the statement "there are infinitely many prime numbers" is false.
   Then there would be only finitely many prime numbers.
   Thus there would be a largest prime number $p$.
   Let $n$ be the number obtained by first multiplying all the natural numbers from 1 to $p$ and then adding 1. *(Establishes notation)*
   Then $n > p$.
   Dividing $n$ by any natural number $k \leq p$ gives a remainder 1.
   Therefore $n$ is not divisible by any natural number $k \leq p$. But $n$ must be divisible by some prime number $q$.
   Therefore there must be a prime number $q > p$.
   This contradicts the hypothesis that $p$ was the largest prime number.
   Therefore there must be infinitely many primes. *q.e.d.*

   **Remark:** This proof was discovered by Euclid about 300 BC.

2. **Prove that there is not a smallest positive rational number.**

   **Proof:** Suppose the statement "there is not a smallest positive rational number" is false.
   Then there is a smallest positive rational number.
   This must be of the form $p/q$, with $p, q \in \mathbb{N}$.
   But $p/(q + 1) < p/q$.
   Thus $p/q$ is not the smallest positive rational number, which is a contradiction.
   Therefore there is not a smallest positive rational number. *q.e.d.*

**Exercise 2** **Proof by contradiction**

1. Show that there is not a largest even natural number.

2. Show that there is not a largest rational number of the form $\frac{p}{p+1}$ with $p \in \mathbb{N}$.

3. Show that there is not a largest negative rational number.

4. Prove that 2 is not the square of a rational number.

2.5 **Induction**

We use induction both to prove statements and to make constructions. Proof by induction is based on the following induction principle, which we take as an
axiom. In other words, we assume it is true without proof!

**Induction Principle:** Suppose given a sequence of statements $S(n)$, one for each $n \in \mathbb{N}$. Suppose that

1. $S(1)$ is true.
2. Whenever $S(n)$ is true then also $S(n + 1)$ is true.

Then $S(n)$ is true for all $n \in \mathbb{N}$.

Recall the definition of the integral power of a real number:

**Definition 5** Let $x$ be a non-zero real number. If $n \in \mathbb{Z}$ then:

$$ x^n = \begin{cases} 
  x, & n = 1 \\
  x x^{n-1}, & n > 1, \\
  1, & n = 0, \\
  1/x^{-n}, & n < 1.
\end{cases} $$

**Example 8** Examples of proof by induction

1. **Statement:** If $n, m \in \mathbb{N}$ and $x$ is a non-zero real number then

$$ x^{n+m} = x^n x^m. $$

**Proof:** We prove this by induction on $m$, and set

$$ S(m) : \text{For every } n, x^{n+m} = x^n x^m. $$

Then $S(1)$ reads for every $n \in \mathbb{N}$, $x^{n+1} = x^n x$, which is true by definition. Now suppose $S(m)$ is true for some $m$. Then because $S(m)$ is true,

$$ x^{n+m+1} = x^{n+m} x = x^n x^m x = x^n x^{m+1}. $$

Thus $S(m+1)$ is true, and so by induction, $S(m)$ is true for all $m$. q.e.d.

2. **Statement:** If $n, m \in \mathbb{N}$ and $x$ is a non-zero real number then

$$ (x^n)^m = x^{nm}. $$

**Proof:** We prove this by induction on $m$, and set

$$ S(m) : (x^n)^m = x^{nm}. $$

Then $S(1)$ reads for any $n \in \mathbb{N}$, $(x^n)^1 = x^n$, which is true by definition. Now suppose $S(m)$ is true for some $m$. Then because $S(m)$ is true, and using the first Statement above, we obtain

$$ x^{n(m+1)} = x^{n+m} x^n = x^{nm} x^n = x^{nm+n} = x^{n(m+1)}. $$

Therefore $S(m+1)$ is true, and so by induction $S(m)$ is true for all $m$. q.e.d.
3. **Statement:** For every \( n \in \mathbb{N} \),
\[
\sum_{k=1}^{n} k^3 = \left( \frac{n(n+1)}{2} \right)^2.
\]

**Proof:** We prove this by induction on \( n \), with \( S(n) \) the equation above.

When \( n = 1 \) both sides of the equation equal 1 and so they equal each other.

Suppose now by induction that \( S(n) \) is true for some \( n \).

Then for this \( n \), since \( S(n) \) is assumed to be true,
\[
\sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + (n+1)^3 = \left( \frac{n(n+1)}{2} \right)^2 + (n+1)^3.
\]

Factoring out \((n+1)^2\) gives
\[
\sum_{k=1}^{n+1} k^3 = (n+1)^2\left( \frac{n^2}{4} + n + 1 \right) = (n+1)^2\left( \frac{n^2 + 4n + 4}{4} \right) = \left( \frac{(n+1)(n+2)}{2} \right)^2.
\]

Thus, since we have assumed \( S(n) \) is true it follows that
\[
\sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + (n+1)^3 = \left( \frac{(n+1)(n+2)}{2} \right)^2.
\]

Therefore \( S(n+1) \) is true, and the formula follows by induction. **q.e.d.**

4. **Statement:** For any \( n \in \mathbb{N} \),
\[
\sum_{k=1}^{n} k \leq \sum_{k=1}^{n} k^2.
\]

**Proof:** When \( n = 1 \) the inequality reduces to \( 1 \leq 1 \), which is true.

Suppose the inequality holds for some \( n \).

Observe that \((n+1)^2 = (n+1)(n+1) = n^2 + 2n + 1 \leq n + 1\).

Therefore
\[
\sum_{k=1}^{n+1} k = \sum_{k=1}^{n} k + (n+1) \leq \sum_{k=1}^{n} k^2 + (n+1)^2 = \sum_{k=1}^{n+1} k^2.
\]

Therefore the inequality holds for \( n + 1 \), and so by induction the inequality holds for all \( n \). **q.e.d.**
An important application of proof by induction is the **Binomial theorem**, which requires a definition and some more notation.

**Definition 6**  
1. For any natural number $n$, **$n$ factorial** is the product of all the natural numbers from 1 to $n$:

$$n! = \prod_{k=1}^{n} k.$$  

Additionally we define 0 factorial to be 1: $0! = 1$.

2. We write

$$\binom{n}{i} = \frac{n!}{i!(n-i)!}.$$  

These numbers are called **binomial coefficients**.

Now we can state the:

**Proposition 2 (Binomial theorem)** For any real numbers $a$ and $b$, and for any $n \in \mathbb{N}$,

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$  

To prove this we first need to establish an important property of the binomial coefficients:

**Lemma 2** For any natural number, $n$, and any natural number $i \leq n + 1$,

$$\binom{n}{i - 1} + \binom{n}{i} = \binom{n + 1}{i}.$$  

**Proof:** First, we use the definition above to rewrite the left hand side as

$$\frac{n!}{(i-1)!(n-(i-1))!} + \frac{n!}{i!(n-i)!}.$$  

Multiply the numerator and the denominator of the first term by $i$ to get

$$\frac{n!}{(i-1)!(n-(i-1))!} = \frac{n! \cdot i}{i!(n+1-i)!}.$$  

Multiply the numerator and the denominator of the second term by $n+1-i$ to get

$$\frac{n!}{i!(n-i)!} = \frac{n! \cdot (n+1-i)}{i!(n+1-i)!}.$$  

Then multiply the numerator and the denominator of the second term by $n+1-i$ to get

$$\frac{n!}{i!(n-i)!} = \frac{n! \cdot (n+1-i)}{i!(n+1-i)!}.$$  

Finally, add together to get

$$\frac{n!}{(i-1)!(n-(i-1))!} + \frac{n!}{i!(n-i)!} = \frac{n! \cdot (i + n + 1 - i)}{(i)!(n+1-i)!} = \frac{(n+1)!}{(i)!(n+1-i)!}.$$  

21
Proof of the binomial theorem: We prove this by induction on \( n \).
When \( n = 1 \) the statement reduces to \( a + b = a + b \).
Next, assume the statement is true for \( n \). Then it follows that
\[
(a + b)^{n+1} = (a + b)(a + b)^n = (a + b) \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.
\]
Multiplying separately by \( a \) and \( b \) and adding, we get
\[
(a + b)^{n+1} = \sum_{i=0}^{n} \binom{n}{i} a^{i+1} b^{n-i} + \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i+1}.
\]
We consider the two terms separately. The second term may be written as
\[
\sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i+1} = b^{n+1} + \sum_{i=1}^{n} \binom{n}{i} a^i b^{n-i+1}.
\]
The first term may be written as
\[
\sum_{i=0}^{n-1} \binom{n}{i} a^{i+1} b^{n-i} + a^{n+1}.
\]
For this first term set \( k = i + 1 \). Then as \( i \) runs from 0 to \( n - 1 \), \( k \) runs from 1 to \( n \), and
\[
n - i = (n + 1) - (i + 1) = n + 1 - k.
\]
Thus the first term may be rewritten as
\[
\sum_{k=1}^{n} \binom{n}{k-1} a^k b^{n+1-k} + a^{n+1}.
\]
Now in this sum, \( k \) is a "dummy variable". We could have used any symbol such as \( m, n, p, \ldots \). In particular we could have used \( i \). Thus this term may be rewritten as
\[
\sum_{i=1}^{n} \binom{n}{i-1} a^i b^{n+1-i} + a^{n+1}.
\]
Finally, adding the two terms together we get
\[
(a + b)^{n+1} = a^{n+1} + \sum_{i=1}^{n} \left( \binom{n}{i-1} + \binom{n}{i} \right) a^i b^{n+1-i} + b^{n+1}.
\]
Hence we may apply Lemma 2 to conclude that
\[
(a + b)^{n+1} = \sum_{i=0}^{n+1} \binom{n+1}{i} a^i b^{n+1-i}.
\]
This shows that the statement is true for \( n + 1 \), and so by induction it is true for all \( n \). \textbf{q.e.d.}

There is another form of the induction principle we shall also use, as illustrated in the next Proposition:

**Proposition 3** Suppose given a sequence of statements \( S(n) \), one for each \( n \in \mathbb{N} \), and that

1. \( S(1) \) is true.
2. Whenever \( S(k) \) is true for \( k \leq n \) then also \( S(n + 1) \) is true.

Then \( S(n) \) is true for all \( n \in \mathbb{N} \).

**Proof:** Let \( T(n) \) be the statement: \( S(k) \) is true for all \( k \leq n \). Since \( S(1) \) is true, so is \( T(1) \).

Suppose by induction that \( T(n) \) is true.
Then \( S(k) \) is true for all \( k \leq n \).
Therefore by hypothesis, \( S(n + 1) \) is true.
Since \( S(k) \) is true for \( k \leq n \) as well, it is true for \( k \leq n + 1 \).
Thus \( T(n + 1) \) is true.
Now it follows from the induction principle that \( T(n) \) is true for all \( n \).
Therefore \( S(n) \) is true for all \( n \). \textbf{q.e.d.}

The principle of induction can also be used for construction, via the following

**Construction by Induction Principle** A sequence \( (x_n)_{n \geq 1} \) can be constructed as follows:

1. First, construct \( x_1 \).
2. Then, assume for some \( n \) that \( x_i \) has been constructed for \( i \leq n \), and give an explicit construction for \( x_{n+1} \).

The induction construction principle then states that this constructs the full sequence.

**Example 9 Construction by induction**

1. An infinite sequence \( (x_n)_{n \geq 1} \) can be constructed by induction as follows:
   (a) Set \( x_1 = 2 \).
   (b) Assume \( x_i \) is constructed for \( i \leq n \) and set
   \[
   x_{n+1} = \sum_{i=1}^{n} (-1)^i x_i.
   \]
2. An infinite sequence \((x_n)_{n \geq 1}\) satisfying \(x_{n+1} > x_n^2\) can be constructed by induction as follows:

   (a) Set \(x_1 = 1\)

   (b) Assume \(x_i\) is constructed for \(i \leq n\) and set

       \[ x_{n+1} = x_n^2 + 1. \]

3. An infinite sequence of natural numbers \((a_k)_{k \geq 1}\) can be constructed by induction by setting

       \[ a_1 = 1, \quad and \]

       \[ a_{k+1} = \begin{cases} \frac{\sum_{i=1}^{i=k} a_i}{k}, & \text{if } \sum_{i=1}^{i=k} a_i \text{ is even;} \\ k, & \text{if } \sum_{i=1}^{i=k} a_i \text{ is odd.} \end{cases} \]

**Exercise 3 Induction**

1. Show by induction that \(\sum_{i=0}^{n} i = \frac{n(n+1)}{2}\).

2. Show by induction that \(\sum_{i=0}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}\).

3. Use the Binomial theorem to show that if \(p, k \in \mathbb{N}\) then

       \[ \left( \frac{p+1}{p} \right)^k > \frac{k}{p}. \]

Conclude that

       \[ \left( \frac{p}{p+1} \right)^k < \frac{p}{k}. \]

4. Show by induction on \(n \in \mathbb{N}\) that the product of \(n\) odd natural numbers is odd.

5. Construct by induction an infinite sequence of natural numbers \((x_n)_{n \geq 1}\) such that for \(n \geq 2\), \(x_n > \sum_{i=1}^{n-1} x_i\).

**2.6 Counterexamples**

Suppose you are given a statement of such as "**all natural numbers are prime**", and you want to prove that it is false. Since the statement asserts a property for all natural numbers, you only have to find a single example (eg. \(6 = 2 \times 3\)) to show that the statement is false.

This is an example of a counterexample: counterexamples are examples which demonstrate that assertions are false. Thus the example asked for in Exercise 2.3 is a counterexample to the assertion that all natural numbers have rational square roots.

There is a much deeper example from the theory of equations. You learned in school that the quadratic equation

\[ ax^2 + bx + c = 0 \]
has two solutions, $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. One might ask whether the solutions to higher order equations can also be expressed in formulas that only use $n^{th}$ roots, for some natural numbers $n$, and indeed this is true for cubic and quartic equations.

However, it is not true for all fifth degree equations, as was shown by Abel and Ruffini in 1824. The easiest way to show this is by an explicit counterexample and a very simple one was proved about a century ago by a famous algebraist, Emil Artin, who showed that roots of the equation $x^5 - x - 1 = 0$ cannot be expressed in this way.

**Exercise 4 Counterexamples**

1. Show that the following statement is false: If $a, b \in \mathbb{N}$ then $a^3 + b^3$ is the cube of a natural number.

2. Prove that the sum of an odd number of natural numbers is odd if each of the natural numbers is odd. What is the converse statement? Show by a counterexample that the converse is false.

3. What is the converse to the statement: "Suppose $n$ is an integer. If $n$ is an odd natural number then $n^2$ is odd". Decide if it is true or false and prove your answer.

4. Prove or disprove the following statement: For every $x \in \mathbb{Q}$ there is a unique $n \in \mathbb{N}$ which is the closest natural number to $x$.

5. Construct an example to show that if in each of two school classes the average GPA of the boys is bigger than that of the girls it may not be the case that when the classes are combined this is still true: i.e., in the combined classes the average GPA of the girls may be bigger than that of the boys.

6. Show by counterexample that the following statement is false: "For every $p, q \in \mathbb{N}$, $p^2 + q^2 > (p + q)^2 - 50p - 51q$.

7. Show by counterexample that the following statement is false: "If $t, u, s \in \mathbb{N}$ satisfy $s > 39, t > s + 5$, and $u > 33$, then $t + u > 82$.

**2.7 The Literature of Mathematics**

Mathematicians have built up a body of knowledge over several thousand years, expressed in theorems and examples, all validated by rigorous proofs, and that once validated are then true forever. This is the literature of mathematics. Each new piece of knowledge depends on what came before, and sometimes it takes many decades for little pieces to fit together to establish some remarkable new phenomenon.
The classical example is 'Fermat’s last theorem'! Fermat was a 17th century number theorist, and after he died in 1637 the following statement was found written in the margin of one of his books: "I have found the most wonderful result, but the margin is too small for me to write down the proof." The simple assertion was this: If $a, b, c, n$ are natural numbers all greater than 1, and if

$$a^n + b^n = c^n,$$

then $n = 2$.

In the following three hundred years the search for a proof was one of the 'holy grails' of mathematics. Much of our knowledge in number theory and geometry was developed in the process. And finally in 1995 Andrew Wiles (Princeton) published a proof using many of the results which had been established over the preceding centuries.

Some of you may wonder about the relation of mathematics to the physical world we live in. Now mathematical knowledge itself is validated just by proofs, and so its correctness stands for all time and does not depend in any way on our physical experience. Nonetheless, much of mathematics is inspired by that physical experience, mathematical knowledge is regularly applied in almost every branch of science and engineering, and indeed many mathematicians focus on building mathematical models and computational algorithms that directly reflect/compute physical phenomena. As a Nobel Laureate in Physics, Eugene Wigner, wrote in 1960 in a paper entitled *The Unreasonable Effectiveness of Mathematics* "...the mathematical formulation of the physicist's often crude experience leads in an uncanny number of cases to an amazingly accurate description of a large class of phenomena".
Chapter 3

Basic Set Theory

3.1 Sets

The basic vocabulary of Mathematics is that of sets and maps. It is used in every branch of the subject.

Definition 7 Sets

1. As stated in Section 2.1, a set is any explicitly specified collection of objects, abstract or concrete.

2. The 'objects' are called the elements of the set.

3. If \( x \) is an element in a set \( S \) we say \( x \) belongs to \( S \), and denote this by \( x \in S \).

4. A finite set is a set with only finitely many elements. In this case \( |S| \) denotes the number of elements in \( S \).

5. A set which is not finite is called infinite.

6. The empty set is the set with no elements. It is denoted by \( \phi \).

Note: To define a set we must specify its elements.

Example 10 Sets
1. The set $S = \{1, 2, 3\}$ has three elements: 1, 2, and 3 and so $|S| = 3$.

2. The empty set $\emptyset$ has no elements and so $|\emptyset| = 0$.

3. Each of $N, Q, R$, and $C$ are infinite sets.

4. The unit interval of all real numbers $x$ satisfying $0 \leq x \leq 1$ is an infinite set.

5. The desks in this classroom are a finite set.

6. The real numbers whose squares are natural numbers are a set.

### 3.2 Operations with sets

**Definition 8** Subsets

A subset of a set $S$ is a set $W$ all of whose elements are elements of $S$. We denote this by $W \subset S$ and we often write

$$W = \{ x \in S \mid \ast \cdots \ast \},$$

where '$\ast \cdots \ast$' specifies which elements of $S$ are in $W$.

**Example 11** Subsets

1. The set $S$ of even integers is the subset of $Z$ given by

$$S = \{ n \in Z \mid n \text{ is divisible by } 2 \}.$$

   It is not a finite set.

2. $N$ is the subset of $Z$ defined by

$$N = \{ n \in Z \mid n > 0 \}.$$

3. $Q$ is the subset of $R$ given by

$$Q = \{ x \in R \mid x = p/q \text{ with } p \in Z \text{ and } q \in N \}.$$

4. The empty set is a subset of every set.

5. Let $L$ be the set whose elements are all the lines in the plane. Thus a line in the plane is a subset of $R^2$, and is also an element in $L$.

6. The vertical lines in the plane are a subset of $L$.

**Definition 9** Products
1. The **product** of two sets $S$ and $T$ is the set $S \times T$ whose elements are the ordered pairs $(x, y)$ with $x \in S$ and $y \in T$. We often write this as

\[ S \times T = \{(x, y) \mid x \in S \text{ and } y \in T\}. \]

**Example 12 Products**

1. The product of the sets $S = \{2, 17, 99\}$ and $T = \{5, \sqrt{2}\}$ is the set

\[ S \times T = \{(2, 5), (17, 5), (99, 5), (2, \sqrt{2}), (17, \sqrt{2}), (99, \sqrt{2})\}. \]

Thus in this example, $|S| = 3, |T| = 2$ and $|S \times T| = 6$.

2. The product $R \times R$ is the set $R^2$, of points in the plane:

\[ R^2 = \{(x, y) \mid x \in R \text{ and } y \in R\}. \]

3. The product of $Q$ with $R$ is the set \{a, y\} $a \in Q, y \in R\} of points in the plane whose x-coordinate is rational.

4. The set $Q$ is the set of real numbers of the form $p/q$ with $p \in Z$ and $q \in N$. This is **not** the product

\[ Z \times N = \{(p, q) \mid p \in Z, q \in N\}, \]

because $(2, 1)$ and $(4, 2)$ are different elements in the product but $2/1 = 4/2$.

**Definition 10 Unions and intersections**

Suppose $S_{\alpha}$ is a collection of sets. Then

1. The **union** of the $S_{\alpha}$ is the set $S = \bigcup_{\alpha} S_{\alpha}$ whose elements are the elements which belong to **at least one** of the $S_{\alpha}$:

\[ \bigcup_{\alpha} S_{\alpha} = \{x \in S \mid x \in S_{\alpha} \text{ for some } S_{\alpha}\}. \]

2. The **intersection** of the $S_{\alpha}$ is the set $\bigcap_{\alpha} S_{\alpha}$ whose elements are the elements which belong to **each** of the $S_{\alpha}$:

\[ \bigcap_{\alpha} S_{\alpha} = \{x \in S \mid x \in S_{\alpha} \text{ for every } S_{\alpha}\}. \]

**Definition 11** The **power set** $P(S)$ is the set whose elements are all the subsets of $S$.

**Example 13 Unions, intersections, and power sets**
1. The union of the sets \( \{1,2,3\} \) and \( \{2,4,8\} \) is \( \{1,2,3,4,8\} \). The intersection of these sets is \( \{2\} \). The original two sets each have three elements, the union has five elements, and the intersection has one element.

2. The union of the sets \( \{1,2\}, \{2,3\} \) and \( \{3,4\} \) is the set \( \{1,2,3,4\} \). The intersection of the first three sets is \( \phi \).

3. Let \( S \) be the set of even integers and let \( T \) be the set of odd integers. Then

\[
S \cup T = \mathbb{Z} \quad \text{and} \quad S \cap T = \phi.
\]

4. For each \( n \in \mathbb{N} \) let \( S_n \subset \mathbb{N} \) be the subset of all natural numbers except \( n \). Then

(a) \[
\biguplus_n S_n = \mathbb{N},
\]

since every natural number is in some (in fact in almost all) \( S_n \).

(b) \[
\bigcap_n S_n = \phi,
\]

because any natural number \( k \) is not \( S_k \) and so is not in every \( S_n \).

