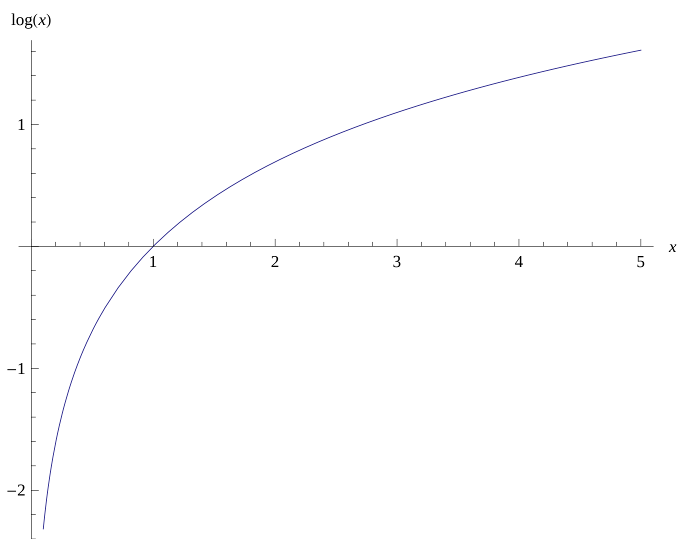
ADVANCED CALCULUS

Lecture notes for
MA 440/540 & 441/541

Rudi Weikard



Based on lecture notes by G. Stolz and G. Weinstein
Version of August 3, 2022

Contents

First things first

The goals

Our goal in this class is threefold:

(1) to obtain a body of knowledge in Advanced Calculus, the basis of the analysis of real-valued functions of one real variable;
(2) to learn how to communicate ideas and facts in both a written and an oral form;
(3) and, perhaps most importantly, to become acquainted with — indeed, to master — the process of creating mathematics.

In conducting this class we shall try to model a mathematical community in which both collaboration and competition are prevalent. This community is — no, you are — on the verge of discovering the foundations for a number of rules and recipes which have been successfully in use for some time. In the process you will recreate a body of knowledge almost as if you were the first to discover it. However, as we have only nine months to do this rather than a century or two, there will be some help available to you, most prominently in the form of these notes which will delineate broadly a path in which discovery will (or could) proceed.

In this course it is allowed and, in fact, required to criticize the person on the board for flaws or incomplete arguments (you are a scientific community). Criticism has to be leveled in a professional manner, in particular, it has to be free from any personal insults. At the same time you have to learn to accept criticism without taking it personally. By learning to stand up for your ideas (or to accept that you made a mistake) you may get something out of this course which is of value not only in mathematics.

The rules

The following rules, based on intellectual and academic honesty, will be in force.

(1) Everybody will have the opportunity to present proofs of theorems. You will have the proof written out on paper and present it with the help of a document camera.
(2) The audience (including the instructor) may challenge a statement made in the course of the proof at any point.
(3) If the presenter is able to defend the challenged statement, he or she proceeds; if not, the presenter must sit down earning no points for this problem and losing the right to present again that day. The challenger may present his or her solution or elect to receive a challenge reward (see rules (10) and (12)).
(4) A proof of a theorem will be considered correct if no one has objections (or further objections). Its written version will then be "published" by uploading it to Canvas (it should have a title and the list of authors). The presenter and, if applicable, his or her collaborators (see rule (9)) will earn a total of 10 points at this time.

(5) During class the instructor has the final decision on determining whether an argument may stand or not. His verdict may still be challenged after a proof is "published" (see rule (6)).
(6) If someone other than an author discovers a flaw in a "published" proof, he or she will get the opportunity to explain the mistake and present a correct proof for a total of 20 points.
(7) While presenting proofs you may only refer to those axioms and theorems in the notes which occur before the one you are working on, to published proofs of such theorems, to the definitions, and to the appendix.
(8) You must give credit where credit is due, i.e., during your presentation you must declare the points at which you had help and by whom.
(9) It is also possible to report joint work. In such a case 4 points will be earned for the presentation while the other 6 are evenly distributed among the collaborators.
(10) The successor of a presenter will be chosen as the student with the smallest number of points among the volunteers taking into account the modification by rules (3), (11) and (12). A die is rolled, if necessary.
(11) You may volunteer for a particular problem by an e-mail to me. This (in the order received) establishes priority among volunteers with the same number of points.
(12) For a student who has earned a challenge reward 20 points will be subtracted from his or her current score for the purpose of determining a successor. At the time such a student is selected to prove a theorem the challenge reward expires.
(13) Class attendance and participation is required. Absences from class are recorded in Canvas. After 10 absences from class 10 points will be subtracted from your class score and the count of absences is set again to zero.
(14) There will be no partial credit except as described above to share credit.

Hints

- As there is no partial credit, be well prepared to answer whatever questions may arise.
- There is no need to take notes, final proofs will be "published". Think along instead.
- Try to earn points early.
- The index at the end of these notes may lead you to some of the necessary definitions (let me know of any omissions).
- Often a proof will be fairly simple once one realizes that a previous result or a previous method of proof can be used. Thus constant participation in class, even when it's other people's turn to present, is highly advised! You should go over the proofs again at home and see if they are still clear to you, possibly rewriting them with more details.
- Mathematical reasoning takes time. You may expect some frustration — without it there would never be a sense of accomplishment. Plan to spend a lot of time thinking about a problem before writing down a final solution.
- Exercises are for private study. Points are only earned for proofs of theorems.
- More difficult problems are marked with a $*$.
- **See me if you need help!**

The language of mathematics

Mathematics is a highly formalized subject. In many ways learning it is similar to learning a foreign language. This language rests on two (tightly interwoven) pillars: logic and set theory. In this course we assume that you have a basic familiarity (possibly unconsciously) with both. Nevertheless Appendix A collects some fundamental material about set theory which you may (and will have to) use in your proofs. If you like you may also find some information in my algebra notes which are on my website. There, at the ends of Sections 1.2 and 1.3, are references to books on these subjects.

Statements (true or false) are the bread and butter of mathematics. The more important kind of statements are the following:

- **Axioms:** In a given mathematical theory some statements are taken for granted. Such statements are called axioms. The axioms, in fact, characterize the theory. Changing an axiom means to consider a different theory. Euclidean geometry, for instance, relies on five axioms. One of them is the axiom of parallels. After trying for centuries to infer the parallel axiom from the other axioms mathematicians of the nineteenth century developed non-Euclidean geometries in which the parallel axiom is replaced by something else.
- **Definitions:** A mathematical definition specifies the meaning of a word or phrase leaving no ambiguity. It may be considered an abbreviation. For instance, the statement "A prime number is a natural number larger than 1 such that if it divides a product of two natural numbers it divides one of the factors." defines the word prime number.
- **Theorems:** A theorem is a true statement of a mathematical theory requiring proof. It is usually of the form "p implies q". For example the theorem "If n is even, then n^2 is divisible by 4." is of this form. Sometimes, when a statement hinges only on the axioms, the theorem could simply be something like "2 is a prime number.".

Many authors also use the words proposition, lemma, and corollary. Logically these are theorems and we will not use these words (but you should still look them up in a dictionary if you ever want to talk to other people).

CHAPTER 1

The real numbers

We will introduce the real numbers $\mathbb{R}$ by a series of axioms, namely the *field axioms*, the *order axioms*, and the *least upper bound axiom*. This means we will be sure of the precise properties of the real numbers. Of course, you probably already have some intuition as to what real numbers are, and these axioms are not meant to substitute for that intuition. However, when writing your proofs, you should make sure that all your statements follow from these axioms or their consequences proved previously. You will be amazed about the rich world being created — by you — from these axioms.

1.1. Field axioms

Definition 1. A *binary operation* on a set A is a function from $A \times A$ to A. It is customary to express a binary operation as $a \star b$ (or with other symbols in place of $\star$). An element $e \in A$ is called an *identity* if $e \star a = a \star e = a$ for all $a \in A$. An element $b \in A$ is called an *inverse* of $a \in A$ if $a \star b = b \star a = e$, assuming e is an identity.

Axiom 1 (Field axioms). For each pair $x, y \in \mathbb{R}$, there is a unique element denoted $x + y \in \mathbb{R}$, called the *sum* of x and y, such that the following *axioms of addition* are satisfied:

- **(A1):** $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathbb{R}$ (*associative law of addition*).
- **(A2):** $x + y = y + x$ for all $x, y \in \mathbb{R}$ (*commutative law of addition*).
- **(A3):** There exists an *additive identity* $0 \in \mathbb{R}$ (in other words: $x + 0 = x$ for all $x \in \mathbb{R}$).
- **(A4):** Each $x \in \mathbb{R}$ has an *additive inverse*, i.e., an inverse with respect to addition.