5. For any set \( S, S \in P(S) \) and \( \phi \in P(S) \) because \( S \) and \( \phi \) are subsets of \( S \).

6. The power set \( P(\{1,2\}) \) is given by

\[
P(\{1,2\}) = \{\phi, \{1\}, \{2\}, \{1,2\}\}.
\]

**Exercise 5 Sets**

1. Let \( S \) be the set of all even integers, and let \( T \) be the set of all integers divisible by 3.

(a) What is \( S \cap T \) ?

(b) Let \( W \) be the set of integers which are not in \( S \cup T \). Is \( W \) infinite or finite?

2. Suppose \( S \) and \( T \) and \( |S| = 3 \) and \( |T| = 4 \)

(a) What is \( |S \times T|\) ?

(b) What is \( |P(T)|\) ?

3. Two sets \( S \) and \( T \) are disjoint if \( S \cap T = \phi \).

(a) If \( S \) and \( T \) are disjoint sets show that for any set \( W \), \( S \times W \) and \( T \times W \) are disjoint.
(b) If \( S \) and \( T \) are disjoint sets for which sets \( W \) are \( S \cup W \) and \( T \cup W \) disjoint?

(c) If \( S \) and \( T \) are disjoint sets show that \( P(S) \cap P(T) = \{ \phi \} \).

4. If \( S \) and \( T \) are finite sets show that

\[
|S \cup T| + |S \cap T| = |S| + |T|.
\]

5. Let \( S \) be a finite set and suppose \( x \notin S \). Show that

\[
|P(S \cup \{x\}| = 2|P(S)|
\]

6. Use the previous problem and induction to prove that if \( S \) is a finite set then \( |P(S)| = 2^{|S|} \).

### 3.3 Maps

**Definition 12 Maps**

1. A map \( \varphi : S \rightarrow T \)

between sets consists of **three** things:

   (a) A set \( S \), called the **domain** of the map;

   (b) A set \( T \), called the **target** of the map; and

   (c) A **rule** which assigns to each element \( x \in S \) a **single** specified element \( y \in T \). The element \( y \) is called the **image** of \( x \) and is denoted by \( y = \varphi(x) \).

2. If \( \varphi : S \rightarrow T \) is a map between sets then the **image** of \( \varphi \) is the subset \( \text{Im} \varphi \subset T \) defined by

\[
\text{Im} \varphi = \{ \varphi(x) \mid x \in S \}.
\]

3. If \( \varphi : S \rightarrow T \) is a map, and if \( U \subset T \) is a subset then we write

\[
\varphi(U) = \{ y \in T \mid y = \varphi(x) \text{ for some } x \in U \}.
\]

In particular, \( \text{Im} \varphi = \varphi(S) \).

4. If \( S \) and \( T \) are sets, the set whose elements are all the maps from \( S \) to \( T \) is denoted by \( T^S \).

**Note:**

1. To define a map we must specify **three** things: the domain, the target, and the rule - see Example 14, below.
2. **Do not confuse** a set, $S$ with a map, $\varphi$.
   
   (a) A set is a collection of objects.
   
   (b) A map, $\varphi : S \rightarrow T$ is a rule connecting each element $x \in S$ to a single element $\varphi(x) \in T$.

**Example 14 Maps**

1. The "rule" which assigns both 1 and 2 to 1 and $k$ to $k$ for $k \geq 2$ is **not a map** from $N$ to $N$ because two different elements are assigned to 1.

2. If $S$ and $T$ are sets then the collection of all maps from $S$ to $T$ is a set.

3. Three maps are defined as follows:
   
   (a) $\varphi : N \rightarrow N; \varphi(n) = n^2 + 1$.
   
   (b) $\psi : Z \rightarrow N; \psi(n) = n^2 + 1$.
   
   (c) $\chi : N \rightarrow Z; \chi(n) = n^2 + 1$.
   
   These three maps are all **different even though the "formula" is the same**, because the two sets are different in each case.

4. A map $f$ from $R$ to $R$ is defined by $f(x) = x^2$, and a map $g$ from $R$ to the set $T$ of all non-negative real numbers defined by $g(x) = x^2$. Note: as above, $f$ and $g$ are **different maps** even though the formula is the same.

5. $S$ is the set of students enrolled in the University of Maryland on Oct. 1, 2016 and $\varphi : S \rightarrow N$ is the rule which assigns to each student the least number of inches which is greater than or equal to their height.

6. $S$ is the set of students enrolled in the University of Maryland on Oct. 1, 2016, $C$ is the set of all countries that existed prior to Oct. 1, 2016, and $\varphi$ is the rule which assigns to each student the country in which they were born.

7. $\psi : N \rightarrow \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ assigns to each $n \in N$ the $n^{th}$ integer in the decimal expansion of $\pi$. 

32
Remark: A map from $\mathbb{R}$ to $\mathbb{R}$ is often called a real valued function, and the formulas you are used to from calculus usually define real valued functions on $\mathbb{R}$. However, not every real valued function has a formula.

**Definition 13** Composites

If $\varphi : S \to T$ and $\psi : U \to W$ are maps, and if $T \subset U$, then the composite is the map $\psi \circ \varphi : S \to W$ defined by

$$(\psi \circ \varphi)(x) = \psi(\varphi(x)), \quad x \in S.$$  

**Important Note:** The composite $\psi \circ \varphi$ is only defined if the target of $\varphi$ is contained in the domain of $\psi$!

**Example 15** Composites

1. The composite $g \circ f$ of a map $f : \mathbb{N} \to \mathbb{R}$ and a map $g : \mathbb{N} \to \mathbb{R}$ is not defined because the domain of $g$ does not contain the target of $f$.

2. If $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are defined by $f(x) = x^2$ and $g(x) = x + 1$ then

$$ (g \circ f)(x) = x^2 + 1 \quad \text{and} \quad (f \circ g)(x) = (x + 1)^2. $$

3. Given $a \in \mathbb{Q}$ we may write $a = p/q$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$. However, $p, q$ are not determined by $a$, since if $a = p/q$ then also $a = 2p/2q$. But we can define maps

$$ f : \mathbb{Q} \to \mathbb{Z} \quad \text{and} \quad g : \mathbb{Q} \to \mathbb{N} $$

by setting $g(a)$ to be the least possible $q$ and $f(a)$ to be the corresponding $p$. Let $h : \mathbb{Z} \subset \mathbb{Q}$ be the inclusion. Then

$$ (f \circ h)(p) = p, \quad p \in \mathbb{Z}. $$

There are three very important classes of maps:

**Definition 14** Onto maps, 1-1 maps, and bijections

1. A map $\varphi : S \to T$ is **onto** if $\text{Im} \varphi = T$. Thus $\varphi$ is onto if and only if every element $y$ can be written in the form $y = \varphi(x)$ for at least one $x \in S$.

2. A map $\varphi : S \to T$ is **1-1** if it maps different elements in $S$ to different elements in $T$. Thus, the map $\varphi$ is 1-1 if and only if for all $x_1, x_2 \in S$,

$$ x_1 \neq x_2 \Rightarrow \varphi(x_1) \neq \varphi(x_2). $$
3. A map $\varphi : S \to T$ is a bijection if and only if it is both onto and 1-1.

Caution: Do not confuse the definition of a map with the definition of a 1-1 map:

1. A map $\varphi : S \to T$ associates to each $x \in S$ a single element $\varphi(x) \in T$.

2. A map $\varphi : S \to T$ is 1-1 if whenever $x_1$ and $x_2$ are different elements in $S$ then the elements $\varphi(x_1)$ and $\varphi(x_2)$ in $T$ are also different.

In other words:

1. If $\varphi$ is a map from $S$ to $T$, then for each $x \in S$ there is a single $y \in T$ such that $\varphi$ maps $x$ to $y$.

2. If $\varphi$ is a 1-1 map from $S$ to $T$, then
   (a) for each $x \in S$ there is a single $y \in T$ such that $\varphi$ maps $x$ to $y$ (because $\varphi$ is a map) AND
   (b) for each $y$ in the image of $\varphi$ there is a single $x \in S$ which is mapped to $y$.

Important Remark: A map $\varphi : S \to T$ is a bijection if and only if:

1. $T = \text{Im } \varphi$ ($\varphi$ is onto) and also
2. For every $y \in \text{Im } \varphi$ there is a single $x$ such that $\varphi(x) = y$. ($\varphi$ is 1-1.)

Putting these two statements together we get

$\varphi$ is a bijection $\iff$ for each $y \in T$ there is a single $x \in S$ which is mapped to $y$.

Example 16 1. In the third example in Example 14, $\varphi$ is 1-1 but not onto, $\psi$ is neither 1-1 or onto, and $\chi$ is 1-1 but not onto.

2. The map $\varphi : \mathbb{N} \to \mathbb{N}$ given by
   $$\varphi(m) = \begin{cases} 
   1, & m = 1, 2, \\
   m - 1, & m \geq 3
   \end{cases}$$
   is onto. It is not 1-1 because both 1 and 2 are mapped to 1.

3. The map $\varphi : \mathbb{Z} \to \mathbb{N}$ given by
   $$\varphi(k) = \begin{cases} 
   -2(k - 1), & k \leq 0, \\
   2k - 1, & k \geq 1
   \end{cases}$$
   is 1-1 and onto. Thus it is a bijection.
4. Let $f : \{1, 2, 3, 4\} \to \{a, b, c, d\}$ be defined by $f(1) = c$, $f(2) = a$, $f(3) = d$, and $f(4) = b$. And let $h : \{a, b, c, d\} \to \{a, b, c, d\}$ be defined by $h(a) = b$, $h(b) = c$, $h(c) = d$, $h(d) = d$. Then

(a) $f \circ h$ is not defined.
(b) $h$ is not onto and is not 1-1.
(c) $h \circ f$ is given by $(h \circ f)(1) = d$, $(h \circ f)(2) = b$, $(h \circ f)(3) = d$, and $(h \circ f)(4) = c$.

5. The map $\varphi : N \to N$ defined by

$$\varphi(n) = n + 1$$

is 1-1 but not onto, since 1 is not in the image.

6. The map $\psi : N \to N$ defined by

$$\psi(1) = \psi(2) = 1 \text{ and } \psi(n) = n - 1 \text{ if } n \geq 3$$

is onto but not 1-1.

**Exercise 6 Maps**

1. Suppose $f : S \to T$ is a map and $U$ and $W$ are subsets of $S$.

(a) If $U \cap W = \emptyset$ does it follow that $f(U) \cap f(W) = \emptyset$?

(b) If $f(U) \cap f(W) = \emptyset$ does it follow that $U \cap W = \emptyset$?

2. List all the maps $\varphi$ from $S = \{1, 2\}$ to $T = \{-1, -2\}$ such that $\text{Im } \varphi = T$.

3. Construct sets $S, T, U, W$ and maps $\varphi : S \to T$ and $\psi : U \to W$ such that the composite $\psi \circ \varphi : S \to W$ is not defined but the composite $\varphi \circ \psi : U \to T$ is defined.

4. If $f : R \to R$ and $g : \{x \in R \mid x \geq 0\} \to R$ are defined by $f(x) = x^2$ and $g(x) = \sqrt{x}$ are $f \circ g$ and $g \circ f$ defined?

5. Construct an example of maps $f : S \to T$ and $g : T \to S$ in which $f$ is 1-1 and $g$ is onto, but the composite $g \circ f$ is neither 1-1 nor onto.

6. Construct a bijection from $Z$ to $N$.

7. If either $S = \emptyset$ or $T = \emptyset$ show that $T^S = \emptyset$.

**Definition 15** Associativity, the identity, and inverses.

1. The **identity map** of a set $S$ is the map $\text{id}_S : S \to S$ defined by

$$\text{id}_S(x) = x, \quad x \in S.$$
2. If \( \phi : S \to T \) is a bijection, then the inverse of \( \phi \) is the map \( \phi^{-1} : T \to S \) defined by 

\[
\phi^{-1}(y) = \text{the unique } x \in S \text{ such that } \phi(x) = y.
\]

**Proposition 4**  Suppose \( \phi : S \to T \), \( \psi : T \to W \), and \( \chi : W \to U \) are maps of sets. Then

1. For \( x \in S \), 
   \[
   \text{id}_T \circ \phi = \phi \quad \text{and} \quad \phi \circ \text{id}_S = \phi.
   \]

2. (Associativity of composition): 
   \[
   (\chi \circ \psi) \circ \phi = \chi \circ (\psi \circ \phi).
   \]

3. If \( \phi \) is a bijection then 
   \[
   \phi^{-1} \circ \phi = \text{id}_S \quad \text{and} \quad \phi \circ \phi^{-1} = \text{id}_T.
   \]

4. If \( \xi : T \to S \) is a map satisfying 
   \[
   \xi \circ \phi = \text{id}_S \quad \text{and} \quad \phi \circ \xi = \text{id}_T,
   \]
   then \( \phi \) is a bijection and \( \xi = \phi^{-1} \).

**Proof:**

1. For \( x \in S \), 
   \[
   (\text{id}_T \circ \phi)(x) = \text{id}_T(\phi(x)) = \phi(x) \quad \text{and} \quad (\phi \circ \text{id}_S)(x) = \phi(\text{id}_S(x)) = \phi(x).
   \]

2. For \( x \in S \), 
   \[
   ((\chi \circ \psi) \circ \phi)(x) = (\chi \circ \psi)(\phi(x)) = \chi(\psi(\phi(x)))
   \]
   and 
   \[
   (\chi \circ (\psi \circ \phi))(x) = \chi((\psi \circ \phi)(x)) = \chi(\psi(\phi(x))).
   \]
   Thus both sides evaluated at any \( x \in S \) give \( \chi(\psi(\phi(x))) \).

3. For \( x \in S \), \( x \) is the unique element mapped to \( \phi(x) \) by \( \phi \). Thus by definition 
   \[
   \phi^{-1}(\phi(x)) = x = \text{id}_S(x).
   \]
   Moreover, if \( y \in T \) the \( y = \phi(x) \) for a unique \( x \in S \). By definition, 
   \( x = \phi^{-1}(y) \). Thus 
   \[
   \phi(\phi^{-1}(y)) = \phi(x) = y = \text{id}_T(y).
   \]
4. If \( y \in T \) then
\[
y = \varphi(\xi(y)) \in \text{Im } \varphi.
\]
Therefore \( \varphi \) is onto.

Moreover, if \( x_1, x_2 \in S \) and \( \varphi(x_1) = \varphi(x_2) \) then
\[
x_1 = \xi(\varphi(x_1)) = \xi(\varphi(x_2)) = x_2.
\]
Therefore \( \varphi \) is 1-1, and so it is a bijection.

Finally,
\[
\xi = \xi \circ \text{id}_T = \xi \circ (\varphi \circ \varphi^{-1}) = (\xi \circ \varphi) \circ \varphi^{-1} = \circ \varphi^{-1}.
\]
q.e.d.

Example 17 Inverses

1. Let \( f : \{1, 2, 3, 4\} \to \{a, b, c, d\} \) be defined by \( f(1) = c, f(2) = a, f(3) = d, \) and \( f(4) = b. \) Then \( f \) is a bijection and its inverse is given by \( f^{-1}(a) = 2, f^{-1}(b) = 4, f^{-1}(c) = 1, f^{-1}(d) = 3. \)

2. Let \( A \) be an \( n \times n \) matrix, and regard \( A \) as a map from the set of column vectors to itself by setting \( A(v) = Av \) (matrix multiplication). If \( B \) is a second \( n \times n \) matrix, then \( (A \circ B)(v) = A(B(v)) = A(Bv) = (AB)v. \) Thus the map \( A \circ B \) is just multiplication by the product matrix \( AB. \)

3. In the example above, suppose \( A \) is an invertible matrix with inverse matrix \( A^{-1}. \) Then using elementary linear algebra it follows that multiplication by \( A \) is a bijection, and that the inverse map is multiplication by \( A^{-1}. \)

4. It follows from what you learn in an elementary calculus course that the map \( f : \mathbb{R} \to \mathbb{R} \) given by \( f(x) = x^3 \) is a bijection. The inverse maps any real number \( x \) to its unique cube root.

5. Let \((x_n)_{n \geq 1}\) be a sequence of real numbers such that for all \( n, x_n < x_{n+1}. \)

Then define \( S \subset \mathbb{R} \) by
\[
S = \{x_n \mid n \geq 1\}.
\]

Define \( f : \mathbb{R} \to \mathbb{R} \) by setting
\[
f(x) = \begin{cases} 
  x, & x \notin S \\
  x_{2k-1}, & x = x_{2k}, \\
  x_{2k}, & x = x_{2k-1}.
\end{cases}
\]

Then \( f \) is a bijection.
Proposition 5 Let $S$ be a non-void set. Then a bijection,
\[ \gamma : P(S) \to \{0, 1\}^S \]
is given by
\[ \gamma(U)(x) = \begin{cases} 
1 & x \in U \\
0 & x \notin U.
\end{cases} \]

Proof: Define $\chi : \{0, 1\}^S \to P(S)$ by setting
\[ \chi(f) = \{ x \in S \mid f(x) = 1 \}, \]
where $f : S \to \{0, 1\}$ is any map. Then for $U \subset S$,
\[ (\chi \circ \gamma)(U) = \{ x \in S \mid \gamma(U)(x) = 1 \} = U. \]
Also, for $f : S \to \{0, 1\}$,
\[ (\gamma \circ \chi(f)) = \gamma(\{ x \in S \mid f(x) = 1 \}) = f. \]
Therefore by Proposition 4.4, $\gamma$ is a bijection with inverse $\chi$. q.e.d.

Theorem 1 If $S$ is any non-void set then no map from $S$ to $P(S)$ is onto.

Proof: In view of Proposition 5 it is sufficient to prove that no map from $S$ to $\{0, 1\}^S$ is onto. Let
\[ \Phi : S \to \{0, 1\}^S \]
be any map. Then for each $x \in S$, $\Phi(x)$ is a map from $S$ to $\{0, 1\}$. Define
\[ f : S \to \{0, 1\} \]
by setting
\[ f(x) = \begin{cases} 
1 & \Phi(x)(x) = 0 \\
0 & \Phi(x)(x) = 1.
\end{cases} \]
Then by construction it is never true for any $x \in S$ that $\Phi(x)(x) = f(x)$. But if $\Phi(y) = f$ for some $y \in S$ then
\[ \Phi(y)(y) = f(y), \]
which is impossible. Thus $f$ is not in $Im \Phi$ and so $\Phi$ is not onto. q.e.d.

Exercise 7 Sets and maps

1. Determine which of the following is a map:
   (a) $\varphi : \mathbb{N} \to \mathbb{N}$, defined by
   \[ \varphi(x) = \begin{cases} 
1, 2 & x = 1, \\
x + 2 & x > 1.
\end{cases} \]
(b) \( \psi : \mathbb{R} \to \mathbb{R} \), defined by
\[
\psi(x) = \begin{cases} 
1 & x = 1, \\
x + 2 & x \neq 1.
\end{cases}
\]

(c) \( \omega : \mathbb{C} \to \mathbb{C} \) defined by:
\[
\omega(x) = \begin{cases} 
5 & x = 1, 2, 3 \\
x + 2 & x \neq 1, 2, 3.
\end{cases}
\]

2. Which of the maps in Problem 1. are 1-1, onto?

3. For \( \psi \) and \( \omega \) in Problem 1. evaluate \( \psi \circ \psi \) and \( \omega \circ \omega \).

4. Below are several attempted proofs showing that the map \( \varphi : \mathbb{N} \to \mathbb{N} \) defined by \( \varphi : n \mapsto n^2 \) is 1-1. Choose the proof that is correctly written.

(a) Recall a map \( \psi : S \to \mathbb{R} \) is 1-1 if and only if for every \( x, y \in S \):
\[ \varphi(x) = \varphi(y) \text{ if and only if } x = y. \] Recall that every natural number has a unique square. Thus we must prove that if \( k, n \in \mathbb{N} \) and if \( k^2 = n^2 \) then \( k = n \). We do this by contradiction. If this is not true then either \( k > n \) or \( k < n \). If \( k > n \) then \( k = n + m \) with \( m \in \mathbb{N} \). It follows that \( k^2 = n^2 + 2nm + m^2 > n^2 \), which is a contradiction. The same argument applies if \( k < n \); just switch the roles of \( k \) and \( n \). q.e.d.

(b) Recall a map \( \varphi : S \to \mathbb{R} \) is 1-1 if and only if for every \( x, y \in S \):
\[ \varphi(x) = \varphi(y) \text{ if } x = y. \] Recall that every natural number has a unique square. Therefore if \( k = n \) then \( k^2 = n^2 \). Therefore the map is 1-1. q.e.d.

(c) Define a map \( \alpha : \mathbb{N} \to \mathbb{N} \) have the rule \( \alpha(x) = n^2 \) for \( n \in \mathbb{N} \). Notice \( (n^2)^{1/2} = n^{2/2} = n \). q.e.d.

(d) Recall a map \( \psi : S \to \mathbb{R} \) is 1-1 if and only if for every \( x, y \in S \):
\[ \varphi(x) = \varphi(y) \text{ if and only if } x = y. \] Notice \( k^2 = n^2 \Rightarrow k = n \). Therefore \( k^2 = n^2 \) if and only if \( k = n \). Therefore the map is 1-1. q.e.d.

5. If \( S \) and \( T \) are finite sets show that \( |S \times T| = |S||T| \).

6. If \( S \) is a finite set show that \( |P(S)| = 2^{|S|} \).

7. Prove that:

(a) for any set \( S \), \( \text{id}_S \) is a bijection.
(b) The composite of onto set maps $\varphi$ and $\psi$ is onto if the target of $\varphi$ is the domain of $\psi$.

(c) The composite of 1-1 set maps is 1-1.

(d) If $\varphi : S \to T$ and $\psi : T \to W$ are bijections show that the inverse of $\varphi$ and that $\psi \circ \varphi$ are bijections.

8. Suppose $\varphi : S \to T$ is a map.

(a) If $T$ is finite and the map is 1-1 show that $S$ is finite and that $|S| \leq |T|$.

(b) If $S$ is finite and the map is onto show that $|T|$ is finite and that $|S| \geq |T|$.

9. Suppose $\varphi : S \to T$ is a map and that $S$ and $T$ are finite sets with the same number of elements. Show that the following three conditions are equivalent:

(a) $\varphi$ is a bijection.

(b) $\varphi$ is onto.

(c) $\varphi$ is 1-1.

10. Construct two maps $\alpha$ and $\beta$, both from the same set $S$ to the same set $T$, and such that $S$ and $T$ are different sets and $\alpha$ is 1-1 but not onto and $\beta$ is onto but not 1-1.

11. If $S$ and $T$ are finite sets show that there are $|T|^{|S|}$ elements in the set of maps from $S$ to $T$.

12. How many elements are there in the subset of bijections of a finite set $S$ to itself?

### 3.4 Equivalence Relations

A relation between two sets, $S$ and $T$ is a generalization of the idea of a map:

**Definition 16** A relation, $R$, between a set $S$ and a set $T$ is a subset $R \subset S \times T$ of the product of $S$ with $T$.

If $(x, y) \in R$ we say that $y$ is **related** to $x$ by the relation $R$, and we write $xRy$.

If $S=T$ we say that $R$ is a relation **in** $S$.  

---

40
Example 18 Suppose $\varphi : S \to T$ is a map. Then the relation 
\[ \{ (x, \varphi(x)) \mid x \in S \} \subset S \times T \]
is called the relation of the map $\varphi$.

Thus relations corresponding to set maps are those subsets of $S \times T$ that satisfy the following property: every $x \in S$ is related to a single $y \in T$.

By contrast, any subset $R \subset S \times T$ is a relation between $S$ and $T$.

Exercise 8 Relations

1. Identify whether which of the following subsets $R \subset S \times T$ is a relation determined by a map:
   
   (a) $S$ is all of $\mathbb{R}$ and $T$ is all of $\mathbb{Z}$, and $R = \{(2, -3), (3, -4), (4, -5)\}$.
   
   (b) $S = \mathbb{N}$, $T = \mathbb{R}$ and $R = \{(n, 1/n) \mid n \in \mathbb{N}\}$.
   
   (c) $S = \mathbb{N}$, $T = \mathbb{N}$ and $R = \{(n^2, 1/n) \mid n \in \mathbb{N}\}$.

2. Show that the relation $\{(x^2, x) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2$ is not the relation of a map.

3. Which are the integers $n \in \mathbb{N}$ such that the relation $\{(x^n, x) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2$ is the relation of a map from $\mathbb{R}$ to $\mathbb{R}$?

4. If $S$ and $T$ are finite sets, how many relations are there between $S$ and $T$? Compare this with the number of maps from $S$ to $T$ — see Exercise 5.

Perhaps you have heard the expression,

You can’t see the forest for the trees.

The speaker is accusing the audience that they are so bound up in the details of the trees that they don’t come to grips with the more global properties of the forests. And in fact, sometimes instead of looking at individual trees, we want to look at the forests as individuals. Similarly, instead of considering individual people we might want to talk about cities, which are collections of people. We might want to describe the properties of species, which are collections of animals, rather than the distinctions between individual animals.

This leads to a very important mathematical idea, which we formalize in the following way:

Definition 17 A partition of a set $S$ is a family of non-empty subsets $S_i \subset S$ such that every element $x \in S$ belongs to exactly one subset $S_i$.

Every partition $\{S_i\}$ of a set $S$ determines a specific relation in $S \times S$; namely, we set $xRy$ if and only if $x$ and $y$ are in the same subset $S_i$. This is called the relation of the partition.
Example 19

1. The single set \( S \) is a partition of a non-empty set \( S \). The corresponding relation is \( xRy \) for every \( x, y \in S \). Thus \( R = S \times S \).

2. The family of subsets \( S_x, x \in S \) of a non-empty set \( S \) is a partition of \( S \), since every \( x \) belongs to a single \( S_x \). The corresponding relation is \( xRy \) if and only if \( x = y \).

3. A partition of the set, \( S \), of sophomore students at the University of Maryland is defined by: For each possible GPA, \( \alpha \), \( S_{\alpha} \) is the set of students whose GPA is \( \alpha \).

Lemma 3 The relation of a partition \( \{S_i\} \) of a set \( S \) satisfies the following properties:

1. For every \( x \in S \), \( xRx \). (The relation is reflexive.)

2. If \( xRy \) then also \( yRx \). (The relation is symmetric.)

3. If \( xRy \) and \( yRz \) then \( xRz \). (The relation is transitive.)

Proof:

1. The relation is reflexive because \( x \) and \( x \) belong to the same \( S_i \).

2. If \( xRy \) then by definition \( x \) and \( y \) are in the same \( S_i \) and so \( yRx \). Therefore the relation is symmetric.

3. We have to show that if \( xRy \) and \( yRz \) then \( xRz \).

   By hypothesis, \( x \) belongs to exactly one \( S_i \).
   Since \( xRy \), \( y \) must also be in that \( S_i \).
   Since \( yRz \), \( z \) must also be in that \( S_i \).
   Thus \( y \) and \( z \) are in the same \( S_i \).
   Therefore by definition, \( xRz \) and the relation is transitive.

q.e.d.

Definition 18 An equivalence relation in a set \( S \) is a reflexive, symmetric, and transitive relation. Equivalence relations are denoted by \( x \sim y \).

Lemma 3 states that the relation of a partition is an equivalence relation. Conversely we have

Proposition 6 Every equivalence relation, \( \sim \), in a set \( S \) is the relation of a unique partition of \( S \).
**Proof:** Fix an equivalence relation, \(\sim\), in \(S\).

For each \(x \in S\) define a subset \(S(x) \subseteq S\) by

\[
S(x) = \{ y \in S \mid y \sim x \}.
\]

Now let \(\{S_i\}\) denote the family of distinct subsets of \(S\) such that each \(S_i = S(x)\) for some \(x \in S\).

We will prove three things:

1. This is a partition of \(S\).
2. \(\sim\) is the relation of the partition.
3. This is the only partition with \(\sim\) as its relation.

**Step 1.** This is a partition of \(S\).

By reflexivity each \(x \in S(x)\) and so the \(S_i\) are not empty and every \(x\) in \(S\) belongs to some \(S_i\).

Thus to show this is a partition we have to show that any element of \(S\) can belong to only one subset \(S_i\).

In other words we have to show that if \(y \in S(x)\) and \(y \in S(z)\) then \(S(x) = S(z)\).

We first prove that for any \(x \in S\), if \(u \in S(x)\) then

\[
S(u) \subseteq S(x).
\]

In fact, since \(u \in S(x)\) we have \(u \sim x\), and if \(v \in S(u)\) we have \(v \sim u\). Thus \(v \sim u \sim x\), and since the relation is transitive, \(v \sim x\).

It follows that \(v \in S(x)\); i.e. \(S(u) \subseteq S(x)\).

Next we prove that if \(u \in S(x)\) then

\[
S(x) \subseteq S(u).
\]

In fact, since \(u \sim x\) and the relation is symmetric, \(x \sim u\).

Therefore \(x \in S(u)\) and by what we have just shown it follows that

\[
S(x) \subseteq S(u).
\]

Altogether we have proved that if \(u \in S(x)\) then \(S(x) = S(u)\).

Finally, if \(y \in S(x)\) and \(y \in S(z)\), then by the equality we just proved

\[
S(y) = S(x) = S(z).
\]
Therefore \( \{S_i\} \) is a partition and Step 1 is proved.

**Step 2.** The equivalence relation \( \sim \) is the relation of the partition.

Denote the relation of the partition by \( R \), so that \( xRy \) if and only if \( x \) and \( y \) are in the same subset \( S_i \) of the partition.

Then by definition,

\[
xRy \iff x, y \in S(w), \text{ some } w \in S.
\]

But we know that if \( x, y \in S(w) \) then \( S(x) = S(w) = S(y) \) and so \( x \sim y \). Thus \( R = \sim \).

**Step 3.** The partition \( \{S_i\} \) is the unique partition which has \( \sim \) as its equivalence relation.

Let \( \{T_\alpha\} \) be a partition of \( S \) with \( \sim \) as its equivalence relation.

If \( x \in S \) then the subset \( T_\alpha \) in the partition containing \( x \) will satisfy

\[
y \sim x \iff y \in T_\alpha.
\]

Thus \( T_\alpha = S(x) \).

It follows that every \( T_\alpha \) is one of the \( S(x) \).

Since every element of \( S \) is in some \( T_\alpha \) it follows that every \( S(x) \) is one of the \( T_\alpha \).

Therefore the partition \( \{T_\alpha\} \) is the partition \( \{S_i\} \). **q.e.d.**

**Definition 19** If \( \sim \) is an equivalence relation in a set \( S \) then the elements \( S_i \) of the corresponding partition are called the **equivalence classes** of the relation.

**Exercise 9 Equivalence relations**

1. Provide a complete proof for the statement in Example ???.3.

2. Two UM students are related if they both take at least one class in common. Is this an equivalence relation?

3. Two UM students are related if they take all their classes in common. Is this an equivalence relation?

4. Suppose \( f : S \to T \) is a map between two sets. Say \( x \) and \( y \) in \( S \) are related if \( f(x) = f(y) \). Is this an equivalence relation?

5. Suppose \( S = \bigcup_{i=1}^n S_i \) is a partition of a set with \( n \) elements for some natural number \( n \).
(a) If each $S_i$ has the same number of elements, $k$, show that $k$ divides $n$ and that $n/k = m$.

(b) If each $S_i$ has $i$ elements find $n$.

(c) If each $S_i$ has $i^3$ elements, find $n$.

(d) If each $S_i$ has an odd number of elements show that $m$ is even if and only if $n$ is even.

6. Recall that $P(Q)$ denotes the set of all subsets of $Q$. Thus the elements of $P(Q)$ are the subsets $S \subset Q$. Define a relation

$$\alpha \subset P(Q) \times P(Q)$$

by setting $S \alpha T$ if

(a) for every $a \in S$ there is some $b \in T$ such that $b \geq a$, and
(b) for every $c \in T$ there is some $d \in S$ such that $d \geq c$.

Then

(a) Show that $\alpha$ is an equivalence relation.
(b) Show that if $\{S_i\}$ is an equivalence class of subsets then

$$\bigcup_i S_i$$

is an element in that equivalence class.
(c) Show that if $S \subset Q$ is not equivalent to $Q$ then for some rational number, $a$, $a > x$ for all $x \in S$.
(d) If $a \in Q$ are the sets

$$S = \{b \in Q \mid b < a\} \quad \text{and} \quad T = \{b \in Q \mid b \leq a\}$$

equivalent?

7. Let $T^S$ denote the set of maps from a set $S$ to a set $T$. Define a relation $R \subset T^S \times T^S$ by setting

$$fRg$$

if $f(x) = g(x)$ except for finitely many points $x \in S$ (depending, of course, on $f$ and $g$). Show this is an equivalence relation.
Chapter 4

The Real Numbers

The bedrock of analysis is an understanding of the real numbers. Notice the difference between the rational numbers and the real numbers: we can describe the rational numbers explicitly as quotients of one integer by another. No such simple expression is available for the reals. We may have an intuitive understanding of these as the points on the real number line, as distances, or as (possibly) infinite decimals, but we cannot use "intuitive understandings" to make rigorous proofs.

We solve this problem by showing that a real number can be approximated arbitrarily well by a rational number. Explicitly:

If $x \in \mathbb{R}$ then for any $\delta > 0$ there is a rational number $a \in \mathbb{Q}$ such that $|x - a| < \delta$.

This allows us to extend properties of the rationals to the reals in a rigorous way.

The concept of "approximation arbitrarily well" is very important throughout the rest of this course. Thus, more generally, if $S \subset \mathbb{R}$ is any subset of the reals we say that $x \in \mathbb{R}$ can be approximated arbitrarily well by numbers in $S$ if for every $\delta > 0$ there is a number $y \in S$ such that $|x - y| < \delta$.

Our first step is to identify the basic properties of the rationals.

4.1 Properties of the Rationals

Here are the two fundamental concepts for the rationals with which we are familiar:
1. **Algebra:** The operations of addition, subtraction, multiplication and division.

2. **Order:** If $b$ and $a$ are rational numbers then
   \[ a < b \iff b - a = m/n \text{ for some } m, n \in \mathbb{N}. \]

**Notation:** We use $a > b$ to mean the same thing as $b < a$.

The basic properties of the ordering in the rationals are contained in the next lemma. While they are utterly and totally familiar we shall give formal proofs to provide more examples of what a proof looks like.

**Lemma 4**

1. For each rational number $a \in Q$ exactly one of the following three possibilities is true:
   
   (a) $a > 0$
   (b) $a = 0$
   (c) $a < 0$

   In particular, if $a, b \in Q$ then
   \[ b > a \iff b - a > 0. \]

2. If $a, b, c$ are any three rational numbers, then:

   (a) If $a < b$ and $b < c$ then $a < c$.
   (b) $a + b < a + c$ if and only if $b < c$.
   (c) If $c > 0$ and $a < b$, then $ca < cb$.
   (d) If $a < b$ then $-a > -b$.

3. If $a \in Q$ then $a > 0 \iff a > 1/n$ for some $n \in \mathbb{N}$.

**Proof:** We prove each statement separately.

1. By definition $a = m/n$ with $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. Thus exactly one of the following three possibilities is true:
   
   (a) $m \in \mathbb{N}$,
   (b) $m = 0$, or
   (c) $-m \in \mathbb{N}$.

   In the first case $a > 0$.
   In the second case, $a = 0$.
   In the third case $a < 0$.
2. (a) If \( a < b \) and \( b < c \) then

\[
b - a = p/q \quad \text{and} \quad c - b = m/n,
\]

with \( p, q, m, n \in \mathbb{N} \).

It follows that

\[
c - a = (c - b) + (b - a) = m/n + p/q > 0.
\]

(b) Since \( (a + c) - (a + b) = c - b \) it follows that

\[(a + c) - (a + b) > 0 \iff c - b > 0.\]

(c) Since \( c > 0 \) and \( b - a > 0 \),

\[
c = p/q \quad \text{and} \quad b - a = m/n \quad \text{for some} \quad p, q, m, n \in \mathbb{N}.
\]

Therefore

\[
(c - b)a = p/qn > 0.
\]

(d) Since \( a < b \),

\[
-a - (-b) = (b - a) > 0.
\]

Now adding \(-b\) to both sides gives (by Lemma 4.2(b)) that \(-a < -b\).

3. If \( a > 0 \), we have for some \( m, n \in \mathbb{N} \) that

\[
a = m/n > 1/n.
\]

Conversely, for any \( n \in \mathbb{N} \), \( 1/n > 0 \).

Thus if \( a > 1/n \) then \( a > 1/n > 0 \).

q.e.d.

4.2 Introducing the Reals

Recall that while we may have an intuitive understanding of the real numbers as the points on the x-axis of the plane, or as distances, or as (possibly) infinite decimals, we cannot use "intuitive understandings" to make rigorous proofs. Thus we introduce the reals and list certain elementary properties as axioms, from which all the other properties will be deduced. In this way we rigorously extend properties of the rationals to the reals.

**Property One:** The real numbers are a set, denoted by \( \mathbb{R} \), containing the rational numbers.
Property Two (Algebra): The operations of addition, subtraction, multiplication and division are defined for real numbers, coinciding with the old operations in the rationals, and with the same properties.

Property Three (Order): The ordering of the rationals extends to an ordering of the real numbers such that

1. For each real number $x$ exactly one of the following three possibilities is true:
   (a) $x > 0$, in which case we say $x$ is positive.
   (b) $x = 0$.
   (c) $x < 0$, in which case we say $x$ is negative.
2. $x < y \iff y - x > 0$.
3. If $x < y$ and $y < z$ then $x < z$.
4. The product of positive real numbers is positive.
5. If $x \in \mathbb{R}$ then there is a natural number $m$ such that
   $$x < m.$$ 

Notation: We shall write $y > x$ to mean $x < y$.

Lemma 5

1. If $x, y, z$ are any three real numbers, then:
   (a) $x + y < x + z \iff y < z$.
   (b) $x > 0 \iff -x < 0$. In this case $1/x > 0$.
   (c) Multiplication by a positive real number preserves inequalities:
      $$z > 0 \quad \text{and} \quad x < y \Rightarrow zx <zy.$$
   (d) If $0 < x < y$ then $1/y < 1/x$.
   (e) Multiplication by $-1$ reverses inequalities, and therefore multiplication by any negative real number reverses inequalities:
      $$z < 0 \quad \text{and} \quad x < y \Rightarrow zx > zy.$$

2. For each $x \in \mathbb{R}$ there is a natural number $n$ such that
   $$-n < x.$$
3. For each positive real number $x$ there are natural numbers $k, m$ such that $\frac{1}{k} < x < m$.

**Proof:** We prove each statement separately.

1. Suppose $x, y, z \in \mathbb{R}$.
   
   (a) Since $(x + z) - (x + y) = z - y$ it follows from Property Three that $x + y < x + z \Leftrightarrow y < z$.
   
   (b) If $0 < x$ then $-x + 0 \leq -x + x$ so that $-x = -x + 0 < -x + x = 0$.
   
   Conversely, if $-x < 0$, then adding $x$ to both sides gives $0 < x$.
   
   Finally, we show by contradiction that $x > 0 \Rightarrow \frac{1}{x} > 0$.
   
   In fact, suppose $\frac{1}{x} \leq 0$. Then by what we just proved, $-1/x \geq 0$.
   
   Since the product of positives is positive (Property 3) it would follow that
   
   $$-1 = (-1/x)x \geq 0,$$
   
   which is false.
   
   It follows by contradiction that it is not true that $1/x \leq 0$.
   
   Thus $1/x > 0$.
   
   (c) Since $x < y$ Property 3 states that $y - x > 0$.
   
   Since $z > 0$ and the product of positive real numbers is positive (Property 3), it follows that $zy - zx = z(y - x) > 0$.
   
   (d) First,
   
   $1/x > 0 \text{ and } 1/y > 0 \Rightarrow 1/x \text{ and } 1/y > 0 \Rightarrow (1/x)(1/y) > 0 \Rightarrow 1/xy > 0$.
   
   Multiplication by $1/xy$ therefore preserves the inequality $x < y$ and this gives $1/y < 1/x$.
   
   (e) First, if $x < y$ then
   
   $$-y = x + (-x - y) < y + (-x - y) = -x,$$
   
   and so multiplication by $-1$ reverses inequalities. Thus in general, if $z < 0$ and $x < y$ then we have $-z > 0$ and so by Property 3,
   
   $$-(zx) = (-z)x < (-z)y = -(zy).$$
   
   Multiplication by $-1$ reverses this inequality and gives $zx > zy$. 50
2. By Property 3 applied to $-x$, for some $n \in \mathbb{N}$ we have $-x < n$. Since multiplication by $-1$ reverses inequalities it follows that 

$$-n < x.$$ 

3. Since $x > 0$, $1/x > 0$. Thus by Property Three there are natural numbers $k, m \in \mathbb{N}$ such that 

$$1/x < k \quad \text{and} \quad x < m.$$ 

By what we proved above this implies that $1/k < x$ and so 

$$1/k < x < m.$$ 

q.e.d.

The preceding Lemma establishes properties you have been taking for granted for years. Henceforth we will use these properties without referring to the Lemma for justification.

The next result is absolutely fundamental for this course. It states that any real number can be approximated arbitrarily well by rational numbers.

**Proposition 7** *(Sandwich theorem)* Suppose $x \in \mathbb{R}$ and $k \in \mathbb{N}$.

1. Then there are rational numbers $a, b$ such that 

$$a < x < b \quad \text{and} \quad b - a < 1/k.$$ 

2. In particular, 

$$0 < x - a < 1/k \quad \text{and} \quad 0 < b - x < 1/k.$$ 

**Proof:**

1. By Property Three and Lemma 5 there are natural numbers $i, j$ such that 

$$-j < x < i.$$ 

Set $p = i(2k + 1)$ and $q = j(2k + 1)$. Then we may rewrite this inequality as 

$$\frac{-q}{2k + 1} < x < \frac{p}{2k + 1}.$$ 

Observe first that if $m \mathbb{Z}$ and $x < m/(2k + 1)$ then 

$$\frac{-q}{2k + 1} < x < \frac{m}{2k + 1},$$ 

and so $-q < m.$
Therefore there is a least integer \( r \) such that

\[ x < \frac{r}{2k+1}. \]

Then

\[ x \geq \frac{r-1}{2k+1} \quad \text{and so} \quad x > \frac{r-2}{2k+1}. \]

Thus

\[ \frac{r-2}{2k+1} < x < \frac{r}{2k+1} \]

and

\[ \frac{r}{2k+1} - \frac{r-2}{2k+1} = \frac{2}{2k+1} < 1/k. \]

2. This is immediate from what was just proved.

q.e.d.

**Proposition 8** If \( x < y \) are real numbers then for some \( b \in \mathbb{Q} \),

\[ x < b < y. \]

**Proof:** By Lemma 5, for some \( k \in \mathbb{N}, 1/k < y - x \). and so \( x + 1/k < y \). Now by the Sandwich theorem, for some \( b \in \mathbb{Q} \),

\[ x < b \quad \text{and} \quad b - x < 1/k. \]

Thus \( b < x + 1/k \) and by choice of \( k, x + 1/k < y \). q.e.d.

**Proposition 9** If \( x > 0 \) and \( y > 1 \) are real numbers then for some \( p \in \mathbb{N} \),

\[ 1/y^p < x. \]

**Proof:** Since \( y > 1 \), by Lemma 5, for some \( k \in \mathbb{N}, y - 1 > 1/k \). Therefore

\[ y > (1/k) + 1. \]

Now apply the Binomial theorem (Proposition 2) to obtain for any \( p \in \mathbb{N} \) that

\[ \left( (1/k) + 1 \right)^p = 1 + p/k + \sum_{i=2}^{p} \binom{p}{i} (1/k)^i > p/k. \]

Moreover, by the Difference theorem (Proposition 1),

\[ y^p - (1/k + 1)^p = (y - (1/k + 1)) \sum_{i=0}^{p-1} y^i (1/k + 1)^{p-1-i} > 0. \]

Therefore

\[ y^p > (1/k + 1)^p > p/k. \]
Now multiply by \( k/py \) to obtain
\[
k/p > 1/y^p.
\]
This holds for any \( p \in \mathbb{N} \).
Finally, by Lemma 5, \( x > 1/n \) for some \( n \in \mathbb{N} \). Now choose \( p = kn \). Then
\[
1/y^p < k/kn = 1/n.
\]
q.e.d.