Furthermore, for each pair $x, y \in \mathbb{R}$, there is a unique element denoted by $x \cdot y \in \mathbb{R}$ (or simply xy), called the *product* of x and y, such that the following *axioms of multiplication* are satisfied:

- **(M1):** $(xy)z = x(yz)$ for all $x, y, z \in \mathbb{R}$ (*associative law of multiplication*).
- **(M2):** $xy = yx$ for all $x, y \in \mathbb{R}$ (*commutative law of multiplication*).
- **(M3):** There exists a *multiplicative identity* $1 \in \mathbb{R} \setminus \{0\}$ (in other words: $1 \cdot x = x$ for all $x \in \mathbb{R}$).
- **(M4):** Each $x \in \mathbb{R} \setminus \{0\}$ has a *multiplicative inverse*, i.e., an inverse with respect to multiplication.

Finally, multiplication and addition satisfy the *distributive law*:

- **(D):** $(x + y)z = xz + yz$ for all $x, y, z \in \mathbb{R}$.

In formulating the distributive law we have made use of the convention to let multiplication take precedence over addition, i.e., $x + yz$ is short for $x + (yz)$.

Any set which satisfies all the above axioms is called a field. Thus $\mathbb{R}$, but also the set of rational numbers with which you are familiar, are fields.

Exercise 1. Show that one may define binary operations in the set $\{0,1\}$ which turn it into a field. (This field is called $\mathbb{Z}_2$.)

Theorem 1. The additive identity in $\mathbb{R}$ is unique.

Exercise 2. Is -0 an additive identity?

Theorem 2. Every real number x has a unique additive inverse.

The unique additive inverse of $x \in \mathbb{R}$ is called the *negative* of x and is denoted by $-x$. For simplicity we will usually write $x - y$ in place of $x + (-y)$.

Theorem 3. $x + y = x + z$ if and only if $y = z$, assuming that $x, y, z \in \mathbb{R}$.

Theorem 4. Suppose $x, y \in \mathbb{R}$. Then the following two statements hold.
(1) If $x + y = x$ then $y = 0$.
(2) If $x + y = 0$ then $y = -x$.

Theorem 5. $-(-x) = x$ for all $x \in \mathbb{R}$.

Theorem 6. There is a unique multiplicative identity in $\mathbb{R}$ and every non-zero real number x has a unique multiplicative inverse.

The unique multiplicative inverse of $x \in \mathbb{R} \setminus \{0\}$ is called the *reciprocal* of x and is denoted by x^{-1}. We will also use the notation $\frac{1}{x}$ and $1/x$ in place of x^{-1} and we will usually write $\frac{x}{y}$ or x/y for xy^{-1}.

Theorem 7. If $x, y, z \in \mathbb{R}$ and $x \neq 0$, then the following statements hold.
(1) $xy = xz$ if and only if $y = z$.
(2) If $xy = x$ then $y = 1$.
(3) If $xy = 1$ then $y = x^{-1}$.
(4) $(x^{-1})^{-1} = x$.

Theorem 8. For every $x \in \mathbb{R}$ we have $0x = 0$.

Theorem 9. 0 does not have a reciprocal and neither is it the reciprocal of any number.

Theorem 10. If x and y are non-zero real numbers, then $xy \neq 0$.

Theorem 11. Let $x, y \in \mathbb{R}$. Then
(1) $(-1)x = -x$,
(2) $(-x)y = -(xy) = x(-y)$, and
(3) $(-x)(-y) = xy$.

Theorem 12. Let $a, b \in \mathbb{R}$ and $x, y \in \mathbb{R} \setminus \{0\}$. Then $a/x + b/y = (ay + bx)/(xy)$.

1.2. Order axioms

We will now state the next set of axioms for the real numbers. We use the following notation: If A is a subset of $\mathbb{R}$ then

$$-A = \{x \in \mathbb{R} : -x \in A\}.$$

Exercise 3. Convince yourself that $a \in -A$ if and only if $-a \in A$.

Theorem 13. Suppose A is a subset of $\mathbb{R}$. Then $-(-A) = A$.

Axiom 2 (Order axioms). The set of real numbers $\mathbb{R}$ is an ordered field. This means that (in addition to being a field) $\mathbb{R}$ has the following property: There is a set $\mathbb{P} \subset \mathbb{R}$ such that

(O1): $-\mathbb{P} \cap \mathbb{P} = \emptyset$,
(O2): $-\mathbb{P} \cup \{0\} \cup \mathbb{P} = \mathbb{R}$, and
(O3): If $a, b \in \mathbb{P}$ then $a + b \in \mathbb{P}$ and $ab \in \mathbb{P}$.

We mention in passing that the rational numbers also form an ordered field.

THEOREM 14. 0 is neither in $\mathbb{P}$ nor in $-\mathbb{P}$.

EXERCISE 4. Show that the field $\mathbb{Z}_2$ in Exercise 1 can not be ordered.