**Exercise 10 Inequalities**

1. Is it true that there are rational numbers \( a < b \) such that for any natural number \( k, b − a < 1/k \)?
2. If \( x, y, z, w \) are positive real numbers such that \( x < y \) and \( z < w \) show that \( xz < yw \). State the converse and either prove it or provide a counterexample to show it is false.
3. Show that if \( x < y \) are real numbers then there are infinitely many rational numbers \( b \) such that \( x < b < y \).
4. Suppose \( c \in \mathbb{Q} \) and \( y, z \) are any real numbers. Show that if \( c > y + z \) then there are rational numbers \( a > y \) and \( b > z \) such that
\[
c > a + b.
\]
5. Show that the product of a positive real number with a negative real number is negative.
6. Let \( x \) and \( y \) be positive real numbers. Show that if \( x > 1 \) then (i) \( xy > y \), (ii) \( 0 < 1/x < 1 \), and (iii) \( 0 < y/x < y \).
7. If \( a < b \) are non-negative real numbers show that for any \( p \in \mathbb{N} \), \( b^p − a^p ≤ p(b − a)b^{p−1} \). (Hint: use the Difference theorem (Proposition 1).)

### 4.3 Absolute value

Recall that the **absolute value** of a real number \( x \) is denoted by \(|x|\) and is defined by:
\[
|x| = x \text{ if } x \geq 0 \text{ and } |x| = −x \text{ if } x < 0.
\]

**Note:**

1. \( −x \) is positive when \( x \) is negative!
2. For every \( x \in \mathbb{R} \):
\[
|x| ≥ 0
\]
and
\[
|x| = 0 ⇔ x = 0.
\]
3. $|x|$ is the larger of $x$ and $-x$.

4. Thus the statement $|x| < \varepsilon$ is the same as $x < \varepsilon$ and $-x < \varepsilon$, and hence $|x| < \varepsilon \Leftrightarrow -\varepsilon < x < \varepsilon$.

**Lemma 6** If $x, y$ are real numbers such that $|x - y| < 1/n$ for every $n \in \mathbb{N}$, then $x = y$.

**Proof:** We prove this by contradiction. Suppose the conclusion $x = y$ is false. Then since $x \neq y$, it follows that $x - y \neq 0$. Therefore $|x - y| > 0$. Now by Property 3, $|x - y| > 1/n$ for some $n \in \mathbb{N}$. This contradicts the statement $|x - y| < 1/n$ for every $n \in \mathbb{N}$. q.e.d.

**Remark:** Lemma 6 illustrates the difference between algebra and analysis. In algebra we show equality directly. In analysis we sometimes show two numbers are the same by showing that their difference is arbitrarily small.

**Lemma 7** For any real numbers $x$ and $y$:

1. $|xy| = |x||y|$.
2. $x^2 \geq 0$ and so $|x^2| = x^2$.
3. $|x + y| \leq |x| + |y|$ (The *triangle inequality*).
4. $||x| - |y|| \leq |x - y|$.

**Proof:** We prove each statement separately.

1. $|xy|$ and $|x||y|$ are both non-negative numbers equal either to $xy$ or to $-xy$. Therefore in this case they are equal.

2. If $x \geq 0$ then this follows because the product of positives is positive. If $x < 0$ then $x = -|x|$ and $|x| > 0$. Thus $x^2 = |x|^2 > 0$.

3. Note that $|x + y|^2 = (x + y)^2 = x^2 + 2xy + y^2 = |x|^2 + 2xy + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2$. 54
Therefore $|x + y|^2 \leq (|x| + |y|)^2$. It follows that

$$0 \leq ((|x| + |y|)^2 - |x + y|^2 = (|x| + |y| - |x + y|)(|x| + |y| - |x + y|)).$$

Since the second factor on the right is not negative it follows that

$$|x| + |y| - |x + y| \geq 0.$$

4. By the triangle inequality,

$$|x - y| + |y| \geq |x - y + y| = |x| \geq |x|,$$

and

$$|x - y| + |-x| \geq |x - y - x| = |y|.$$

Therefore

$$|x - y| \geq |x| - |y| \text{ and } |x - y| \geq |y| - |x| = |y| - |x|$$

Therefore

$$|x - y| \geq ||x| - |y||.$$

q.e.d.

The next Proposition is also fundamental for this course.

**Proposition 10** Suppose $a \in \mathbb{R}$ and $\varepsilon \in \mathbb{R}$ is positive. Then for $x \in \mathbb{R}$

$$|x - a| < \varepsilon \iff a - \varepsilon < x < a + \varepsilon.$$

**Proof:** The statement $|x - a| < \varepsilon$ is the same as

$$x - a < \varepsilon \text{ and } a - x < \varepsilon.$$

But

$$x - a < \varepsilon \iff x < a + \varepsilon \text{ and } a - x < \varepsilon \iff x > a - \varepsilon.$$

Thus

$$|x - a| < \varepsilon \iff a - \varepsilon < x < a + \varepsilon.$$

q.e.d.

**Exercise 11 Absolute value**

1. If $x$ is a positive real number show that for some $\varepsilon > 0$,

$$y \in \mathbb{R} \text{ and } |x - y| < \varepsilon \Rightarrow y > 0.$$
2. If \( x, z \in \mathbb{R} \) show that for each \( \varepsilon > 0 \) there is some \( \delta > 0 \) such that if \( y \in \mathbb{R} \) satisfies \( |y - x| < \delta \) then \( |zy - zx| < \varepsilon \).

3. If \( x \in \mathbb{R} \) show that for each \( \varepsilon > 0 \) there is some \( \delta > 0 \) such that if \( y \in \mathbb{R} \) satisfies \( |y - x| < \delta \) then \( |y^2 - x^2| < \varepsilon \). (Hint: Use \( y^2 - x^2 = (y-x)(y+x) \).)

4. If \( x \in \mathbb{R} \) and \( x \neq 0 \) show that for each \( \varepsilon > 0 \) there is some \( \delta > 0 \) such that if \( y \in \mathbb{R} \) satisfies \( |y - x| < \delta \) then \( y \neq 0 \) and \( |1/y - 1/x| < \varepsilon \).

### 4.4 Bounds

**Definition 20 Bounds**

1. A nonvoid subset \( S \subset \mathbb{R} \) is **bounded above** if for some \( b \in \mathbb{R} \),
   \[
   x \leq b \quad \text{for all} \quad x \in S.
   \]
   In this case we write \( S \leq b \) or \( b \geq S \),
   and say \( b \) is an **upper bound** for \( S \).

2. A nonvoid subset \( S \subset \mathbb{R} \) is **bounded below** if for some \( a \in \mathbb{R} \),
   \[
   a \leq x \quad \text{for all} \quad x \in S.
   \]
   In this case we write \( a \leq S \) or \( S \geq a \),
   and say \( a \) is a **lower bound** for \( S \).

3. A subset \( S \subset \mathbb{R} \) is **bounded** if it is both bounded above and bounded below.

**Lemma 8** Suppose \( S \) and \( T \) are non-void subsets of \( \mathbb{R} \). Then every element of \( S \) is a lower bound for \( T \) if and only if every element of \( T \) is an upper bound for \( S \).

In this case we write \( S \leq T \) or \( T \geq S \).

**Proof:** The statement that every element of \( S \) is a lower bound for \( T \) is true if and only if \( x \leq T \) for every \( x \in S \).

This is equivalent to the statement:

\[
  x \leq y \quad \text{for all} \quad x \in S \quad \text{and all} \quad y \in T.
\]

But this is equivalent to saying that every \( y \in T \) is an upper bound for \( S \).

q.e.d.

**Example 20 Bounds**
1. Any real number $x \leq 1$ is a lower bound for $\mathbb{N}$.

2. $\mathbb{N}$ is not bounded above.
   In fact if $x$ is any real number then by Property 3.5 there is some natural number $m$ such that $x < m$.
   Thus $x$ is not an upper bound for $\mathbb{N}$, and so $\mathbb{N}$ does not have an upper bound.

3. $S = \{1/n \mid n \in \mathbb{N}\}$ is bounded above by any number $\geq 1$ and bounded below by any number $\leq 0$.

4. The rationals are not bounded below or above.

5. The set $S$ of real numbers whose squares are less than 2 is bounded above by 2 and below by $-2$.
   In fact, if $x > 2$ then $x^2 = xx > 2x > 4$ and so $x \notin S$.
   Also, if $x < -2$ then $-x > 2$ and $x^2 = xx = (-x)(-x) > 4$ and again, $x \notin S$.
   Thus
   $$-2 \leq S \leq 2.$$

Exercise 12 Bounds

1. Show that if $x$ is a positive real number then $x + 1$ is not a lower bound for $\mathbb{N}$.

2. Show that if $b$ is an upper bound for a non-void set $S \subset \mathbb{R}$ then every $c > b$ is also an upper bound for $S$.

3. Give an example of a non-void bounded set $S \subset \mathbb{R}$ which contains an upper bound but does not contain a lower bound.

4. Show that a non-void set $S \subset \mathbb{R}$ cannot contain two different upper bounds.

5. Show that the set of real numbers of the form $x/y$ with $|x| > |y| > 0$ is not bounded above or below.

6. Show that the set of real numbers of the form $x/y$ with $|y| > |x| > 0$ is bounded. Find explicit upper and lower bounds and prove your answers are correct.

7. Suppose $S \subset T \subset \mathbb{R}$ are non-void sets. If $S \leq T$ show that $S$ has exactly one element.

8. More generally, suppose $S \subset \mathbb{R}$ and $T \subset \mathbb{R}$ are non-void sets. If $S \leq T$ show that $S \cap T$ can have at most one element. If $S \cap T$ is non void show that its unique element is an upper bound for $S$ and a lower bound for $T$.  

57
4.5 Least upper and greatest lower bounds

In everything we have done so far there is nothing to suggest that there are real numbers which are not rationals!

All we have assumed about the real numbers at this point is listed in Properties One, Two, and Three at the start of Section 4.2.

We correct this with one final property, still as an axiom. This is the fundamental property of the real numbers and is the key fact which makes analysis possible.

**Property Four:** Let $S$ be a non empty subset of $\mathbb{R}$ which is bounded above. Then $S$ has an upper bound, $b$ with the following property:

If $y$ is any upper bound for $S$ then $b \leq y$.

The upper bound $b$ in Property Four is unique.

**Proof:** Suppose $c$ satisfies the same condition as $b$. Then $c$ is an upper bound for $S$. Therefore $b \leq c$. But $b$ is an upper bound for $S$. Therefore $c \leq b$. Therefore $b = c$.

q.e.d.

**Definition 21** The real number $b$ in Property Four is called the least upper bound for $S$ and is denoted by $\text{lub}(S)$ or by $\text{sup}(S)$.

**Important Remark:** If $S$ is bounded above, sometimes $\text{lub}(S)$ is in $S$ and sometimes it is not! For example if

$$S = \{x \in \mathbb{R} \mid x \leq 1\}$$

or if

$$S = \{x \in \mathbb{R} \mid x < 1\},$$

then in both cases $\text{lub}(S) = 1$. However, in the first case, $1 \in S$ and in the second case $1 \notin S$.

Analogous to least upper bounds, a non empty set which is bounded below has a greatest lower bound:

**Proposition 11** Suppose $S$ is a non empty subset of $\mathbb{R}$ which is bounded below. Then there is a unique lower bound, $a$ with the following property:

If $y$ is any lower bound for $S$ then $y \leq a$. 

58
Definition 22 The real number $a$ in Proposition 7 is called the greatest lower bound for $S$ and is denoted by $\text{glb}(S)$ or by $\text{inf}(S)$.

Proof of Proposition 11: Let $-S = \{-x \mid x \in S\}$.
If $y$ is a lower bound for $S$ then $y \leq S$.
Since multiplication by $-1$ reverses inequalities, $-y \geq -S$.
Therefore $-S$ is bounded above, and so by Property Four, it has a least upper bound, $b$.
Set $a = -b$.
Since $b \geq -S$, it follows that $a = -b \leq S$.

Now let $y$ be any lower bound for $S$. Then

$$-y \geq -S.$$ 

Thus $-y$ is an upper bound for $-S$ and so $-y \geq b$.
Therefore $y \leq -b = a$ and so $a = \text{glb}(S)$. Its uniqueness is proved in the exact same way Lemma 9 is proved.

q.e.d.

We recap what it means for $b$ to be the least upper bound of a non-void set $S$. It means exactly this:

$$S \leq b$$ 

and

If $u < b$ then $u$ is not an upper bound for $S$.

A useful way of expressing this is given in the next Proposition:

Proposition 12 Suppose $S$ is a non-void subset of $R$. Then

1. $b$ is the least upper bound of $S$ if and only if
   
   (a) $S \leq b$, and
   
   (b) for all $\varepsilon > 0$ there is an $x \in S$ such that $|b - x| < \varepsilon$.

2. $a$ is the greatest lower bound of $S$ if and only if
   
   (a) $S \geq a$, and
   
   (b) for all $\varepsilon > 0$ there is an $x \in S$ such that $|a - x| < \varepsilon$.

Proof: For the first assertion, suppose the two conditions above are satisfied.
Then $b$ is an upper bound.
Moreover, if $u < b$, then $b - u > 0$. Therefore by hypothesis, for some $x \in$
S, |b - x| < b - u.
Since b - x > 0 we have

\[ b - x = |b - x| < b - u, \]

and so u < x and u is not an upper bound for S.

Conversely, suppose b is the least upper bound for S. Then b ≥ S.
Moreover, if ε > 0, then b - ε < b and so b - ε is not an upper bound for S.
Therefore, for some x ∈ S, b - ε < x. It follows that |b - x| = b - x < ε.

The second assertion is proved in the same way.

q.e.d.

Example 21

1. glb(N) = 1.
   In fact, since 1 ∈ N, any lower bound x for N satisfies x ≤ 1
   But 1 ≤ n for all n ∈ N. Thus 1 is a lower bound for N and so it is the
   greatest lower bound: glb(N) = 1.

2. Let S = {(x + 1)/x | x ∈ R and x ≥ 1}. Then 1 = glb(S).
   In fact for x ≥ 1
   \[ 1 ≤ 1 + 1/x. \]
   Moreover, if y > 1 then by Property 3 there is some n ∈ N with 1/n < y - 1. Thus
   \[ 1 + 1/n < y \]
   and so y is not a lower bound for S. Thus 1 is the greatest lower bound.

Exercise 13 glb and lub

1. If b ∈ S is an upper bound for a non-void set S ⊆ R show that it is the
   least upper bound.
2. Show that 0 is the greatest lower bound for {1/n | n ∈ N}.
3. Let a ∈ R and let S = {x ∈ R | |x - a| < 1/2}. Show that S is bounded.
   Then find lub(S) and glb(S) and prove your answers are correct.
4. Let S be a a non-void subset of R which is bounded above. Show that the
   following conditions are equivalent on an upper bound, b, for S:
   (a) b = lub(S).
   (b) For each ε > 0 there is an element x ∈ S such that b - ε < x ≤ b.
(c) For each $\varepsilon > 0$ there is an element $x \in S$ such that $0 \leq b - x < \varepsilon$.

5. Suppose $S \subset T$ are subsets of $R$:

(a) If $T \subset R$ is a non-void set bounded below show that $S$ is bounded below and that $\text{glb}(T) \leq \text{glb}(S)$;

(b) If $T$ is bounded above show that $S$ is bounded above and that $\text{lub}(T) \geq \text{lub}(S)$.

4.6 Powers

We begin by recalling from Sec. 2.3 the definition of integral powers of a real number $x$, which we do by induction.

**Definition 23** Let $x$ be a non-zero real number. If $n \in \mathbb{N}$ then $x^n$ is defined inductively by

$$x^1 = x \quad \text{and} \quad x^{n+1} = x \cdot x^n.$$  

Then we set $x^0 = 1$ and $x^{-n} = 1/x^n$.

Our objective here is to extend the definition to rational powers of a positive real number, $x$. To do this we first need to prove:

**Proposition 13** Let $S = \{x \in R \mid x > 0\}$ be the set of positive real numbers. Then for any natural number $n \in \mathbb{N}$:

1. For $x, y \in S$,

   $$x < y \iff x^n < y^n.$$  

2. The map

   $$S \to S, \quad x \mapsto x^n$$

   is a bijection.

**Proof:**

1. The Difference theorem (Proposition 1) gives

   $$y^n - x^n = (y - x) \sum_{i=0}^{n-1} y^i x^{n-1-i}.$$  

   Since both $x$ and $y$ are positive, $\sum_{i=0}^{n-1} y^i x^{n-1-i}$ is positive. Therefore

   $$y^n - x^n > 0 \iff y - x > 0.$$  

2. We first observe that because of the assertion above, the map is 1-1.

   Now we prove that our map is onto.
We need to show that if $y$ is a positive real number and $n \in \mathbb{N}$ then there is a positive real number $x$ such that $x^n = y$.

For this define a set $T$ by

$$T = \{ x \in S \mid x^n > y \}.$$  

Observe that $T \neq \emptyset$. In fact, since $y > 0$, it follows that $y + 1 > 1$ and so $(y + 1)^n \geq y + 1 > y$. Thus $y + 1 \in T$.

On the other hand, $0 \leq T$ and so $T$ is bounded below. By Proposition 11, $T$ has a greatest lower bound, $\text{glb}(T)$. Write $x = \text{glb}(T)$. We show that $x^n = y$.

By Property 3.1, exactly one of the following three possibilities must hold:

$$x^n < y, \ x^n > y, \text{ or } x^n = y.$$  

To prove our assertion we show by contradiction that the two inequalities above are false, and so it must be true that $x^n = y$. Our strategy is as follows.

(a) We show that if $x^n < y$ then $x$ is not the greatest lower bound of $T$, which is a contradiction.

To do this, we will show that if $x^n < y$ then there is a positive real number $c$ such that $x^n < c^n < y$.

Indeed, if such a $c$ exists then for $z \in T$,

$$c^n < y < z^n,$$

and so by the first assertion of the Proposition,

$$c < z \quad \text{for all} \quad z \in T.$$  

Thus $c$ is a lower bound for $T$. But since $x^n < c^n$ we have $x < c$ and so $x$ would not be the greatest lower bound.

(b) We show that if $x^n > y$ then $x$ is not a lower bound of $T$, which is also a contradiction.

To do this, we will show that if $x^n > y$ then there is a positive real number $b$ such that $y < b^n < x^n$.
But if such a $b$ exists then by the definition of $T$, $b \in T$. Now by the first assertion of the Proposition,

$$b < x,$$

and so $x$ is not a lower bound for $T$.

**In summary**, if we can construct $c$ and $b$ as above then the possibilities $x^n < y$, $x^n > y$, are excluded and so it follows that $x^n = y$.

It remains to construct $c$ and $b$.

**To construct $c$** so that $x^n < c^n < y$ recall that we are assuming $x^n < y$.

Thus we may choose $k \in \mathbb{N}$ so that $1/k < x$ and so that

$$1/k < \frac{y - x^n}{2^n x^n - 1}.$$  

Set $c = x + 1/k$. Then $c > x$ and so

$$c^n > x^n \quad \text{and} \quad (1/k)x^{n-1}2^n < y - x^n.$$  

Now, by the Binomial theorem, (Proposition 2),

$$c^n - x^n = \sum_{i=0}^{n} \frac{n!}{i!(n-i)!}x^i(1/k)^{n-i} - x^n = \frac{1}{k} \sum_{i=0}^{n-1} \frac{n!}{i!(n-i)!}x^i(1/k)^{n-i-1}.$$  

Since $0 < 1/k < x$ it follows that $x^i(1/k)^{n-i-1} < x^{n-1}$.

Therefore, again by Proposition 2,

$$c^n - x^n < \frac{1}{k} \sum_{i=0}^{n-1} \frac{n!}{i!(n-i)!}x^{n-1} < (1/k)x^{n-1} \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} = (1/k)x^{n-1}(1+1)^n.$$  

Thus

$$c^n - x^n < (1/k)x^{n-1}2^n < y - x^n,$$

and so $c^n < y$. This completes the construction of $c$.

**To construct $b$** recall that we are assuming $x^n > y$.

Thus we may choose $k \in \mathbb{N}$ so that $0 < 1/k < x$ and

$$1/k < \frac{x^n - y}{nx^n - 1}.$$  

63
Set $b = x - 1/k$. Then $0 < b < x$ and so
\[ x - b = 1/k \quad \text{and} \quad b^n < x^n. \]
Thus by the Difference theorem (Proposition 1), since $x - b = 1/k$,
\[ x^n - b^n = (x - b) \sum_{i=0}^{n-1} x^i b^{n-1-i} \leq (x - b)nx^{n-1} = \frac{nx^{n-1}}{k} < x^n - y. \]
Therefore $y < b^n$.

This completes the construction of $b$ and the proof of the Proposition.

q.e.d.

**Definition 24** If $y$ is a positive real number and $n$ is a natural number then the unique positive number $x$ such that $x^n = y$ is called the $n$th root of $y$ and is denoted by $x = y^{1/n}$.

Proposition 8 allows us to construct our first real number that is not rational. It shows that there is a unique positive real number $x$ such that $x^2 = 2$. Thus now we know that $\sqrt{2}$ exists. On the other hand, according to Exercise 2.4, $x$ is not a rational number.

We can now define any rational power of a positive real number. For this we need the following:

**Lemma 10** Suppose $n, q \in \mathbb{N}$ and $m, p \in \mathbb{Z}$. If $x$ is a positive real number and $m/n = p/q$ then
\[ (x^{1/n})^m = (x^{1/q})^p. \]

**Proof:** First note that by Example 8
\[ ((x^{1/n})^m)^{nq} = (x^{1/n})^{mnq} = ((x^{1/n})^n)^{mq} = x^{mq}. \]
In the same way,
\[ ((x^{1/q})^p)^{nq} = (x^{1/q})^{pqn} = ((x^{1/q})^q)^{pn} = x^{pn}. \]
Since $m/n = p/q$ it follows that $mq = pn$. Therefore
\[ ((x^{1/n})^m)^{nq} = ((x^{1/q})^p)^{nq}. \]
By Proposition 13 raising to the $(nq)^{th}$ power is a 1-1 map. Therefore
\[ (x^{1/n})^m = (x^{1/q})^p, \]
and the Lemma is proved. q.e.d.

Now let \( a \) be any rational number. We can express \( a \) in the form \( a = m/n \) with \( m \in \mathbb{Z} \) and \( n \in \mathbb{N} \). Lemma 10 shows that if \( x \) is a positive real number then \((x^{1/n})^m\) does not depend on the choice of \( m \) and \( n \). Thus we may make the

**Definition 25** Let \( x \) be any positive real number and let \( a \) be any rational number. Then

\[
x^a = (x^{1/n})^m,
\]

where \( a = m/n \) and \( m \in \mathbb{Z} \) and \( n \in \mathbb{N} \).

**Exercise 14 Powers**

1. Show that if \( a, b \) are rational numbers and \( x \) is a positive real number then

   \[
x^a x^b = x^{a+b}.
\]

2. Show that if \( a, b \) are rational numbers and \( x \) is a positive real number then

   \[
   (x^a)^b = x^{ab}.
\]

3. Show that if \( x \in \mathbb{R} \) satisfies \( x > 1 \) and if \( c \in \mathbb{Q} \) is positive, then \( x^c > 1 \).

4. Show that if \( x > 1 \) is a real number and if \( a < b \) are positive rational numbers then \( 0 < x^a < x^b \).

5. Show that if \( a \in \mathbb{Q} \) is positive and if \( 0 < x < y \) then \( x^a < y^a \).

6. If \( k \in \mathbb{N} \) and \( x > 0 \) is a real number, show by induction on \( k \) that

   \[
x^k \leq x \quad \text{if} \quad x < 1 \quad \text{and} \quad x^k \geq x \quad \text{if} \quad x > 1.
\]

7. Identify what is wrong with the following attempt to prove the assertion in the previous exercise:

   **Proof:** Let \( S(k) \) be the statement \( x^k \leq x \) if \( x < 1 \). When \( k = 1 \) then \( S(1) \) reads \( x \leq x \) which is true. Since \( x < 1 \), it follows that \( x^k < 1 \). Therefore \( x^k \leq x < 1 \). q.e.d.

8. If \( x, \varepsilon \in \mathbb{R} \) are positive and if \( x < 1 \) show that for some \( N \in \mathbb{N} \),

   \[
x^a < \varepsilon \quad \text{if} \quad a \in \mathbb{Q} \quad \text{and} \quad a \geq N.
\]

9. Let \( p, q \in \mathbb{N} \) and \( x > 0 \) be a real number. When is \( x^{p/q} > x \) and when is \( x^{p/q} < x \) ? Prove your answer is correct.

10. Let \( x \in \mathbb{R} \) and let \( S \) be the set of those rationals of the form \( u = p/10^k \) satisfying the three conditions: (i) \( p \in \mathbb{Z} \), (ii) \( k \in \mathbb{N} \), and (iii) \( u < x \). Show that \( x = \text{lub}(S) \). Hint: Follow the following steps:
(a) Let \( \varepsilon > 0 \) be an arbitrary positive number. Use Proposition 8 to conclude that there are rational numbers \( a, b \) such that \( x - \varepsilon < a < b < x \).

(b) Show that for some \( k \in \mathbb{N} \), \( 10^k b - 10^k a > 2 \). Explain why this implies that some integer \( p \) satisfies \( 10^k a < p < 10^k b \).

(c) Conclude that \( p/10^k \in S \) and that \( x - \varepsilon < p/10^k < x \).

4.7 Constructing the real numbers

Call a set \( S \) of rationals that is bounded above full if
\[
y \in S \text{ and } z < y \text{ implies that } z \in S.
\]

Then simply define the real numbers to be the subsets of the rationals that are bounded above and full, and identify each rational \( p/q \) with the set of all rationals \( \leq p/q \). Intuitively we have just "filled in the holes" among the rationals.

If \( S \) and \( T \) are sets of rationals that are bounded above and full, then we think of these sets as real numbers \( x \) and \( y \) and define \( x < y \) if \( S \subset T \), and \( x > y \) if \( S \supseteq T \). Then set \( x + y \) to be the set
\[
S + T = \{ p/q + m/n \mid p/q \in S \text{ and } m/n \in T \},
\]
and define the other algebraic operations in a similar way.

Of course, we then need to prove that all these definitions are well-defined, that all algebraic rules are still true, that the properties above for the order hold, and that the operations and order for the rationals remain as before. And then, of course, we need to prove Properties One through Four. All this is long and boring, so instead we will simply take for granted the existence of the real numbers satisfying those properties.
Chapter 5

Infinite Sequences

5.1 Convergent sequences

Definition 26 An infinite sequence \((x_i)_{i \geq k}\) of elements in a set \(S\) is a list of elements \(x_i \in S\), with \(i \in \mathbb{Z}\) and \(i \geq k\).

We say the sequence begins with \(x_k\). For any \(p \geq k\), the infinite sequence \((x_i)_{i \geq p}\) is called the \(p\)th tail of the original sequence.

Remark: The set of elements appearing in an infinite sequence may be finite, as illustrated by the infinite sequence 1, 0, 1, 0, 1, 0, ... in which only two integers, 0 and 1 appear.

In the world of applications we never need to know the exact value of a real number - we only need to know it within a given tolerance. Specs for any engineering design always specify heights, lengths, weights etc. up to so many fractions of an inch or so many millimeters, or so many fractions of a gram.

Infinite sequences \((x_i)\) of real numbers are therefore a fundamental tool for applications because they can be used to approximate a real number \(x\) to within a given tolerance. Intuitively that means that the error, \(|x_i - x|\), gets arbitrarily small as \(i\) gets larger. In other words, if we want to approximate \(x\) within a given tolerance we may simply use one of the \(x_i\) as long as \(i\) is sufficiently large.

Infinite sequences are also a fundamental tool in analysis, and for this we need as always to formalize the intuitive idea above:

Definition 27 An infinite sequence \((x_i)_{i \geq k}\) of real numbers converges to \(x \in \mathbb{R}\) if for each \(\varepsilon > 0\) there is some integer \(N \geq k\) (usually depending on \(\varepsilon\)) such that

\[ |x_i - x| < \varepsilon \quad \text{if} \quad i \geq N. \]

Lemma 11 If an infinite sequence \((x_i)\) of real numbers converges, it converges to a unique real number \(x\).
**Proof:** Suppose the sequence \((x_i)\) converges to both \(x\) and \(y\). Then for each \(\varepsilon > 0\) there is some \(N\) such that for \(i \geq N\) both \(|x_i - x| < \varepsilon/2\) and \(|x_i - y| < \varepsilon/2\).

It follows that

\[
|x - y| = |x - x_i + x_i - y| \leq |x_i - x| + |x_i - y| < \varepsilon.
\]

Since this is true for each \(\varepsilon\), \(x = y\).

**q.e.d.**

**Definition 28** If \((x_i)\) is a convergent sequence, the unique \(x\) to which it converges is called the **limit** of the sequence and is denoted by \(\lim_{i \to \infty} x_i\).

**Example 22** The infinite sequence \(1, 2, 3, \ldots\) of natural numbers does not converge.

**Proof:** This is proved by contradiction. Assume the sequence converges to some \(x\). Then for some \(N\),

\[
n \geq N \Rightarrow |x_n - x| < 1.
\]

Thus for \(n \geq N\),

\[
|x_n| = |x_n - x + x| \leq |x_n - x| + |x| < 1 + |x|.
\]

But by Property Three, for some \(k\),

\[
1 + |x| < k < k + N = x_{k+N}.
\]

This is the desired contradiction. **q.e.d.**

**Example 23**

1. The infinite sequence 0, 1, −1, 2, −2, 3, −3, . . . which lists all the integers. This sequence does not converge.
2. The infinite sequence \((x_n)_{n \geq 1}\) defined by \(x_n = 1/n\). This sequence converges to 0.
3. The infinite sequence \((x_n)_{n \geq 1}\) defined inductively by \(x_1 = 1\), and

\[
x_n = \sum_{i=1}^{n-1} 2^{x_i}.
\]

This sequence does not converge.

4. The infinite sequence \((x_n)_{n \geq 1}\) defined by: \(x_n\) is the \(n\)th largest prime number. This sequence does not converge.
5. The infinite sequence \((x_n)_{n \geq 1}\) defined by: \(x_n = 1 + (-1/2)^n\). This sequence converges to 1.

**Lemma 12** If \((x_i)\) and \((y_i)\) are infinite sequences of real numbers converging respectively to \(x\) and \(y\), then \((x_i y_i)\) converges to \(xy\).

**Proof:** If \(\varepsilon > 0\), choose \(N \in \mathbb{N}\) so that if \(n \geq N\) then
\[
|x_n - x| < \frac{\varepsilon}{2(|y| + 1)}, \quad |y_n - y| < \frac{\varepsilon}{2(|x| + 1)},
\]
and
\[
|y_n - y| < 1.
\]
Then for \(n \geq N\),
\[
|y_n - y| = |y_n - y + y| \leq |y_n - y| + |y| < |y| + 1.
\]
Therefore for \(n \geq N\),
\[
|x_n y_n - xy| = |x_n y_n - x y_n + x y_n - xy| \leq |x_n - x||y_n| + |x||y_n - y| < \frac{\varepsilon}{2(|y| + 1)} (|y| + 1) + |x| \frac{\varepsilon}{2(|x| + 1)} \leq \varepsilon.
\]
q.e.d

**Exercise 15 Convergence**
As always, provide complete proofs for your answers.

1. In each of the examples in 23 supply proofs for the statements about convergence.

2. Suppose \((x_n)_{n \geq 1}\) is an infinite sequence of real numbers converging to \(x\).
   Define the sequence \((y_n)_{n \geq 1}\) by \(y_n = x_{2n}\). Is this sequence convergent, and if so, what is its limit? Prove your answer is correct.

3. Suppose \((x_n)_{n \geq 1}\) is an infinite sequence of real numbers. Suppose the sequence \(y_n = x_{3n}\) converges to \(y\) and that the sequence \(z_n = x_{3n+2}\) converges to \(z\). If \(y \neq z\) show that the original sequence \((x_n)\) is not convergent.

4. Show that if \(x\) is any real number then there is an infinite sequence of rational numbers converging to \(x\).

5. Suppose \((x_n)_{n \geq 1}\) is an infinite sequence of real numbers converging to \(x\).
   Define a sequence \((y_n)_{n \geq 1}\) by \(y_n = x_{n+1} - x_n\). Prove that the sequence \((y_n)_{n \geq 1}\) is convergent, and find its limit. Prove your answer is correct.

6. Suppose \((x_i)\) and \((y_i)\) are infinite sequences of real numbers converging respectively to \(x\) and \(y\).
   
   (a) Show that \((x_i + y_i)\) converges to \(x + y\).

   (b) If \(y \neq 0\) show that for some \(n \in \mathbb{N}\), \(y_i \neq 0\) for \(i \geq n\). In this case prove that the infinite sequence \((1/y_i)_{i \geq n}\) converges to \(1/y\).
(a) If \((x_i)\) is a convergent sequence and if each \(x_i \leq b\), show that \(\lim_{i} (x_i) \leq b\).

(b) If \((x_i)\) is a convergent sequence and if each \(x_i \geq a\), show that \(\lim_{i} (x_i) \geq a\).

7. If \((x_i)\) is an infinite sequence converging to \(x\), show that the sequence \((|x_i|)\) is convergent, and find its limit.

5.2 Bounded sequences

Definition 29

1. An infinite sequence, \((x_i)_{i \geq k}\), of real numbers is **bounded above** if for some real number \(b\),
   
   \[ x_i \leq b \quad \text{for all } i \geq k. \]

2. An infinite sequence, \((x_i)_{i \geq k}\), of real numbers is **bounded below** if for some real number \(a\),
   
   \[ a \leq x_i \quad \text{for all } i \geq k. \]

3. An infinite sequence is **bounded** if it is both bounded above and bounded below.

Definition 30

1. An infinite sequence, \((x_i)_{i \geq k}\), of real numbers is **increasing** if each \(x_i \leq x_{i+1}\).

2. An infinite sequence, \((x_i)_{i \geq k}\), of real numbers is **decreasing** if each \(x_i \geq x_{i+1}\).

Lemma 13

1. An increasing sequence, \((x_i)_{i \geq k}\), which is bounded above converges, and
   
   \[ \lim_{i} (x_i) = \text{lub}\{x_i \mid i \geq k\}. \]

2. A decreasing sequence, \((y_i)_{i \geq k}\), which is bounded below converges, and
   
   \[ \lim_{i} (y_i) = \text{glb}\{y_i \mid i \geq k\}. \]

Proof:

1. Set \(S = \{x_i \mid i \geq k\}\).
   
   Then \(S\) is bounded above.
   
   Fix any \(\varepsilon > 0\).
   
   By the definition of \(\text{lub}(S)\), \(\text{lub}(S) - \varepsilon\) is not an upper bound for \(S\).
   
   Therefore there is some \(x_p\) such that
   
   \[ x_p > \text{lub}(S) - \varepsilon. \]
Since the sequence is increasing, for any $i \geq p$,

$$x_i > \text{lub}(S) - \varepsilon.$$  

On the other hand, since each $x_i \in S$, for $i \geq p$

$$x_i \leq \text{lub}(S).$$

Therefore, for $i \geq p$,

$$0 \leq \text{lub}(S) - x_i < \varepsilon.$$  

In particular,

$$|\text{lub}(S) - x_i| < \varepsilon, \quad i \geq p.$$  

Since $\varepsilon$ was any positive real number, this is precisely the statement that the sequence converges to $\text{lub}(S)$.

2. Set $T = \{y_i \mid i \geq k\}$.

Then $T$ is bounded below.

Set $-T = \{-y_i \mid i \geq k\}$.

Since $T$ is bounded below and multiplication by $-1$ reverses inequalities,

$$\text{lub}(-T) = -\text{glb}(T).$$

Moreover, the sequence $(y_i)_{i \geq k}$ is increasing.

Thus by the first part of the Lemma, if $\varepsilon > 0$ there is a $p \in \mathbb{N}$ such that

$$|\text{lub}(-T) - (-y_i)| < \varepsilon, \quad i \geq p.$$  

Since $\text{lub}(-T) = -\text{glb}(T)$ it follows that $\text{glb}(T) = -\text{lub}(-T)$. Therefore, for $i \geq p$,

$$|y_i - \text{glb}(T)| < \varepsilon, \quad i \geq p.$$  

Thus the sequence $(y_i)_{i \geq k}$ converges to $\text{glb}(T)$.

\textbf{q.e.d.}

\textbf{Exercise 16}

1. Show that convergent sequences are bounded. (Hint: If $x_n$ converges to $x$ then for some $N, x - 1 < x_n < x + 1$ if $n \geq N$. Then show that the entire sequence is bounded.

2. Which of the following infinite sequences are bounded, which are increasing, which are decreasing, and which converge to a limit? If the sequence is convergent, find the limit. As always prove your answers.
(a) 1, −1, 1, −1, 1, −1, ...
(b) $(x_q)_{q \geq 1} = x^{1/q}$ where $0 < x < 1$ is fixed.
(c) $(x_q)_{q \geq 1} = x^{1/q}$ where $x = 1$.
(d) $(x_q)_{q \geq 1} = x^{1/q}$ where $x > 1$ is fixed.

3. Suppose $a \in Q$ is positive. If $(x_i)$ is a sequence of positive real numbers converging to $x > 0$, show that the sequence $(x_i^a)$ converges and find its limit, proving your answer. Is the converse true? (Hint: use the following steps:
(a) Show by induction that for $n \in N$, $(x_i^n)$ converges and find its limit.
(b) Show that for $n \in N, (x_i^{1/2})$ converges and find its limit.
(c) Complete the proof when the exponent is a positive rational.

4. Suppose $x$ is a real number and that $0 < x < 1$.
(a) Show that the set $S = \{x^n \mid n \in N\}$ is bounded below.
(b) Show that $\text{glb}(S) = \text{glb}(\{x^n \mid n \geq 2\})$.
(c) Show that $\text{glb}(S) = x \cdot \text{glb}(S)$, and conclude that $\text{glb}(S) = 0$.
(d) Prove that the sequence $(x^n)_{n \geq 1}$ converges and find its limit. Prove your answer.

**Proposition 14** A sequence $(x_n)$ of real numbers converges if and only if for each $\varepsilon > 0$ there is some integer $N$ such that $|x_n - x_m| < \varepsilon$ for all $n, m \geq N$.

**Proof:** Suppose first that the sequence converges to $x$.
Then for any $\varepsilon > 0$, there is some $N$, such that

$$|x_n - x| < \varepsilon/2 \text{ if } i \geq N.$$ 

Thus if $n, m \geq N$,

$$|x_n - x_m| = |x_n - x + x - x_m| \leq |x_n - x| + |x_m - x| < \varepsilon.$$ 

Conversely, suppose that for each $\varepsilon > 0$ there is some integer $N$ such that

$$|x_n - x_m| < \varepsilon \text{ if } n, m \geq N.$$ 

Define sets $S_n$ by setting

$$S_n = \{x_i \mid i \geq n\}.$$ 

We first show that $S_1$ is bounded. By hypothesis, for some $N$, if $n \geq N$, then

$$|x_n - x_N| < 1.$$ 

Thus for $n \geq N$,

$$|x_n| \leq |x_n - x_N| + |x_N| < 1 + |x_N|.$$ 

72
Since for $n < N$, $|x_n| \leq \text{Max}\{|x_1|, \ldots, |x_{N-1}|\}$, it follows that for all $n$,

$$|x_n| < 1 + |x_N| + \text{Max}\{|x_1|, \ldots, |x_{N-1}|\}. $$

Thus $S_1$ is bounded.

Now, since each $S_n \subset S_1$, it follows that each $S_n$ is bounded. Thus we may set

$$a_n = \text{glb}(S_n).$$

Since each $S_{n+1} \subset S_n$, it follows that $a_n$ is a lower bound for $S_n$. But $a_{n+1}$ is the greatest lower bound for $S_{n+1}$, and therefore

$$\cdots \leq a_n \leq a_{n+1} \leq \cdots$$

In other words, $(a_n)$ is an increasing sequence.

Moreover, since $a_n$ is a lower bound for $S_n$, for all $n$ we have

$$a_n \leq x_n < 1 + |x_N| + \text{Max}\{|x_1|, \ldots, |x_{N-1}|\}. $$

Thus the sequence $a_n$ is bounded above. Apply Lemma 13 to conclude that the sequence $a_n$ is convergent to $a = \text{lub}\{a_n\}$.

Finally, fix any $\varepsilon > 0$ and choose $N$ so that

$$|a_N - a| < \varepsilon/3 \text{ and } |x_n - x_m| < \varepsilon/3 \text{ if } n, m \geq N.$$

Since $a_N$ is the glb for $S_N$ it follows that for some $m \geq N$ we have

$$|x_m - a_N| < \varepsilon/3$$

Therefore, for $n \geq N$,

$$|x_n - a| \leq |x_n - x_m| + |x_m - a_N| + |a_N - a| < \varepsilon.$$

In other words, the sequence $x_n$ converges to $a$. \textbf{q.e.d.}

\section{The Intersection Theorem}

We now come to a result, which again uses Property Four about the reals. To state the Theorem we recall the

\textbf{Notation:} If $a \leq b$ are real numbers, then

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$
Theorem 2 Let

\[ S_m \supset S_{m+1} \supset S_{m+2} \supset \cdots \supset S_k \supset \cdots \]

be an infinite sequence of non-empty bounded subsets of \( \mathbb{R} \). Set \( a_k = \text{glb}(S_k) \) and \( b_k = \text{lub}(S_k) \).

Then

1. \((a_k)\) is an increasing sequence bounded above, and convergent to \( a = \text{lub}\{a_k\} \);
2. \((b_k)\) is a decreasing sequence bounded below, and convergent to \( b = \text{glb}\{b_k\} \);
3. \( b - a \geq 0 \);
4. \((b_k - a_k)\) is a decreasing sequence of non-negative numbers convergent to \( b - a \);
5. \( \bigcap_k [a_k, b_k] = [a, b] \).

**Proof:** By definition, \( a_k \) is a lower bound for \( S_k \) and \( b_k \) is an upper bound for \( S_k \). Therefore

\[ a_k \leq b_k. \]

Since \( S_k \supset S_{k+1} \) it follows that \( a_k \) is a lower bound for \( S_{k+1} \) and \( b_k \) is an upper bound for \( S_{k+1} \). But by definition, \( a_{k+1} \) is the greatest lower bound for \( S_{k+1} \) and \( b_{k+1} \) is the least upper bound for \( S_{k+1} \). Therefore

\[ a_k \leq a_{k+1} \leq b_{k+1} \leq b_k. \]

Thus \((a_k)\) is an increasing sequence. Moreover,

\[ a_k \leq b_k \leq b_1. \]

Thus the sequence \((a_k)\) is both increasing and bounded above. Therefore by Lemma 13, the sequence converges to \( a = \text{lub}\{a_k\} \).

Similarly, \((b_k)\) is a decreasing sequence and since for all \( k \)

\[ a_1 \leq a_k \leq b_k. \]

The sequence \((b_k)\) is also bounded below. Thus by Lemma 13, the sequence \((b_k)\) converges to \( b = \text{glb}\{b_k\} \).

In particular the sequence \((b_k - a_k)\) converges to \( b - a \), and since each \( b_k - a_k \geq 0 \), it follows that \( b - a \geq 0 \).

It remains to show that
\[ \bigcap_k [a_k, b_k] = [a, b]. \]

But if \( x \in [a, b] \) then for all \( k \),
\[ a_k \leq a \leq x \leq b \leq b_k, \]
and so \( x \in [a_k, b_k] \). Thus \( x \in \bigcap_k [a_k, b_k] \); i.e.,
\[ \bigcap_k [a_k, b_k] \supseteq [a, b]. \]

On the other hand, suppose \( x \in \bigcap_k [a_k, b_k] \). Then for all \( k \)
\[ a_k \leq x \leq b_k. \]
Thus \( x \) is an upper bound for the set \( \{a_k\} \) and a lower bound for the set \( \{b_k\} \).
It follows that
\[ x \geq a \text{ and } x \leq b. \]
Thus \( x \in [a, b] \), and
\[ \bigcap_k [a_k, b_k] \subset [a, b]. \]

This completes the proof of the theorem.
q.e.d.

**Corollary** Let \((S_k)_{k \geq m}\) be as in Theorem 2, and suppose \( \lim_k (b_k - a_k) = 0 \).
Then
\[ \bigcap_k [a_k, b_k] \]
is a single point \( c \). If \((x_k)\) is an infinite sequence with \( x_k \in [a_k, b_k] \) then the sequence \((x_k)\) converges to \( c \).

**Proof:** It follows from Theorem 2 that \( b - a = 0 \); i.e., \( b = a \).

Denote this point by \( c \). Then \([a, b] = \{c\}\), and so
\[ \bigcap_k [a_k, b_k] = \{c\}. \]

Now, let \( \varepsilon > 0 \) be any positive number.
Then choose \( N \) so that \( b_k - a_k < \varepsilon \) for \( k \geq N \).
Since \( c \in [a_k, b_k] \) and \( x_k \in [a_k, b_k] \) it follows that if \( k \geq n \), then
\[ |x_k - c| < \varepsilon. \]
Thus the sequence \((x_k)\) converges to \( c \).
q.e.d.
Exercise 17

1. Construct a sequence of open intervals $I(n)$ of length $l(n)$ such that the sequence $(l(n))$ converges to zero, and such that each $(I(n))$ and $I(n+1)$ have a point in common, but such that

$$\bigcap_n I(n) = \emptyset.$$

2. Let $S_n$ be the set of rational numbers $a$ satisfying $\sqrt{2}-1/n < a < \sqrt{2+1/n}$. Find $\cap_n S_n$.

3. Let $S_n$ be the sets in the previous problem. If $x_n \in S_n$, does the sequence $(x_n)$ converge? If so, find its limit.