DEFINITION 2. The elements in $\mathbb{P}$ are called *positive* and those in $-\mathbb{P}$ are called *negative*. The *non-negative* numbers are those in $(-\mathbb{P})^c = \mathbb{P} \cup \{0\}$ while the *non-positive* numbers are those in $\mathbb{P}^c = -\mathbb{P} \cup \{0\}$.

DEFINITION 3. Let a and b be real numbers. We say that $a < b$ or, equivalently, $b > a$ if $b - a \in \mathbb{P}$. We say that $a \leq b$ or, equivalently, $b \geq a$ if $b - a \in \mathbb{P} \cup \{0\}$.

THEOREM 15. $1 > 0$.

THEOREM 16. Suppose $a, b \in \mathbb{R}$. Then $a \leq b$ if and only if $a < b$ or $a = b$.

THEOREM 17. Let $x, y \in \mathbb{R}$. Then either $x \leq y$ or $y \leq x$ (or both).

THEOREM 18. Let $x, y \in \mathbb{R}$. If $x \leq y$ and $y \leq x$ then $x = y$.

THEOREM 19. For any two real numbers x and y exactly one of the following three statements is true: $x < y$, $x = y$, or $x > y$.

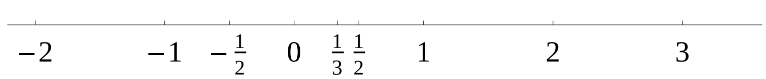
THEOREM 20. Let $x, y, z \in \mathbb{R}$. If $x < y$ and $y < z$, then $x < z$.

THEOREM 21. Let $x, y, z \in \mathbb{R}$. Then $x + y < x + z$ if and only if $y < z$.

THEOREM 22. Let $x, y, z \in \mathbb{R}$.
(1) If $0 < x$ and $y < z$, then $xy < xz$.
(2) If $x < 0$ and $y < z$, then $xz < xy$.

THEOREM 23. If $x, y \in \mathbb{R}$ and $0 < x < y$, then $0 < y^{-1} < x^{-1}$.

These consequences of the order axioms show that one can represent the real numbers on the familiar number line. The relationship $x < y$ is to be interpreted as "x lies to the left of y". It is often helpful to think along these lines when trying to devise proofs. Here $2 = 1 + 1$ and $3 = 2 + 1$.



EXERCISE 5. Identify the theorems which guarantee the ordering indicated in this sketch of the number line.

THEOREM 24. If $x, y \in \mathbb{R}$ and $x < y$, then there exists a number $z \in \mathbb{R}$ such that $x < z < y$.

This shows that any ordered field will have many elements.

DEFINITION 4. Let a, b be real numbers. Each of the following types of subsets of $\mathbb{R}$ is called a *finite interval*: $(a, b) = \{x \in \mathbb{R} : a < x < b\}$, $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$, $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$, and $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$. The following sets are called *infinite intervals*: $(-\infty, a) = \{x \in \mathbb{R} : x < a\}$, $(-\infty, a] = \{x \in \mathbb{R} : x \leq a\}$, $(a, \infty) = \{x \in \mathbb{R} : a < x\}$, $[a, \infty) = \{x \in \mathbb{R} : a \leq x\}$, and $(-\infty, \infty) = \mathbb{R}$. The intervals (a, b), $(-\infty, a)$, and (a, ∞) are called *open intervals* while $[a, b]$, $(-\infty, a]$, and $[a, \infty)$ are called *closed intervals*. The interval $(-\infty, \infty)$ is considered both an open and a closed interval.

Note that the empty set is also considered to be an interval, in fact both an open and a closed interval since $\emptyset = (a, b) = [a, b]$ if $a > b$.

1.3. The induction principle

Definition 5. A subset S of $\mathbb{R}$ is called *inductive*, if $1 \in S$ and if $x + 1 \in S$ whenever $x \in S$.

Theorem 25. The sets $\mathbb{R}$, $\mathbb{P}$, and $[1, \infty)$ are all inductive.

Definition 6. The intersection of all inductive sets is called the set of *natural numbers* and is denoted by $\mathbb{N}$. The set $\mathbb{Z} = -\mathbb{N} \cup \{0\} \cup \mathbb{N}$ is called the set of *integers* or *whole numbers*. The set $\mathbb{Q} = \{m/n : m \in \mathbb{Z}, n \in \mathbb{N}\}$ is called the set of *rational numbers*.

Exercise 6. If the sets A and B are inductive, then so is $A \cap B$.

Theorem 26. $\mathbb{N}$ is inductive. In particular, 1 is a natural number.

Theorem 27. If n is a natural number, then $n \geq 1$.

Theorem 28 (The induction principle). If $M \subset \mathbb{N}$ is inductive, then $M = \mathbb{N}$.