4. Let $S_n$ be the set of non-zero numbers in $[-1/n, 1/n]$. What is the intersection of these sets?

5. Let $S_n$ be the open interval $(1, 1+1/n)$. What is the intersection of these sets?

5.4 Subsequences

Definition 31 A subsequence of a sequence $(x_n)_{n \geq k}$ is a sequence of the form $(y_i = x_{n_i})_{i \geq r}$, in which $n_r < n_{r+1} < n_{r+2} \ldots$ is a sequence of integers.

Exercise 18 Subsequences

1. Construct a sequence which has two convergent subsequences converging to different limits.

2. Construct a sequence which for each $n \in \mathbb{N}$ has a subsequence converging to $n$.

3. Show that any subsequence of a convergent sequence converges to the same limit.

4. Show that a sequence which is not bounded above has an increasing subsequence which is not bounded above.

5. Let $(x_n)$ be a sequence such that the set $S = \{x_n\}$ is finite. Show that there is a subsequence $(x_{n_i})$ such that for all $i, x_{n_i} = x_{n_1}$.

Theorem 3 Every bounded sequence contains a convergent subsequence.

Proof: Let $(x_n)_{n \geq 1}$ be a bounded sequence, and let $S = \{x_n \mid n \geq 1\}$. Since the sequence is bounded there are numbers $a_1 < b_1$ such that

$S \subset [a_1, b_1]$. 

76
We first construct by induction on \( n \) a decreasing sequence of intervals

\[
[a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n] \cdots
\]

with the following two properties:

1. Each interval \([a_n, b_n]\) contains \( x_k \) for infinitely many \( k \), and

2. 

\[
b_n - a_n = \frac{b_1 - a_1}{2^{n-1}}.
\]

By hypothesis, \( x_k \in [a_1, b_1] \) for all \( k \).

Now suppose by induction that the intervals \([a_i, b_i]\) are constructed for \( i \leq n \). By construction \([a_n, b_n]\) contains \( x_k \) for infinitely many \( k \).

Then there are two mutually exclusive possibilities: either

\[
[a_n, \frac{a_n + b_n}{2}] \text{ contains } x_k \text{ for infinitely many } k,
\]
or

\[
[a_n, (a_n + b_n)/2] \text{ contains } x_k \text{ for only finitely many } k.
\]

In the second case,

\[
[a_n = \frac{a_n + b_n}{2}, b_n]
\]

must contain \( x_k \) for infinitely many \( k \).

In the first case set

\[
a_{n+1} = a_n \quad \text{and} \quad b_{n+1} = \frac{a_n + b_n}{2}.
\]

In the second case set

\[
a_{n+1} = \frac{a_n + b_n}{2} \quad \text{and} \quad b_{n+1} = b_n.
\]

By our induction hypothesis \( b_n - a_n = (b_1 - a_1)/2^{n-1} \), and therefore

\[
b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2} = \frac{b_1 - a_1}{2^n}.
\]

This completes the inductive construction.

Our second step is to construct a subsequence \((x_{n_i})\) so that for all \( i \):

\[
x_{n_i} \in [a_i, b_i].
\]

First, set \( x_{n_1} = x_1 \). Then

\[
x_{n_1} = x_1 \in S \subset [a_1, b_1]
\]
Then suppose by induction that the $x_{n_i}$ are constructed for $i \leq r$.

Since $[a_r+1, b_r+1]$ contains $x_k$ for infinitely many $k$ there is a least integer $p > n_r$ such that $x_p \in [a_r+1, b_r+1]$.

Set $x_{n_{r+1}} = x_p$.

This completes the inductive construction of the subsequence.

It remains to show that this subsequence is convergent.

Since

$$b_r - a_r = \frac{b_1 - a_1}{2^{r-1}},$$

it follows that the sequence $b_r - a_r$ converges to zero. Thus by the Corollary to Theorem 2, since $x_{n_r} \in [a_r, b_r]$, the sequence $(x_{n_r})$ converges. q.e.d.
Chapter 6

Continuous Functions of a Real Variable

6.1 Real-valued Functions of a Real Variable

Functions from the reals to the reals are a central tool in almost every discipline that uses mathematics. Among the many examples are:

- Position and speed of a particle as a function of time.
- Crop yield as a function of total precipitation.
- Grade on an exam as a function of time spent studying.
- Cost to the national health system as a function of the amount of pollution in the air.
- Fraction of the population that is illiterate as a function of the average time per pupil spent in class on reading and writing.
- Wave length of light reaching us from a star as a function of its distance away.

As you may easily imagine, functions of a real variable are also a core part of mathematics itself. In fact this field is a prime example of how mathematics interacts with other disciplines: many of the problems and theorems in mathematics in this area are inspired by questions and phenomena from outside, while the results and techniques that mathematicians discover frequently get applied elsewhere.
This chapter focuses on two key concepts in the analysis of functions: limits and continuity. But first we establish some basic definitions.

**Definition 32 Intervals**

1. An interval is a subset of \( \mathbb{R} \) of one of the following forms (where \( a \) and \( b \) are any real numbers):
   (a) \( \mathbb{R} \)
   (b) \([a,b]\), \((a,b]\), \([a,b)\) or \((a,b)\)
   (c) \([a,\infty)\), \((a,\infty)\), \((\infty,b]\) or \((\infty,b)\)

2. The intervals in (b) are called finite; the first is closed, the next two are half closed, and the fourth is open.

3. The numbers \( a \) and \( b \) are called the end points of their intervals.

4. The closure of an interval \( D \) is the union of \( D \) together with any end points. It is denoted by \( \bar{D} \).

**Definition 33** A real-valued function is a set map

\[ f : D \to S \]

from an interval \( D \) to a subset \( S \subset \mathbb{R} \). \( D \) is called the domain of \( f \) and \( S \) is its target.

**Definition 34** If \( f : D \to S \) is a real-valued function and if \( E \) is a second interval contained in \( D \) then the restriction of \( f \) to \( E \) is the real function \( g : E \to S \) defined by \( g(x) = f(x) \), \( x \in E \).

**Example 24 Real-valued functions**

1. If \( f : D \to S \) and \( g : D \to S \) are real-valued functions then \( f + g : D \to \mathbb{R} \) and \( fg : D \to \mathbb{R} \) are the functions defined by
   \[ (f + g)(x) = f(x) + g(x) \quad \text{and} \quad (fg)(x) = f(x)g(x). \]

2. If \( f : D \to S \) is a real-valued function and if \( f(x) \neq 0 \) for all \( x \in D \) then \( 1/f : D \to \mathbb{R} \) is the function defined by \( (1/f)(x) = 1/f(x) \).

3. \( f : \mathbb{R} \to \mathbb{R} \) defined by
   \[ f(x) = \begin{cases} 0, & x \in \mathbb{Q}, \\ 1, & x \notin \mathbb{Q}. \end{cases} \]

4. \( f : (0,\infty) \to \mathbb{R} \) defined by: \( f(x) = 1/x \).
5. \( f : (0, 1) \to \mathbb{R} \) defined by: \( f(x) \) is the number in the 25th decimal place of \( x \).

6. 

\[
f(x) = \begin{cases} 
0, & x < 0, \\
1, & x \geq 0.
\end{cases}
\]

7. \( f(x) = \sum_{i=0}^{n} \lambda_k x^k \), where the \( \lambda_k \) are real numbers and \( \lambda_n \neq 0 \). Such a function is called a polynomial of degree \( n \).

### 6.2 Limits

Suppose \( f : D \to S \) is a real-valued function. An important question which arises is:

Can we use the values of \( f(x) \) near a point \( c \in D \) to determine a "limiting value"?

There is a good practical reason for this question. In practice one can never make an exact measurement at a specific point, \( c \). The best one can do is make measurements nearby and hope they give a good approximation to a measurement at \( c \) itself.

This concept has two fundamental ingredients:

1. We consider only values of the function near \( c \) and not the value at \( c \).

2. The values of the function at points near \( c \) must bunch ever more closely together as the points get closer to \( c \).

This idea is formalized in the following way:

**Definition 35** Suppose \( f : D \to S \) is a real-valued function defined in an interval \( D \), and that \( c \in D \). Then \( f(x) \to u \) as \( x \to c \) if for any \( \varepsilon > 0 \) there is some \( \delta > 0 \) such that

\[
|f(x) - u| < \varepsilon \quad \text{if} \quad x \in D \quad \text{and} \quad 0 < |x - c| < \delta.
\]

**Note:** Usually \( \delta \) will depend on \( \varepsilon \).

**Lemma 14** Suppose \( f : D \to S \) is a real-valued function defined in an interval \( D \). If for some \( c \in D \),

\[
f(x) \to u_1 \quad \text{and} \quad f(x) \to u_2
\]

as \( x \to c \). Then \( u_1 = u_2 \).
Proof: Fix \( \varepsilon > 0 \). Choose \( \delta_1 > 0 \) and \( \delta_2 > 0 \) so that for \( x \in D \),

\[
|f(x) - u_1| < \varepsilon/2 \quad \text{if} \quad 0 < |x - c| < \delta_1, \quad \text{and} \quad |f(x) - u_2| < \varepsilon/2 \quad \text{if} \quad 0 < |x - c| < \delta_2.
\]

Set \( \delta \) to be the lesser of \( \delta_1 \) and \( \delta_2 \). Then for \( x \in D \) and \( 0 < |x - c| < \delta \) we have

\[
|u_2 - u_1| = |u_2 - f(x) + f(x) - u_1| \leq |u_2 - f(x)| + |f(x) - u_1| < \varepsilon.
\]

Thus \( |u_2 - u_1| < \varepsilon \) for all \( \varepsilon > 0 \) and so \( u_2 = u_1 \). q.e.d.

Definition 36 If \( f(x) \to u \) as \( x \to c \), then \( u \) is the limit of \( f(x) \) as \( x \) approaches \( c \) and we write

\[
l_{x \to c} f(x) = u.
\]

Important Remarks:

1. The definition of limit specifies the condition

\[
0 < |x - c| < \delta.
\]

Thus the limit depends on the values of \( f(x) \) when \( x \) is close to but different from \( c \). Even when \( c \) is in the interval \( D \) the value of \( f(c) \) is irrelevant to the definition of \( \lim_{x \to c} f(x) \).

2. If \( f : D \to S \) is a real-valued function and \( c \in \overline{D} \), then there are two possibilities, and either could be correct:
   
   (a) The limit of \( f(x) \) as \( x \to c \) exists, or
   (b) This limit does not exist.

3. If \( f : D \to S \) is a real-valued function and \( c \in D \) then there are three possibilities:
   
   (a) The limit of \( f(x) \) as \( x \to c \) exists and \( \lim_{x \to c} f(x) = f(c) \), or
   (b) The limit of \( f(x) \) as \( x \to c \) exists and \( \lim_{x \to c} f(x) \neq f(c) \), or
   (c) The limit of \( f(x) \) as \( x \to c \) does not exist.

Example 25 Limits

1. \( D = \mathbb{R}, S = \mathbb{R} \) and \( f(x) = 5x + 1 \). Then

\[
\lim_{x \to 1} f(x) = 6.
\]

In fact given \( \varepsilon > 0 \) set \( \delta = \varepsilon/5 \).

If \( 0 < |x - 1| < \delta \) then

\[
|f(x) - 6| = 5|x - 1| < 5(\varepsilon/5) = \varepsilon.
\]
2. $D = S = \mathbb{R}$ and

$$f(x) = \begin{cases} 
5x + 1, & x \neq 1, \\
17, & x = 1. 
\end{cases}$$

Then, as above, $\lim_{x \to 1} f(x) = 6$, but $f(1) = 17$.

3. $D = (-\infty, 1), S = \mathbb{R}$ and $f(x) = 5x + 1$. Then

$$\lim_{x \to 1} f(x) = 6.$$

(Same proof as above.)

**Review of Negations:**

The negation of the statement "A is true" is the statement "A is false".

In analysis we frequently prove a statement is true by showing that the negation of the statement leads to a contradiction, and for this it is often helpful to restate the negation in a more useful form. Here is a typical example:

**Statement A:** $\lim_{x \to c} f(x) = u$.

So what does it mean for Statement A to be false?

If Statement A is true then for each $\varepsilon > 0$ there is a $\delta > 0$ satisfying the requirements of the definition above.

Therefore to say Statement A is false is precisely to say there must be some $\varepsilon > 0$ so that no $\delta > 0$ will work for that particular $\varepsilon$.

Now what does it mean for $\delta$ to work for that $\varepsilon$?

It means precisely that for every $x \in D$ such that $0 < |x - c| < \delta$ it is true that $|f(x) - u| < \varepsilon$.

Thus to say that no $\delta > 0$ will work for that $\varepsilon$ means that for each $\delta > 0$ there is some $x \in D$ such that

$$0 < |x - c| < \delta \quad \text{and} \quad |f(x) - u| \geq \varepsilon.$$

Altogether then we have:

**Statement A** is false if and only if Statement B below is true:

**Statement B:** for some $\varepsilon > 0$ and every $\delta > 0$ there is some $x \in D$ such
that 0 < |x - c| < δ and |f(x) - u| ≥ ε.

Statement B is the **negation** of Statement A.

**Example 26**  \( D = S = \mathbb{R} \) and

\[
f(x) = \begin{cases} 
1/|x|, & x \neq 0, \\
0, & x = 0.
\end{cases}
\]

*In this example \( \lim_{x \to 0} f(x) \) does not exist.*

We have to prove that any \( u \in \mathbb{R} \) is not the limit of \( 1/|x| \) as \( x \to 0 \). We fix some \( u \) and prove the negation is true. Thus we have to show that for some \( \varepsilon > 0 \) and every \( \delta > 0 \) there is some \( x \in \mathbb{R} \) such that

\[
0 < |x| < \delta \quad \text{and} \quad |1/|x| - u| > \varepsilon.
\]

First, choose \( \varepsilon = 1 \). Then, for any \( \delta > 0 \) choose \( x \) so that

\[
|x| < \delta \quad \text{and} \quad |x| < \frac{1}{|u| + 1}.
\]

Then \( 1/|x| > |u| + 1 \). Thus

\[
|1/|x| - u| \geq 1/|x| - u > 1 + |u| - u \geq 1 = \varepsilon.
\]

Therefore \( f(x) \) does not converge to \( u \) as \( x \to 0 \). Since \( u \) was any real number, the limit does not exist.

In the previous chapter we introduced the limit of an infinite sequence, and now we have another kind of limit. It is natural to wonder if these two kinds of limit are connected, and indeed they are. The next Proposition shows how.

**Proposition 15** Let \( f : D \to S \) be a real-valued function defined in an interval \( D \), and suppose \( c \in D \). Then

\[
\lim_{x \to c} f(x) = u
\]

if and only if

\[
\lim_{n} f(x_n) = u
\]

whenever \( (x_n) \) is a sequence of points in \( D \) different from \( c \), but convergent to \( c \).

**Proof:** We have two prove two things:

1. If \( \lim_{x \to c} f(x) = u \) then \( \lim_{n} f(x_n) = u \) whenever \( (x_n) \) is a sequence of points in \( D \) different from \( c \) but convergent to \( c \).
2. If \( \lim_{n} f(x_n) = u \) whenever \((x_n)\) is a sequence of points in \( D \) different from \( c \) but convergent to \( c \), then \( \lim_{x \to c} f(x) = u \).

For the first statement our hypothesis is \( \lim_{x \to c} f(x) = u \).
Then if \((x_n)\) is any sequence of points in \( D \) different from but converging to \( c \), we need to prove that
\[
\lim_{n} f(x_n) = u.
\]
In other words, for any \( \varepsilon > 0 \) we need to show that for some \( N \),
\[
|f(x_n) - u| < \varepsilon \quad \text{if} \quad n \geq N.
\]
Fix some \( \varepsilon > 0 \).
We are supposing \( f(x) \) has \( u \) as limit when \( x \) approaches \( c \).
It follows that for some \( \delta > 0 \),
\[
|f(x) - u| < \varepsilon \quad \text{if} \quad 0 < |x - c| < \delta.
\]
Since the sequence \((x_n)\) converges to \( c \) and no \( x_n = c \) there is some \( N \in \mathbb{N} \) for which
\[
0 < |x_n - u| < \delta \quad \text{if} \quad n \geq N.
\]
Therefore \( |f(x_n) - u| < \varepsilon \) if \( n \geq N \). This proves that the sequence \( f(x_n) \) converges to \( u \).

For the second statement our hypothesis is that whenever \((x_n)\) is a sequence of points in \( D \) different from but converging to \( c \) then
\[
\lim_{n} f(x_n) = u,
\]
and we need to prove that
\[
\lim_{x \to c} f(x) = u.
\]
We prove this by showing that the negation of this statement is false, and so the statement must be true.

As described above the negation is Statement B:

For some \( \varepsilon > 0 \) and every \( \delta > 0 \) there is some \( x \in D \) such that:
\[
0 < |x - c| < \delta \quad \text{and} \quad |f(x) - u| \geq \varepsilon.
\]
We prove this is false by contradiction: we assume Statement B is true and deduce a contradiction.

Now if Statement B is true, then for each \( n \in \mathbb{N} \) there is some point \( x_n \in D \) such that
\[
0 < |x_n - c| < \frac{1}{n} \quad \text{and} \quad |f(x_n) - u| \geq \varepsilon.
\]
This defines a sequence \((x_n)\) of points in \(D\) different from \(c\).
Moreover, for any real number \(\sigma > 0\) there is some \(k \in \mathbb{N}\) such that \(1/k < \sigma\)
(Property Three for the real numbers).
Therefore, for \(n \geq k\)
\[|x_n - c| < 1/n \leq 1/k < \sigma.\]
Thus the sequence \((x_n)\) converges to \(c\).

Now our hypothesis states that since \((x_n)\) converges to \(c\), it follows that \(f(x_n)\) converges to \(u\). But we constructed the sequence so that
\[|f(x_n) - u| \geq \varepsilon\]
for all \(n\), which contradicts our hypothesis.

It follows that Statement B cannot be correct and therefore that \(\lim_{x \to c} f(x) = u\).
q.e.d.

**Exercise 19 Limits**

1. If \(c < d\) are points in an interval \(D\), show that \([c, d] \subset D\).

2. Suppose \(f : D \to S\) is a real-valued function defined in an interval \((a, b)\) and that for some \(u \in \mathbb{R}\), \(f(x) < u\) for all \(x \in (a, b)\). If \(\lim_{x \to b} f(x)\)
exists, show that
\[\lim_{x \to b} f(x) \leq u.\]

3. For each of the following choices of \(D\), \(f\), and \(c\), and with \(S = \mathbb{R}\) decide if \(\lim_{x \to c} f(x)\) exists and when it does find the limit.
   
   (a) \(D = (-1, 0), c = 0, f(x) = 1, x \in D\).
   
   (b) \(D = [0, 1), c = 0, \text{ and } f(x) = 0, x \in D\).
   
   (c) \(D = (-1, 1), c = 0, \text{ and } f(x) = \begin{cases} 1, & x \in (-1, 0], \\ 0, & x \in (0, 1). \end{cases}\)

4. Suppose \(f : D \to S\) and \(g : D \to S\) are real-valued functions in an interval \(D\). If \(c \in D\) and if \(\lim_{x \to c} f(x)\) and \(\lim_{x \to c} g(x)\) exist, show that
\[\lim_{x \to c} (f(x) + g(x)) = \lim_{x \to c} f(x) + \lim_{x \to c} g(x),\]
and
\[\lim_{x \to c} (f(x)g(x)) = (\lim_{x \to c} f(x))(\lim_{x \to c} g(x)).\]
In particular conclude that these limits exist.
5. Suppose $f : D \to S$ is a real-valued function in an interval $D$. Suppose further that for some $c \in D$ and some $\alpha > 0$, $f(x) \neq 0$ if $0 \leq |x - c| < \alpha$. Show that

$$\lim_{x \to c} \frac{1}{f(x)}$$

exists if $\lim_{x \to c} f(x)$ exists and $\lim_{x \to c} f(x) \neq 0$.

6. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a polynomial. Use Exercise 19.4 to show that for all $y \in \mathbb{R},$

$$\lim_{x \to y} f(x) = f(y).$$

7. Suppose $z \in \mathbb{R}$ and $f : D \to \mathbb{R}$ is a real-valued function in an interval. Assume that for each $0 < \varepsilon < 1$ there is a $\delta > 0$ such that if $0 < |x - c| < \delta$ then $|f(x) - z| < \varepsilon$. Show that $\lim_{x \to c} f(x) = z$.

8. Suppose $x, y$ are positive real numbers and $n \in \mathbb{N}$.

(a) If $n \geq 2$ use the Difference theorem to show that $|x^{1/n} - y^{1/n}| < \frac{|x-y|}{y^{1/n} - x^{1/n}}$.

(b) Show that $\lim_{x \to y} x^{1/n} = y^{1/n}$.

9. Do the following limits exist?

$$\lim_{x \to 0} \sin \left( \frac{1}{x} \right) \text{ and } \lim_{x \to 0} x \sin \left( \frac{1}{x} \right).$$

You may use the following facts:

(a) $-1 \leq \sin(y) \leq 1$ for all $y \in \mathbb{R}$.

(b) $\sin \left( \frac{\pi k}{2} \right) = \begin{cases} 
0, & k = 2n, \\
1, & k = 4n + 1, \\
-1, & k = 4n - 1.
\end{cases}$

10. Define a real-valued function $f : \mathbb{R} \to \mathbb{R}$ in the interval by

$$f(x) = \begin{cases} 
0, & x < 0, \\
1, & x \geq 0.
\end{cases}$$

For which $c \in \mathbb{R}$ does $\lim_{x \to c} f(x)$ exist?

11. Let $n \in \mathbb{N}$ and define a function $f$ in $(-\infty, a)$ by

$$f(x) = \frac{x^n - a^n}{x - a}.$$

Show that $\lim_{x \to a} f(x)$ exists, find the limit and prove your answer is correct.
6.3 Continuity

Let $f : D \rightarrow S$ be a real-valued function defined in an interval, $D$. If $c \in D$ and if $\lim_{x \to c} f(x)$ exists then we have two numbers:

$$\lim_{x \to c} f(x) \quad \text{and} \quad f(c).$$

As we saw in the previous exercises, these two numbers may be different!

However, if these numbers are the same then we can approximate the value at $c$ of $f$ by the values nearby $c$. Functions for which these two numbers are the same have many good properties, and play a crucial role in mathematics. In fact they have a name:

**Definition 37** A real-valued function $f : D \rightarrow S$ in an interval $D$ is **continuous** if for each $c \in D$ the limit $\lim_{x \to c} f(x)$ exists and

$$\lim_{x \to c} f(x) = f(c).$$

**Lemma 15** Suppose $f : D \rightarrow S$ is a continuous function in an interval $D$. Then

1. For each $\varepsilon > 0$ and $c \in D$ there is a $\delta > 0$ such that if $x \in D$ and $|x - c| < \delta$ then

   $$|f(x) - f(c)| < \varepsilon.$$

2. If $(x_n)$ is a sequence of points in $D$ converging to $c \in D$ then the sequence $(f(x_n))$ converges to $f(c)$.

**Remark:**

1. The $\delta$ in the Lemma will usually depend on both $\varepsilon$ and $x$.

2. The Lemma states that continuous functions preserve convergent sequences!

**Proof:** of Lemma 15

1. Since $f$ is continuous, $\lim_{x \to c} f(x) = f(c)$. In particular, the limit exists. Fix any $\varepsilon > 0$.

   By the definition of limit, there is some $\delta > 0$ such that if $x \in D$ and $0 < |x - c| < \delta$, then

   $$|f(x) - f(c)| < \varepsilon.$$

   But this inequality is trivially true if $x = c$. 

88
2. Fix any \( \varepsilon > 0 \).
   Choose \( \delta > 0 \) so that
   \[
   |f(x) - f(c)| < \varepsilon \quad \text{if} \quad x \in D \quad \text{and} \quad |x - c| < \delta.
   \]
   Since the sequence \((x_n)\) in our hypothesis converges to \( c \in D \), it follows
   that for some \( N \),
   \[
   |x_n - c| < \delta \quad \text{if} \quad n \geq N.
   \]
   Thus for \( n \geq N \),
   \[
   |f(x_n) - f(c)| < \varepsilon.
   \]
   Therefore \((f(x_n))\) converges to \( f(x) \).

q.e.d.

Exercise 20 Continuous functions

1. Show that if \( \lambda \in \mathbb{R} \) the constant function \( \mathbb{R} \to \lambda \) is continuous.

2. If \( f : D \to S \) and \( g : D \to S \) are continuous functions in an interval \( D \),
   show that the functions \( f + g \) and \( fg \) are continuous.

3. If \( f : D \to S \) is a continuous function in an interval \( D \), and if \( f(x) \neq 0 \)
   for all \( x \in D \), show that \( 1/f \) is continuous.

4. Show that polynomials are continuous functions in \( \mathbb{R} \).

5. If \( n \in \mathbb{N} \) show that \( x^{1/n} \) is a continuous function in \((0, \infty)\).

6. Show that \( \sin x \) and \( \cos x \) are continuous functions. (You may use the
   trigonometric formulae for \( \sin(x+y) \) and \( \cos(x+y) \).)

7. If \( f : D \to S \) is a continuous function in an interval \( D \) and if \( g : \mathbb{R} \to T \)
   is a continuous function in \( \mathbb{R} \), show that the composite \( g \circ f \) is continuous.

8. If \( f : D \to S \) is a continuous function defined in an interval \( D \), show that
   the restriction of \( f \) to a second interval contained in \( D \) is continuous.

9. If \( f : D \to S \) is a continuous function defined in \( D = [a, b) \) and if
   \( \lim_{x \to a} f(x) = a \), show that a continuous function \( g : [a, b] \to \mathbb{R} \) is defined by
   \[
   g(x) = \begin{cases} 
   f(x), & x \in [a, b), \\
   a, & x = b.
   \end{cases}
   \]

10. Suppose \( f : D \to S \) is a continuous function in an interval \( D \).
    
    (a) Show that the function \( |f| : D \to \mathbb{R} \) defined by \( |f(x)| = |f(x)| \) is
    continuous.
(b) Show that the functions $f_+: D \to \mathbb{R}$ and $f_-: D \to \mathbb{R}$ defined by

$$
f_+(x) = \begin{cases} f(x), & f(x) \geq 0, \\ 0, & f(x) < 0. \end{cases}$$

and

$$
f_-(x) = \begin{cases} -f(x), & f(x) \leq 0, \\ 0, & f(x) > 0. \end{cases}$$

are continuous. (Hint: Consider $1/2(f + |f|)$.)

(c) Show that $f = f_+ - f_-$. 

Recall that a continuous function is a real-valued function whose value at a given point is the limit of the values at points approaching the given point. Continuous functions in a closed interval automatically satisfy a stronger condition:

**Proposition 16** Suppose $f : D \to S$ is a continuous function defined in a closed interval $D = [a, b]$. Then for all $\varepsilon > 0$ there is some $\delta > 0$ such that

$$|f(y) - f(x)| < \varepsilon \quad \text{if} \quad x, y \in [a, b] \quad \text{and} \quad |x - y| < \delta.$$ 

**Remark:** In contrast with Lemma 15, because $[a, b]$ is closed the $\delta$ does not depend on $x$.

**Proof of Proposition 16:** We assume the Proposition is false and derive a contradiction.

If the Proposition is false there is some $\varepsilon > 0$ such that for each $\delta > 0$ there is a corresponding pair of points $x, y \in [a, b]$ satisfying

$$|f(x) - f(y)| \geq \varepsilon \quad \text{and} \quad |x - y| < \delta.$$ 

In particular, we may find sequences $(x_n)$ and $(y_n)$ of points in $[a, b]$ such that

$$|f(x_n) - f(y_n)| \geq \varepsilon \quad \text{and} \quad |x_n - y_n| < \frac{1}{n}.$$ 

Since the sequences $(x_n)$ and $(y_n)$ are bounded it follows from Theorem 3 that the sequence $(x_n)$ contains a convergent subsequence $(u_i) = (x_{n_i})$. In the same way the sequence $(v_i) = (y_{n_i})$ contains a convergent subsequence $(v_{i_k})$.

Now Exercise 15.3 asserts that the subsequence $(u_{i_k})$ is also convergent.

Set

$$\lim_k(u_{i_k}) = x \quad \text{and} \quad \lim_k(v_{i_k}) = y.$$
Since the $x_n$ and the $y_n$ are in $[a, b]$, it follows from Exercise 15 that $x \in [a, b]$ and $y \in [a, b]$.

Since $|x_n - y_n| < 1/n$ it follows that

$$\lim_{k} |u_{i_k} - v_{i_k}| = 0,$$

and so $x = y$.

Since $f$ is continuous by Lemma 14.2 it preserves convergent sequences and so

$$f(x) = \lim_{k} (f(u_{i_k}) = \lim_{k} (f(v_{i_k}) = f(y).$$

But this is impossible because

$$|f(y_n) - f(x_n)| \geq \varepsilon$$

for all $n$.

This is the desired contradiction.

q.e.d.

### 6.4 Continuous functions preserve intervals

In this section we establish an important and fundamental property of continuous functions.

**Theorem 4** Suppose $f : D \to S$ is a continuous function defined in a closed interval $D = [a, b]$. Then the image of $f$ is a closed interval:

$$f([a, b]) = [c, d]$$

for some real numbers $c \leq d$.

**Consequences of Theorem 4.**

1. **Maximum and minimum values.** Since $[c, d]$ is the image of $f$ there must be points $u, v \in [a, b]$ such that

$$f(u) = c \quad \text{and} \quad f(v) = d.$$

Thus

$$f(u) \leq f(x) \leq f(v)$$

for all $x \in [a, b]$.

2. **Intermediate values.** Since $[c, d]$ is the image of $f$ it follows that for every $y \in [c, d]$ there is some $x \in [a, b]$ such that

$$f(x) = y.$$
Important Remarks: Let \( f : D \to S \) be a continuous function defined in the interval \( D = [a, b] \).

1. Then the Theorem states that
   - (a) The set of values of \( f \) is bounded.
   - (b) \( f \) has a minimum and maximum value.
   - (c) Every number between the minimum and maximum value is a value of \( f \).

2. However, the points \( u, v \) where \( f \) attains its minimum and maximum values may be anywhere inside the interval \([a, b]\), and it may well happen that \( v < u \).

Example 27

1. \([a, b] = [-1, 1]\) and
   \[
   f(x) = \begin{cases} 
   x, & x \leq 0, \\
   -2x, & x \geq 0.
   \end{cases}
   \]
   Here the image of \( f \) is the interval \([-2, 0]\) and \( f \) attains its minimum at 1 and its maximum at 0.

2. \( f \) is defined in \((0, 1)\) by \( f(x) = 1/x \). Here the values of \( f \) are not bounded above but are bounded below by 1; however, \( f \) does not attain a minimum value since
   \[
   \text{glb} \{f(x)\} = 1
   \]
   and 1 is not a value of \( f \).

While the properties above are consequences of Theorem 4 we actually have to prove them first and then use them to prove the Theorem! We do this now.

Proposition 17 Suppose \( f : D \to S \) is a continuous function defined in a closed interval \( D = [a, b] \). Then for some \( u, v \in [a, b] \),

\[
f(u) \leq f(x) \leq f(v)
\]

for all \( x \in [a, b] \).

Proof: Set \( S = \{f(x) \mid x \in [a, b]\} \). We show first by contradiction that \( S \) is bounded above.

In fact, if \( S \) is not bounded above then for each \( n \in \mathbb{N} \) there is an \( x_n \in [a, b] \) such that \( f(x_n) > n \). By Theorem 3 the sequence \((x_n)\) contains a convergent subsequence \((x_{n_i})\), and by Exercise 15 the limit \( z \) of this subsequence is in \([a, b]\). Since continuous functions preserve convergent sequences, (Lemma 15.2) the sequence \((f(x_{n_i}))\) is convergent. But

\[
f(x_{n_i}) > n_i
\]
and so this sequence is not bounded above. Therefore it cannot be convergent, which is a contradiction, and so $S$ must be bounded above.

Since $S$ is bounded above, for each $n$ there is some $y_n \in S$ such that

$$0 \leq \text{lub}(S) - y_n < \frac{1}{n}.$$  

In particular, $\lim_n y_n = \text{lub}(S)$.

On the other hand, because $y_n \in S$ we may write $y_n = f(x_n)$ for some $x_n \in [a,b]$. Then $(x_n)$ is a bounded sequence.

Therefore by Theorem 3 there is a subsequence $x_{n_k}$ converging to some point $v$. Since each $x_{n_k} \in [a,b]$, it follows from Exercise 15 that $v \in [a,b]$.

Because continuous functions preserve convergent sequences (Lemma 15.2) the sequence $f(x_{n_k})$ converges to $f(v)$.

But $f(x_{n_k})$ is a subsequence of the convergent sequence $f(x_n)$. Therefore by Exercise 15.3, this subsequence has the same limit. Therefore

$$f(v) = \lim_k f(x_{n_k}) = \lim_n f(x_n) = \lim_n y_n = \text{lub}(S).$$

In other words, for all $x \in [a,b]$,

$$f(x) \leq f(v).$$

The proof of the existence of $u$ is identical.

**q.e.d.**

**Definition 38** The points $u, v$ in Proposition 14 are called respectively an absolute minimum point and an absolute maximum point for $f$.

**Proposition 18** If $f: D \to S$ is a continuous function defined in an interval $D$ and if $f(x) < f(y)$ for some $x, y \in D$, then for any $w$ with $f(x) < w < f(y)$ there is some $z \in D$ with $f(z) = w$.

**Proof:** We argue by contradiction, and so suppose that there is no $z \in D$ with $f(z) = w$.

Let $d = |y - x|$. We construct by induction a sequence of intervals $I_n \subset D$, $n \geq 0$ with endpoints $x_n$ and $y_n$ such that

$$|y_n - x_n| = \frac{d}{2^n}, \text{ and}$$

$$f(x_n) < w < f(y_n).$$

(Note: it may happen that $x_n < y_n$ or that $y_n < x_n$!!)

First set $x_0 = x$ and $y_0 = y$.  

93
Now suppose $x_n$ and $y_n$ are constructed. Since $w$ is not a value, either
\[ f\left(\frac{x_n + y_n}{2}\right) < w \quad \text{or} \quad f\left(\frac{x_n + y_n}{2}\right) > w. \]

In the first case, set
\[ x_{n+1} = \frac{x_n + y_n}{2} \quad \text{and} \quad y_{n+1} = y_n. \]

In the second case, set
\[ x_{n+1} = x_n \quad \text{and} \quad y_{n+1} = \frac{x_n + y_n}{2}. \]

Now by Theorem 2, $\bigcap I_n$ is a single point $c$. Moreover, the Corollary to that Theorem states that
\[ \lim_{n \to \infty} x_n = c = \lim_{n \to \infty} y_n. \]

Since $f$ is continuous, it follows that
\[ \lim_{n \to \infty} f(x_n) = f(c) = \lim_{n \to \infty} f(y_n). \]

But $f(x_n) < w$ for all $n$, and thus $f(c) \leq w$. Also $f(y_n) > w$ for all $n$, and thus $f(c) \geq w$.

It follows that $f(c) = w$, contradicting our hypothesis that $w$ was not a value of $f$. q.e.d.

**Proof of Theorem 4**: By Proposition 17, $f$ has an absolute minimum at some $u \in [a, b]$ and an absolute maximum at some $v \in [a, b]$:
\[ f(u) \leq f(x) \leq f(v) \]
for all $x \in [a, b]$. If $w \in [f(u), f(v)]$, then either $w = f(u)$ or $w = f(v)$ or $f(u) < w < f(v)$. But in this third case it follows from Proposition 18 that $w = f(x)$ for some $x \in [a, b]$.

**Exercise 21**

In these exercises you may assume the standard properties of $\sin x$ and $\cos x$.

1. Let $f : D \to S$ be any real-valued function defined in an interval $D$.
   
   (a) Show that if $f$ attains a maximum then that maximum is $\text{lub}(S)$ where
   \[ S = \{ f(x) \mid x \in D \}. \]
   
   (b) If $f$ attains a minimum value, what is it?

2. (a) Use Exercise 18.5 to show that $\sin(1/x)$ is a continuous function in $(0, \infty)$.
(b) Show that this function attains a maximum and a minimum value.
(c) Show that it is not the restriction of a continuous function defined in $[0, \infty)$.

3. Show that $(1 - x)\sin(1/x)$ is a continuous function in $(0, 1]$ which does not attain a maximum or minimum value.

6.5 Exponential functions

Suppose $x > 1$ is real number. For every rational number $a$, a real number, $x^a$ is defined (25), and shown in Exercise 12, to have the following properties:

1. $x^0 = 1$ and $x^1 = x$.
2. If $a, b \in \mathbb{Q}$ and $a < b$ then
   
   \[0 < x^a < x^b.\]

3. $x^a x^b = x^{a+b}$ for any $a, b \in \mathbb{Q}$.
4. $(x^a)^b = x^{ab}$ for any $a, b \in \mathbb{Q}$.

We now extend the construction to all real exponents to define a function

$$ \exp_x : \mathbb{R} \to \mathbb{R} $$

as follows: For any $y \in \mathbb{R}$, set

$$ S(y) = \{ a \in \mathbb{Q} \mid a \geq y \} \quad \text{and} \quad E(y) = \{ x^a \mid a \in S(y) \}.$$

If $a \in \mathbb{Q}$ satisfies $a < y$ then $a$ is a lower bound for $S(y)$.

Thus it follows from Exercise 12 that $x^a$ is a lower bound for $E(y)$.

Now define $\exp_x$ by setting:

$$ \exp_x(y) = \inf(E(y)), \quad y \in \mathbb{R}. $$

**Theorem 5** Let $x > 1$ be a real number. Then the function $\exp_x : \mathbb{R} \to \mathbb{R}$ has the following properties:

1. $\exp_x(a) = x^a$ if $a$ is rational.

2. If $y < z$ are real numbers then

   \[0 < \exp_x(y) < \exp_x(z).\]

3. $\exp_x(y)\exp_x(z) = \exp_x(y + z)$ for all $y, z \in \mathbb{R}$.  

95
4. For all $u, y \in \mathbb{R}$
\[
\frac{|\exp_x(u) - \exp_x(y)|}{|u - y|} \leq x \exp_x(y) \quad \text{if} \quad 0 < |u - y| < 1.
\]

5. $\exp_x$ is a continuous function.

**Proof:** As usual, we prove the theorem taking the statements in order.

1. $\exp_x(a) = x^a$ if $a \in \mathbb{Q}$.

   If $a \in \mathbb{Q}$ then $S(a) = [a, \infty) \cap \mathbb{Q}$. Therefore in this case
   \[
   E(a) = \{x^b \mid b \geq a \text{ and } b \in \mathbb{Q}\}.
   \]
   Therefore
   \[
   \exp_x(a) = \inf(E(a)) = x^a.
   \]

2. If $y < z$ are real numbers then
   \[
   0 < \exp_x(y) < \exp_x(z).
   \]
   Choose rational numbers $a, b$ so that
   \[
   a < y < b < z.
   \]
   Then, as observed in the construction of $\exp_x$,
   \[
   x^a \leq E(y).
   \]
   Thus if $c$ is a rational number such that $a < c < y$ it follows from Exercise 12 that
   \[
   0 < x^a < x^c \leq \exp_x(y).
   \]
   In the same way it follows that
   \[
   x^b < \exp_x(z).
   \]
   Finally, since $y < b$ we have $b \in S(y)$ and $x^b \in E(y)$. Therefore
   \[
   0 < \exp_x(y) \leq x^b < \exp_x(z).
   \]

3. $\exp_x(y)\exp_x(z) = \exp_x(y + z)$ for all $y, z \in \mathbb{R}$.

   If $a \in S(y)$ and $b \in S(z)$ then $a + b \in S(x + y)$. It follows that for $a \in S(y)$ and $b \in S(z)$,
   \[
   \exp_x(y + z) = \inf(E(y + z)) \leq x^{a+b} = x^a x^b.
   \]
Therefore $\exp_z(y+z)$ is a lower bound for $E(y)x^b$. It follows that
\[ \exp_z(y+z) \leq \exp_z(y)x^b. \]
Repeating this argument yields
\[ \exp_z(y+z) \leq \exp_z(y)\exp_z(z). \]
On the other hand, suppose $c \in S(y+z)$. It follows from Exercise 8.4 that for some $a \in S(y)$ and some $b \in S(z)$,
\[ c \geq a + b. \]
Therefore
\[ x^c \geq x^{a+b} = x^a x^b \geq \exp_z(y)\exp_z(z). \]
Therefore $\exp_z(y)\exp_z(z)$ is a lower bound for $E(y+z)$ and so
\[ \exp_z(y)\exp_z(z) \leq \exp_z(y+z). \]
Combined with the previous inequality this proves the assertion.

4. For all $u, y \in \mathbb{R}$,
\[ \frac{|\exp_x(u) - \exp_x(y)|}{|u - y|} \leq x \exp_x(y) \text{ if } 0 < |u - y| < 1. \]

For this assertion we need two lemmas.

**Lemma 16** For any natural number $p$ and any positive real number $t$,
\[ (1 - pt)(1 + t)^p < 1. \]

**Proof:** We use induction on $p$.
When $p = 1$ the Lemma follows from
\[ (1 - t)(1 + t) = 1 - t^2 < 1. \]
Now suppose by induction that the Lemma is true for some $p$.
Then, since
\[ (1 - (p+1)t) < (1 - (p+1)t + pt^2) = (1 - t)(1 - pt) \]
it follows from our induction hypothesis that
\[ (1 - (p+1)t)(1 + t)^{p+1} < (1 - t)(1 - pt)(1 + t)^p(1 + t) < (1 - t)(1 + t) < 1, \]
which closes the induction and completes the proof of the Lemma.

q.e.d.
Lemma 17 Suppose a and b are rational numbers and x > 1 is a real number. If 0 < a < b, then

$$\frac{x^a - 1}{a} < \frac{x^b - 1}{b}.$$

Proof: Since both a and b are rational and positive we may find natural numbers k, m, n such that

$$a = \frac{m}{k} \quad \text{and} \quad b = \frac{n}{k}.$$ 

Since a < b it follows that m < n. Therefore it is sufficient to consider the case that n = m + 1 and prove that

$$\frac{x^{m/k} - 1}{m/k} < \frac{x^{(m+1)/k} - 1}{(m+1)/k}.$$ 

Set $w = x^{1/k}$. Then multiply both sides by 1/k to see that it is enough to show that

$$\frac{w^m - 1}{m} < \frac{w^{m+1} - 1}{m+1},$$

or, equivalently,

$$(m + 1)(w^m - 1) < m(w^{m+1} - 1).$$

Since $x > 1$ it follows from Exercise 12 that $w > 1$. Write $w = 1 + t$ with $t > 0$.

Now Lemma 16 gives

$$(m + 1 - mw)w^m = (1 - mt)(1 + t)^m < 1.$$ 

It follows that

$$(m + 1)w^m < 1 + mw^{m+1},$$

and therefore

$$(m + 1)(w^m - 1) < 1 + mw^{m+1} - (m + 1) = m(w^{m+1} - 1),$$

which is the inequality we needed to prove. q.e.d.

We now return to the proof of the statement: For all $u, y \in \mathbb{R}$

$$\frac{|\exp_x(u) - \exp_x(y)|}{|u - y|} \leq x\exp_x(y) \quad \text{if} \quad 0 < |u - y| < 1.$$
Suppose $0 < h < 1$ is a rational number. Apply Lemma 17 with $a = h$ and $b = 1$ to obtain

$$\frac{\exp_x(y + h) - \exp_x(y)}{h} = \exp_x(y) \frac{\exp_x(h) - 1}{h} < \exp_x(y) \frac{\exp_x(1) - 1}{1} < x \exp_x(y),$$

and

$$\frac{\exp_x(y) - \exp_x(y - h)}{h} = \exp_x(y-h) \frac{\exp_x(h) - 1}{h} < \exp_x(y) \frac{\exp_x(1) - 1}{1} < x \exp_x(y).$$

Now if $|u - y| = t < h < 1$, then it follows from the second assertion of the Theorem that

$$0 < \exp_x(y - h) < \exp_x(y - t) < \exp_x(y) < \exp_x(y + t) < \exp_x(y + h).$$

Therefore

$$|\exp_x(u) - \exp_x(y)| < h x \exp_x(y).$$

It follows that

$$\frac{|\exp_x(u) - \exp_x(y)|}{x \exp_x(y)} < h$$

for every rational $h$ such that $h \geq h < 1$. Therefore

$$|\exp_x(u) - \exp_x(y)| \leq |u - y| x \exp_x(y),$$

which proves the assertion.