The induction principle gives rise to an important method of proof, the so called induction proofs of which we will see many: $M = \mathbb{N}$, if M has the three properties (i) $M \subset \mathbb{N}$, (ii) $1 \in M$, and (iii) $\forall n \in M : n + 1 \in M$.

Definition 7. If $n \in \mathbb{N}$ we call $n + 1$ the *successor* of n and n the *predecessor* of $n + 1$.

Theorem 29. 1 is the only natural number without a predecessor.

Theorem 30. $(n, n + 1) \cap \mathbb{N} = \emptyset$ whenever $n \in \mathbb{N}$.

Induction may also be used to define concepts. We give some important examples (note that these depend on the recursion theorem, see Appendix A).

Definition 8. Let n be a natural number and g a function from $\mathbb{N}$ to $\mathbb{R}$. The sum of the first n terms of g, denoted by $\sum_{k=1}^{n} g(k)$, is defined inductively by

$$\sum_{k=1}^{1} g(k) = g(1) \text{ and } \sum_{k=1}^{n+1} g(k) = g(n + 1) + \sum_{k=1}^{n} g(k).$$

Also, the product of the first n terms of g, denoted by $\prod_{k=1}^{n} g(k)$, is defined inductively by

$$\prod_{k=1}^{1} g(k) = g(1) \text{ and } \prod_{k=1}^{n+1} g(k) = g(n + 1) \prod_{k=1}^{n} g(k).$$

Definition 9 (Factorial). For natural numbers n one defines the number $n!$ (pronounced '*n factorial*') by

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

In particular, $1! = 1$, $2! = 2$, and $3! = 2 \cdot 3 = 6$. One also defines $0! = 1$ (get used to it).

DEFINITION 10 (Powers). If $x \in \mathbb{R}$ and $n \in \mathbb{N}$, then the number $\prod_{k=1}^{n} x$ is called the n-th *power* of x. It is denoted by x^n. We also define $x^0 = 1$ and, for $x \neq 0$, $x^{-n} = (x^{-1})^n$. In particular $0^0 = 1$ (get used to that, too).

THEOREM 31. For any $n \in \mathbb{N}$

$$\sum_{k=1}^{n} ca_k = c\sum_{k=1}^{n} a_k$$

when c and the a_k are real numbers.

THEOREM 32. For any $n \in \mathbb{N}$ and real numbers a_k and b_k

$$\sum_{k=1}^{n}(a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k.$$

THEOREM 33. If the a_k are real numbers and $n \in \mathbb{N}$, then

$$a_n - a_0 = \sum_{k=1}^{n}(a_k - a_{k-1}).$$

A sum like the one on the right of the previous equation is called a *telescoping sum.*

THEOREM 34. If m and n are natural numbers, then so are $n + m$ and nm.

THEOREM 35. The sum of the first n natural numbers is $n(n+1)/2$, i.e.,

$$\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$$

for all $n \in \mathbb{N}$.

THEOREM 36. Let a be real number different from 1 and n a natural number. Then

$$a^0 + \sum_{k=1}^{n} a^k = \frac{1 - a^{n+1}}{1 - a}.$$

THEOREM 37 (Bernoulli's[1] inequality). If $a > -1$ and $n \in \mathbb{N}$, then $(1 + a)^n \geq 1 + na$.

THEOREM 38. For all $n \in \mathbb{N}$ it is true that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

THEOREM 39. If $a \neq 0$ and $m, n \in \mathbb{Z}$, then $a^m a^n = a^{m+n}$ and $(a^m)^n = a^{mn}$. Moreover, if $a, n \in \mathbb{N}$, then $a^n \in \mathbb{N}$.

Next, we present the well ordering principle. It is actually equivalent to the induction principle and sometimes comes handy in proofs.

THEOREM 40 (The well-ordering principle).* Let $S \subset \mathbb{N}$ and $S \neq \emptyset$. Then S contains a smallest element (denoted by $\min S$).

THEOREM 41. Suppose $0 < q \in \mathbb{Q}$. Then there exist $m, n \in \mathbb{N}$ such that $m/n = q$ and, if $k/\ell = q$ for $k, \ell \in \mathbb{N}$, then $m \leq k$ and $n \leq \ell$.

[1] Jacob Bernoulli (1654 – 1705)

In the situation of the previous theorem m/n is called the representation of q in *lowest terms*. A negative rational number q also has a representation in lowest terms, namely $-m/n$, if m/n is the representation in lowest terms of $-q > 0$.

THEOREM 42. Suppose $m, n \in \mathbb{R} \setminus \{0\}$ and $m^2 = 2n^2$. Then $n \neq m$ and $m/n = (2n - m)/(m - n)$.

THEOREM 43. There is no rational number r for which $r^2 = 2$.

Is there are real number whose square is 2? It will turn out that introducing the real numbers allows to handle this deficiency. In fact, that's why real numbers were invented.

1.4. Counting and infinity

This section deals with counting and the notion of infinity. For a given $n \in \mathbb{N}$ let $Z_n = \{k \in \mathbb{N} : k \leq n\}$.

DEFINITION 11. A set X is called finite if it is empty or if there exists an $n \in \mathbb{N}$ and a surjective function $\varphi : Z_n \to X$. Otherwise the set is called infinite.

If X is a non-empty finite set, then $\min\{k \in \mathbb{N} : \exists \text{ surjective function } \varphi : Z_k \to X\}$ is called the number of elements[2] in X and is denoted by $\#X$. We also define $\#\emptyset = 0$.

EXERCISE 7. Find out what Z_1, Z_2, and Z_3 actually are. Then find all functions $\varphi : Z_1 \to Z_1$ and all functions $\psi : Z_2 \to Z_2$. Which of these are injective and which are surjective? Also, find all injective functions $\tau : Z_3 \to Z_3$.

EXERCISE 8. Think about how counting is related to surjective, injective, and bijective functions.

THEOREM 44. $\#Z_n \leq n$.

THEOREM 45. If $S \subset Z_n$, then $\#S \leq n$.

THEOREM 46. If $\varphi : Z_k \to Z_m$ is surjective, then there is a surjective $\psi : Z_k \to Z_m$ such that $\psi(k) = m$.

THEOREM 47. $\#Z_n = n$.

THEOREM 48. Suppose X is a non-empty finite set. If there is a bijective function from Z_n to X, then $\#X = n$. Conversely, if $\#X = n$, then there is a bijective $\varphi : Z_n \to X$.

THEOREM 49. $\mathbb{N}$ is infinite.

THEOREM 50. Suppose X and Y are sets and $X \subset Y$. Then the following two statements are true:

(1) If Y is finite, then X is finite.
(2) If X is infinite, then Y is infinite.

DEFINITION 12. A set X is called *countable* if it is empty or if there exists a surjective function $\phi : \mathbb{N} \to X$. Otherwise it is called *uncountable*. A set which is countable but not finite is called *countably infinite*.

THEOREM 51. Every finite set is countable and $\mathbb{N}$ is countably infinite.

THEOREM 52. Suppose $k, \ell, m, n \in \mathbb{N}$. Define the function $\varphi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $\varphi(m, n) = m + a(m + n - 2)$ where $a(s) = \sum_{j=1}^{s} j$ for $s \in \mathbb{N}$ and $a(0) = 0$. If $k + \ell > m + n$, then $\varphi(k, \ell) > \varphi(m, n)$.

[2] This number exists by Theorem 40.

THEOREM 53. Let a be defined as in Theorem 52. If $k \in \mathbb{N}$, then $\{m \in \mathbb{N} : k \leq a(m)\}$ is not empty.

THEOREM 54. $\mathbb{Z}$, $\mathbb{N} \times \mathbb{N}$, and $\mathbb{Q}$ are countably infinite.

1.5. The least upper bound axiom

Note that the set of rational numbers, just like the set of real numbers, is an ordered field (identify the set $\mathbb{P}$ which shows this). So far, for all we know, these sets could be identical. Only the least upper bound axiom, which is our final axiom, distinguishes between the two.

DEFINITION 13. Let S be a subset of $\mathbb{R}$. The number b is called an *upper bound* of S if $x \leq b$ for all $x \in S$. It is called a *lower bound* of S if $x \geq b$ for all $x \in S$. If S has an upper (or lower) bound it is called *bounded above* (or *bounded below*). S is called *bounded* if it has both an upper and a lower bound.

EXERCISE 9. Every element of $\mathbb{P}$ is an upper bound of $-\mathbb{P}$. What is the least upper bound of $-\mathbb{P}$?

DEFINITION 14. Let S be a subset of $\mathbb{R}$, which is bounded above, and U the set of all its upper bounds. A number c is called a *least upper bound* of S or a *supremum* of S if it is an upper bound of S and a lower bound of U. Similarly, if S is bounded below and L is the set of all lower bounds of S, then a number c which is a lower bound of S and an upper bound of L is called a *greatest lower bound* of S or an *infimum* of S.

THEOREM 55. Suppose $a, b \in \mathbb{R}$ and $a < b$. Then b is a least upper bound and a is a greatest lower bound of (a, b).