5. $\exp_x$ is a continuous function.

Fix $y \in \mathbb{R}$ and $\varepsilon > 0$. Then choose $0 < \delta < 1$ so that

$$\delta < \frac{\varepsilon}{x \exp_x(y)}.$$

Then if $|u - y| < \delta$,

$$|\exp_x(u) - \exp_x(y)| < \delta x \exp_x(y) < \varepsilon.$$

Therefore $\exp_x$ is continuous.

q.e.d.

**Definition 39** A real-valued function $f$ in an interval $D$ is called **increasing** (resp., **decreasing**) if whenever $x < y$ it follows that $f(x) \leq f(y)$ (resp., $f(x) \geq f(y)$). In either case $f$ is called a **monotone** function.

**Exercise 22**
1. Show that if a continuous function in \([a, b]\) is a 1-1 map it is monotone.

2. For \(0 < x \leq 1\) define a function \(\exp_x\) by setting \(\exp_1(y) = 1, y \in \mathbb{R}\) and

\[
\exp_x(y) = \frac{1}{x^y}
\]

if \(0 < x < 1\) and \(y \in \mathbb{R}\). Then show for \(0 < x < 1\) that \(\exp_x\) satisfies statements 1,3,5 of Theorem 5.

3. If \(0 < x < 1\) show that \(\exp_x\) is a monotone function.

4. If \(x > 0\) and \(y, z\) are any real numbers show that

\[
(x^y)^z = x^{yz}.
\]
Chapter 7

Derivatives

Recall that the graph of a real function $f$ defined in an interval $D$ is the set of points $\{(x, f(x)) \mid x \in D\}$. While every such function has a graph, the function with domain $[0, 1]$ given by

$$f(x) = \begin{cases} 
0, & x \in \mathbb{Q}, \\
1, & x \notin \mathbb{Q}.
\end{cases}$$

has a graph which cannot be drawn in any reasonable way.

By contrast continuous functions have graphs that are often reasonable curves, but can have corners that make them jagged, such as in the example below:

If we want smoother graphs we need to restrict to differentiable functions. These then turn out to have other really nice properties which are essential for much of analysis and for many applications.

The idea is very simple: a real-valued function defined in an interval $D$ is differentiable if it has a derivative at every point in $D$, and it has a derivative at a point $x \in D$ if its graph has a non-vertical tangent line at $(x, f(x))$. This eliminates the possibility of corners! The derivative is then the slope of the tangent line.

We have to formalize this intuitive idea. If $y$ is a point near $x$ the secant determined by $x$ and $y$ is the straight line through the points $(y, f(y))$ and $(x, f(x))$. Thus the slope of the secant is

$$\frac{f(y) - f(x)}{y - x}.$$
If the graph has a tangent line then the secants should approach the tangent line as \( y \) approaches \( x \) and so the slopes of the secants should approach the slope of the tangent line. We have thus finally developed our intuitive idea to the point where we can formalize it:

**Definition 40** Let \( f \) be a real-valued function defined in an interval, \( D \). Then \( f \) is **differentiable** at \( x \in D \) if

\[
\lim_{y \to x} \frac{f(y) - f(x)}{y - x}
\]

exists. In this case we call the limit the derivative of \( f \) at \( x \), and denote it by \( f'(x) \).

**Definition 41** A real-valued function defined in an interval \( D \) is **differentiable** if it is differentiable at every point in \( D \).

**Proposition 19** A differentiable function is continuous.

**Proof:** Suppose \( f \) is a differentiable function in an interval \( D \). If \( a \in D \) then by definition

\[
\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a).
\]

Thus for any \( \omega > 0 \) there is some \( \eta \) such that

\[
|\frac{f(x) - f(a)}{x - a} - f'(a)| < \omega \quad \text{if} \quad 0 < |x - a| < \eta.
\]

It follows that

\[
|\frac{f(x) - f(a)}{x - a}| = |\frac{f(x) - f(a)}{x - a} - f'(a) + f'(a)| < \omega + |f'(a)| \quad \text{if} \quad 0 < |x - a| < \eta.
\]

Multiply by \( |x - a| \) to obtain

\[
|f(x) - f(a)| < (\omega + |f'(a)|)|x - a| \quad \text{if} \quad 0 < |x - a| < \eta.
\]

Now if \( \epsilon > 0 \), set

\[
\delta = \min\left(\frac{\epsilon}{\omega + |f'(a)|}, \eta\right).
\]

Then for \( 0 < |x - a| < \delta \) we have \( 0 < |x - a| < \eta \) and so

\[
|f(x) - f(a)| < (\omega + |f'(a)|)|x - a| < (\omega + |f'(a)|)\frac{\epsilon}{\omega + |f'(a)|} < \epsilon.
\]

Thus \( \lim_{x \to a} f(x) = a \). q.e.d.

**Remarks:** The derivative at a point \( x \) of a function \( f \) can also be thought of as its instantaneous rate of change. In fact, the formula above for the slope of a secant shows that that slope is the "average" rate of change in \( f \) between \( y \) and \( x \). Thus it seems reasonable to interpret the limit as the instantaneous
rate of change. Magnificently, as we shall see, this actually works in practice and makes the derivative of central importance for applications.

Observe that if $f$ is a differentiable function in an interval $D$, then a new real-valued function in $D$ is given by

$$x \mapsto f'(x).$$

This function is called the derivative of $f$ and is denoted by $f'$ or by $df/dx$. If $f'$ is continuous the $f$ is continuously differentiable; if $f'$ is itself differentiable then $f$ is twice differentiable and the derivative of $f'$ is written $f''$ and called the second derivative of $f$.

Remark: The next Proposition shows how the search for a maximum or a minimum can be reduced to solving the equation $f'(c) = 0$. This is, of course, of fundamental importance for many applications where the location of maxima and minima is important.

**Proposition 20** Suppose $f$ is a real-valued function in an interval $D$ with a derivative at a point $c \in D$. If for some $a < c < b$, $(a, b) \subset D$, and if either

$$f(x) \leq f(c), \quad x \in (a, b),$$

or else

$$f(x) \geq f(c), \quad x \in (a, b),$$

then $f'(c) = 0$.

**Proof:** Suppose $f(x) \leq f(c), \ x \in (a, b)$. If $y \in (a, c)$ then

$$\frac{f(y) - f(c)}{y - c} \geq 0$$

and so taking the limit as $y$ approaches $c$ from the left we find that $f'(c) \geq 0$. On the other hand, if $y \in (c, b)$ then

$$\frac{f(y) - f(c)}{y - c} \leq 0$$

and so $f'(c) \leq 0$. Thus $f'(c) = 0$.

The proof when $f(x) \geq f(c), \ x \in (a, b)$ is identical.

q.e.d.

This Proposition is also the main step in the proof of the important

**Theorem 6 (Mean Value Theorem).** Suppose $f$ is a differentiable function in an interval $[a, b]$. Then for some $z$ with $a < z < b$,

$$f'(z) = \frac{f(b) - f(a)}{b - a}.$$
We first consider the special case that \( f(a) = f(b) \).

**Proposition 21 (Rolles Theorem)** Suppose \( f \) is a differentiable function in an interval \([a, b]\) and \( f(a) = f(b) \). Then for some \( z \in (a, b) \),

\[
f'(z) = 0.
\]

**Proof:** According to Proposition 19, \( f \) is continuous. Thus by Proposition 17 there are points \( u, v \) in \([a, b]\) such that

\[
f(u) \leq f(x) \leq f(v)
\]

for all \( x \in [a, b] \).

Now if both \( u \) and \( v \) are endpoints of \([a, b]\) then \( f(u) = f(v) \), and it follows that \( f \) is constant. In this case \( f' \) is identically zero, as is immediate from the definition of derivative.

Otherwise either \( u \) or \( v \) is a point \( y \in (a, b) \). But now Proposition 20 asserts that \( f'(y) = 0 \). q.e.d.

**Proof of Theorem 6** Define a new function \( g \) in \([a, b]\) by

\[
g(x) = f(x) - \frac{x-a}{b-a}(f(b) - f(a)).
\]

It is straightforward to verify (see Exercise 21 below) that \( g \) is differentiable and that

\[
g'(x) = f'(x) - \frac{f(b) - f(a)}{b-a}.
\]

But it is immediate from the definition of \( g \) that \( g(a) = f(a) = g(b) \). Thus by Rolles Theorem, \( g'(z) = 0 \) for some \( z \in (a, b) \). Now the formula above for \( g'(x) \) shows that

\[
f'(z) = \frac{f(b) - f(a)}{b-a}.
\]

This completes the proof. q.e.d.

**Exercise 23 Derivatives**

1. Show that the sum and product of differentiable functions is differentiable, and establish the formulae for the derivatives.

2. If \( f \) is a continuous function in an interval \([a, b]\) with a derivative at \( c \in (a, b) \), show that a continuous function \( g \) in \([a, b]\) is defined by

\[
g(x) = \begin{cases} 
  \frac{f(x) - f(c)}{x-c} & \text{if } x \neq c, \\
  f'(c) & \text{if } x = c.
\end{cases}
\]
3. Show that if $f$ is a differentiable function in $(a, b)$ and if $f'(x) > 0$ for all $x \in (a, b)$, then for any $a < x < y < b$, $f(x) < f(y)$.

4. If $f$ is a differentiable function in $[a, b]$ and $f'(x) = 0$ for all $x \in (a, b)$, show that $f(x) = f(y)$ for all $x, y \in (a, b)$. Then use continuity to conclude that $f(x) = f(a)$ for all $x \in [a, b]$.

5. Suppose $f$ and $g$ are differentiable functions in an interval $[a, b]$. If $f(a) = g(a)$ and if $f'(x) = g'(x)$ for all $x \in (a, b)$, show that $f(x) = g(x)$ for all $x \in [a, b]$. 


Chapter 8

Area and Integrals

All the way back in the distant mists of public school we learned that the area of a rectangle is its length times its width. So how do we go about defining the area of a more complicated region, \( R \), contained inside a large rectangle in the plane?

One very simple way works as follows: first, grid the big rectangle into non-overlapping smaller ones. Then add the areas of the little rectangles that touch \( R \) and call that an upper area. Next, add the areas of the little rectangles entirely contained in \( R \), and call that a lower area.

Intuitively, every upper area should be bigger than the area of \( R \) and every lower area would be less than the area of \( R \). Thus, if we set

\[
U = \text{\{upper areas\}} \quad \text{and} \quad L = \text{\{lower areas\}},
\]

then we would expect that

\[
\text{glb}(U) \geq \text{area}(R) \geq \text{lub}(L).
\]

If we get lucky and it happens that \( \text{inf}(U) = \text{sup}(L) \) we would define this number to be the area of \( R \).

If we are unlucky then there would be two possibilities: either we need to find a different method, or else it simply is not possible to assign an area to \( R \).

The first situation is exemplified by the region \( R \) bounded by the \( x \)-axis, the \( y \)-axis, the line \( x = 1 \) and the graph of the function \( f \) defined by

\[
f(x) = \begin{cases} 
0, & x \in \mathbb{Q} \\
1, & x \notin \mathbb{Q}.
\end{cases}
\]

Here our method does not work, but there is, in fact, a more sophisticated way of defining area, and that method gives this region an area of one.
There are however, even more complicated regions for which there is no reasonable way to assign an area!

Nonetheless our method does work for the region between the graph of a continuous function and the x-axis, except that we have to count the area below the x-axis negatively and it is this 'signed' area we use to define the integral. We proceed to formalize this intuitive idea.

**Definition 42 Upper and lower sums.**

1. A partition of an interval \([a, x]\) is a sequence
   
   \[ a = x_0 < x_1 < x_2 < \cdots < x_r = x. \]

2. If \(f\) is a real-valued function defined in an interval \([a, x]\), then an upper sum, \(u\), and a lower sum, \(l\), for \(f\) are numbers respectively of the form
   
   \[ u = \sum_{i=1}^{r} y_i (x_i - x_{i-1}) \]
   
   and
   
   \[ l = \sum_{i=1}^{r} w_i (x_i - x_{i-1}) \]

   where the \(x_i\) are a partition of \([a, x]\) and

   \[ w_i \leq f([x_{i-1}, x_i]) \leq y_i. \]

**Proposition 22** Suppose \(f\) is a continuous function in an interval \([a, x]\), and let \(U\) and \(L\) denote the set of upper sums and the set of lower sums for \(f\). Then

\[ \operatorname{glb}(U) = \operatorname{lub}(L). \]

**Proof:** We proceed in two steps.

**Step One:** Every upper sum is larger than every lower sum.

Suppose an upper sum is given by

\[ u = \sum_{i=1}^{r} y_i (x_i - x_{i-1}) \]

and a lower sum is given by

\[ l = \sum_{j=1}^{p} w_j (x'_j - x'_{j-1}) \]

defined respectively by partitions \((x_i)\) and \((x'_j)\). Then we can "merge" the two partitions to get a partition \(x''_k\), \(0 \leq k \leq q\) in which each \(x''_k\) is either an \(x_i\) or an
and each $x_i$ and $x_j$ is one of the $x''_k$. Call the intervals $[x''_{k-1}, x''_k]$ the “little” intervals.

In particular, each interval $[x_{i-1}, x_i]$ is a union of little intervals. Set

$$y''_k = y_i$$

if $[x''_{k-1}, x''_k] \subset [x_{i-1}, x_i]$. Then

$$u = \sum_{k=1}^q y''_k (x''_k - x''_{k-1}).$$

Now define $w''_k$ in the same way and observe that

$$l = \sum_{k=1}^q w''_k (x''_k - x''_{k-1}).$$

Finally, recall that each little interval is contained in some $[x_{i-1}, x_i]$ and in some $[y_{j-1}, y_j]$. It follows that

$$w''_k \leq f([x''_{k-1}, x''_k]) \leq y''_k,$$

and so

$$l = \sum_{k=1}^q w''_k (x''_k - x''_{k-1}) \leq \sum_{k=1}^q y''_k (x''_k - x''_{k-1}) \leq u.$$

**Step Two**: Completion of the proof.

It follows from Step One that $lub(L) \leq glb(U)$. Now recall from Property 3 that for any $\delta > 0$ there is some $r \in \mathbb{N}$ such that $1/r < \delta$. Thus it follows from Proposition 16 that given $\varepsilon > 0$ there is some $r \in \mathbb{N}$ such that for all $s, t \in [a, x]$,

$$|f(s) - f(t)| < \frac{\varepsilon}{x - a} \quad \text{if} \quad |s - t| < \frac{x - a}{r}.$$

Now divide $[a, x]$ into $r$ intervals of even length $(x - a)/r$ with endpoints

$$a = x_0 < x_1 < x_2 < \cdots < x_r = x.$$

Then the restriction of $f$ to a subinterval $[x_{i-1}, x_i]$ has a maximum value $y_i$ and a minimum value $w_i$, which it assumes at points $z$ and $z'$ in the subinterval (Proposition 17). Thus the inequality above implies that

$$y_i - w_i < \frac{\varepsilon}{x - a}.$$

Therefore, with these choices the corresponding upper and lower sums satisfy

$$u - l = \sum_{i=1}^r (y_i - w_i)(x_i - x_{i-1}) < \frac{\varepsilon}{x - a} \sum_{i=1}^r (x_i - x_{i-1}) = \varepsilon.$$
It follows that $\text{glb}(U) - \text{lub}(L) < \varepsilon$ for all $\varepsilon > 0$, and so $\text{glb}(U) = \text{lub}(L)$.

q.e.d.

**Definition 43** Let $f$ be a continuous function in an interval $[a, x]$. Then the number $\text{glb}(U) = \text{lub}(L)$ of Proposition 22 is called the integral of $f$ and is denoted by

$$\int_a^x f(t)dt.$$ 

**Exercise 24 Integrals**

1. If $f$ is a continuous function in an interval $[a, b]$ with maximum value $y$ and minimum value $w$, show that

$$w(b - a) \leq \int_a^b f(t)dt \leq y(b - a).$$

2. If $f$ is a continuous function in an interval $[a, b]$ show that

$$\int_a^a f(t)dt = 0 \quad \text{and for} \quad c \in (a, b), \quad \int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$ 

3. If $f$ and $g$ are continuous functions in an interval $[a, b]$ and if $\lambda$ and $\mu$ are real numbers show that

$$\int_a^b (\lambda f(t) + \mu g(t))dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt.$$ 

4. If $f$ and $g$ are continuous functions in an interval $[a, b]$, and if $f(t) \leq g(t)$ for all $t \in [a, b]$, show that

$$\int_a^b f(t)dt \leq \int_a^b g(t)dt.$$ 

If the two integrals are equal show that $f(t) = g(t)$ for all $t \in [a, b]$.

**Theorem 7** Let $f$ be a continuous function defined in an interval $[a, b]$. Then

$$F(x) = \int_a^x f(t)dt$$

is a differentiable function in $[a, b]$, and $F'(x) = f(x)$.

**Proof:** Fix $c \in [a, b]$. We have to show that

$$\lim_{s \to c} \frac{F(s) - F(c)}{s - c} = f(c),$$
and we consider separately the cases where \( s \) approaches \( c \) from the right and from the left. Consider first the case of an approach from the right.

For \( s > c \) let \( y(s) \) and \( w(s) \) be the maximum and minimum values assumed by \( f \) in \( c, s \]. By Exercise 22.2,

\[
\int_a^s f(t)dt - \int_a^c f(t)dt = \int_c^s f(t)dt.
\]

Therefore by Exercise 22.1,

\[
w(s) - f(c) \leq \frac{\int_c^s f(t)dt}{s - c} - f(c) \leq y(s) - f(c).
\]

It follows that

\[
w(s) - f(c) \leq \frac{\int_a^s f(t)dt - \int_a^c f(t)dt}{s - c} - f(c) \leq y(s) - f(c).
\]

In other words,

\[
w(s) - f(c) \leq \frac{F(s) - F(c)}{s - c} - f(c) \leq y(s) - f(c).
\]

Now as \( s \) approaches \( c \) both \( w(s) - f(c) \) and \( y(s) - f(c) \) approach zero, and so

\[
\frac{F(s) - F(c)}{s - c}
\]

approaches \( f(c) \). The same argument establishes this if \( s \) approaches \( C \) from the left. q.e.d.

We can now establish the

**Theorem 8 (Fundamental Theorem of the Calculus)** Suppose \( f \) is a continuously differentiable function in \([a, b]\). Then

\[
\int_a^b f'(t)dt = f(b) - f(a).
\]

**Proof:** In fact, we shall prove that for all \( x \in [a, b] \),

\[
\int_a^x f'(t)dt = f(x) - f(a).
\]

Now the right hand side defines a function \( g(x) \) which by Theorem 7 is differentiable and satisfies

\[
g'(x) = f'(x), \quad x \in [a, b].
\]
Thus the derivative of $g - f$ is identically zero, and so by Exercise 18.4 it follows that $g - f$ is a constant function. But $g(a) = 0$ and so $g(x) - f(x) = -f(a)$ for all $x \in [a, b]$. In other words,

$$g(x) = f(x) - f(a)$$

for all $x \in [a, b]$. q.e.d.

**Remark:** As you will recall from your calculus classes, while derivatives tend to be computable it is Theorem 8 which makes it possible to find formulæ for integrals!

Recall the Mean Value theorem, which states that if $f : [a, b] \to \mathbb{R}$ is differentiable then for $x \in [a, b]$ then for some $z \in [a, b]$,

$$f(x) = f(a) + f'(z)(x - a).$$

We generalize this in the

**Theorem 9** Suppose $f : [a, b] \to \mathbb{R}$ has $n + 1$ derivatives $f^{(k)}$, $1 \leq k \leq n + 1$, and that $f^{(n+1)}$ is continuous. The for each $x \in [a, b]$ there is some $z \in [a, b]$ such that

$$f(x) = \sum_{k=0}^{n} f^{(k)}(x-a)^{k} + f^{(n+1)}(z)(x-a)^{n+1}$$

**Proof:** When $n = 0$ this is just the Mean Value Theorem. Suppose it holds for some $n$, and that $f$ satisfies the hypotheses for $n + 1$. Then the derivative, $f'$ satisfies the hypotheses for $n$. Since $f'^{k} = f^{k+1}$, our induction hypothesis gives

$$f'(x) = \sum_{k=0}^{n} f^{(k+1)}(x-a)^{k} + f^{(n+2)}(z)(x-a)^{n+1}.$$  

We have assumed that $f^{(n+2)}$ is continuous. Thus it has an absolute minimum value $c$ and an absolute maximum value $d$ in $[a, b]$. In particular, from the equation above we obtain

$$c(x-a)^{n+1} \leq f'(x) - \sum_{k=0}^{n} f^{(k+1)}(x-a)^{k} \leq d(x-a)^{n+1}.$$

Now these inequalities hold for every $x \in [a, b]$, and so

$$\int_{a}^{x} c \frac{(t-a)^{n+1}}{(n+1)!} \leq \int_{a}^{x} f'(t) - \sum_{k=0}^{n} f^{(k+1)} \frac{(t-a)^{k}}{k!} \leq \int_{a}^{x} d \frac{(x-a)^{n+1}}{(n+1)!}.$$  

Apply the Fundamental Theorem of the Calculus to obtain

$$c \frac{(x-a)^{n+2}}{(n+2)!} \leq f(x) - f(a) - \sum_{k=0}^{n} f^{(k+1)} \frac{(x-a)^{k+1}}{(k+1)!} \leq d \frac{(x-a)^{n+2}}{(n+2)!}.$$  

111
Finally, the Intermediate Value Theorem for $f^{(n+2)}$ implies that for some $z \in [a, b],$

$$f(x) - f(a) - \sum_{k=0}^{n} f^{(k+1)}(x-a)^{k+1}(x-a) = f^{(n+2)}(z)(x-a)^{n+2}(n+2)!. $$

Since

$$f(a) + \sum_{k=0}^{n} f^{(k+1)}(x-a)^{k+1}(x-a) = f(a) + \sum_{k=1}^{n} f^{(k)}(x-a)^{k}(k)!,$$

the formula is established for $n + 1$. q.e.d.