THEOREM 56. If a least upper or a greatest lower bound of S exists, then it is unique.

We are now ready to state our final axiom of the real numbers. This means that all desired results on the real numbers have to be proven by using nothing more than Axioms 1 — 3 (and results which are known to follow from these axioms).

AXIOM 3 (Least Upper Bound Axiom). Any non-empty subset of $\mathbb{R}$ which is bounded above has a least upper bound.

Both $\mathbb{Q}$ and $\mathbb{R}$ satisfy Axioms 1 and 2 (they are both ordered fields). It is Axiom 3 which makes all the difference (a tremendous difference as we will see) between the rational and the real numbers.

THEOREM 57. Let S be a subset of $\mathbb{R}$. The set S is bounded above if and only if $-S$ is bounded below and the number b is an upper bound of S if and only if $-b$ is a lower bound of $-S$.

THEOREM 58. If S is a non-empty subset of $\mathbb{R}$ which is bounded below, then S has a greatest lower bound.

DEFINITION 15. If S is a subset of $\mathbb{R}$ which has a least upper bound we will denote this uniquely defined number by $\sup S$. If $\sup S$ is an element of S, it is called a *maximum* of S and is denoted by $\max S$. Similarly, the greatest lower bound of a set S (when it exists) is denoted by $\inf S$. If $\inf S$ is an element of S, it is called a *minimum* of S and is denoted by $\min S$.

THEOREM 59. If either of the numbers $\sup S$ and $\inf(-S)$ exist, then so does the other. Moreover, in this case, $\sup S = -\inf(-S)$.

Theorem 60. Let S be a bounded non-empty subset of $\mathbb{R}$. Then $\inf S \leq \sup S$.

Theorem 61. Suppose that $\emptyset \neq T \subset S \subset \mathbb{R}$.

(1) If S is bounded above, then T is bounded above and $\sup T \leq \sup S$.
(2) If S is bounded below, then T is bounded below and $\inf T \geq \inf S$.

Theorem 62 (Archimedian[3] property). Let $a, b \in \mathbb{R}$ and $a > 0$. Then there exists a natural number n such that $na > b$.

The following innocent looking theorem will be very important later on when we will look at limits.

Theorem 63. Let x be a real number such that $0 \leq x < 1/n$ for every natural number n. Then $x = 0$.

Theorem 64. If x and y are real numbers and $x < y$, then there exists $q \in \mathbb{Q}$ such that $x < q < y$.

Because of this last theorem one says that the rational numbers are *dense* in $\mathbb{R}$ but recall from Theorem 43 that despite of this there are still holes on the number line. This is why the real numbers are sorely needed.

Our next goal is to prove the existence of square roots of positive real numbers. We will then use this to show that the rational numbers do not satisfy the least upper bound axiom.

Theorem 65. Suppose $x, y \geq 0$ are real numbers. Then $x < y$ if and only if $x^2 < y^2$.

Theorem 66. Let $y > 0$ be a real number and $E = \{z \in \mathbb{R} : z > 0, z^2 < y\}$. Then E is not empty and bounded above.

Theorem 67.* Let $y > 0$ be a real number. Then there exists a real number $x > 0$ such that $x^2 = y$.

Definition 16. Let x and y be real numbers. x is called a square root of y if $x^2 = y$.

Theorem 68. A real number $y < 0$ has no square root. The number $y = 0$ has exactly one square root, namely 0. A real number $y > 0$ has exactly two square roots; these are negatives of each other.

Definition 17. If $x > 0$, then we denote by $\sqrt{x}$ the unique positive square root of x. Also $\sqrt{0} = 0$.

We will usually call this number the square root of $y \geq 0$, even if we actually should be more precise and call it the *non-negative square root*.

Theorem 69. If $a \geq 0$ and $b \geq 0$, then $\sqrt{ab} = \sqrt{a}\sqrt{b}$.

Theorem 70. The rational numbers $\mathbb{Q}$ do not satisfy the Least Upper Bound Axiom. More precisely, if $A \subset \mathbb{Q}$ is bounded above and U is the set of all rational upper bounds of A, then U may not have a least element.

As the final subject of this chapter we will discuss absolute values.

Definition 18. The *absolute value* of $x \in \mathbb{R}$ is defined by

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

[3]Archimedes of Syracuse (ca. 287 BC – 212 BC)

THEOREM 71. Let x, y be any real numbers. Then $\sqrt{x^2} = |x|$ and $|xy| = |x|\,|y|$.

THEOREM 72. Suppose a and x are real numbers and c is a positive real number. Then $-|x| \leq \pm x \leq |x|$. Moreover, the interval $(a-c, a+c)$ equals the set $\{x \in \mathbb{R} : |x-a| < c\}$.

THEOREM 73 (*Triangle inequality*). If $x, y \in \mathbb{R}$, then

$$|x+y| \leq |x| + |y|$$

and

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

DEFINITION 19. Let X be a set. A function $d : X \times X \to [0, \infty)$ is called a *metric* or *distance function* if it has the following properties:

(1) $d(x, y) = 0$ if and only if $x = y$.
(2) $d(x, y) = d(y, x)$ for all $x, y \in X$.
(3) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

(X, d) is then called a metric space. The number $d(x, y)$ is called the *distance* between x and y.

Property (3) is also called *triangle inequality*.

THEOREM 74. The function $\mathbb{R} \times \mathbb{R} \to [0, \infty) : (x, y) \mapsto |x-y|$ is a metric.

Note that, by Theorem 72, the interval $(a-c, a+c)$ is the set of all real numbers whose distance from a is less than c.

CHAPTER 2

Sequences and series

2.1. Sequences

Definition 20. A *sequence* of real numbers is a function which maps $\mathbb{N}$ to $\mathbb{R}$.

For a sequence $x : \mathbb{N} \to \mathbb{R}$ we will usually write $n \mapsto x_n$, where x_n is traditionally used for the value $x(n)$ of the function x at n.

Exercise 10. For the sequence $x : n \mapsto 1/n$ and $y : n \mapsto 1/n^2$ consider where the points x_n and y_n fall on the number line.

We have the impression that these sequences approach (whatever that might mean) zero. How could we make such a notion precise?

Exercise 11. Show that $|1/n - 0| < 1/100$ for all $n > 100$ and $|1/n - 0| < 1/1000$ for all $n > 1000$.

Exercise 12. Intuitively, what number L do the entries of the sequence $n \mapsto x_n = (n^2 + (-1)^n)/n^2$ approach?

Exercise 13. Consider the sequence from the previous exercise and suppose $\varepsilon = 1/100$. Find an $N \in \mathbb{R}$ such that $|x_n - L| < \varepsilon$ for all $n > N$. What about $\varepsilon = 1/500$? What will happen for even smaller ε? Note that ε describes the error we allow.

Definition 21. Let $x : n \mapsto x_n$ be a sequence. We say that x *converges* to the real number L if for every positive real number ε there is an $N \in \mathbb{R}$ such that $|x_n - L| < \varepsilon$ whenever $n > N$; or, more concisely,

$$\forall \varepsilon > 0 : \exists N \in \mathbb{R} : \forall n > N : |x_n - L| < \varepsilon.$$

We say that the sequence x *converges* or is *convergent* if there is a real number L such that x converges to L. If a sequence is not convergent, we say that it *diverges* or is *divergent*.

Theorem 75. If $c \in \mathbb{R}$, then $x : n \mapsto x_n = c$ converges to c.

Theorem 76. The sequence $n \mapsto 1/n$ converges to 0.

Theorem 77. If the sequence x converges to L_1 and also to L_2, then $L_1 = L_2$.

Definition 22. If a sequence x converges to L, then L is called the *limit* of x and we write

$$\lim_{n\to\infty} x_n = L.$$

Definition 23. We say that a sequence is bounded above (or below) if its range is bounded above (or below). A sequence is called *bounded* if it is bounded both above and below.

Theorem 78. A finite set of real numbers is bounded.

THEOREM 79. If a sequence converges, then it is bounded.

THEOREM 80. The sequence $n \mapsto n$ diverges.

THEOREM 81. The sequence $n \mapsto (-1)^n$ diverges.

THEOREM 82. If $\lim_{n\to\infty} x_n = L$, then $\lim_{n\to\infty} |x_n| = |L|$.

THEOREM 83. If $\lim_{n\to\infty} x_n = L$ and $\lim_{n\to\infty} y_n = M$, then $\lim_{n\to\infty}(x_n+y_n) = L+M$.

THEOREM 84. If $c \in \mathbb{R}$ and $\lim_{n\to\infty} x_n = L$, then $\lim_{n\to\infty} cx_n = cL$.

THEOREM 85. If $\lim_{n\to\infty} x_n = L$ and $\lim_{n\to\infty} y_n = M$, then $\lim_{n\to\infty} x_n y_n = LM$.

THEOREM 86. If $\lim_{n\to\infty} x_n = 0$ and the sequence y is bounded, then $\lim_{n\to\infty} x_n y_n = 0$.

THEOREM 87. If $x_n \neq 0$ for all $n \in \mathbb{N}$ and if $\lim_{n\to\infty} x_n = L \neq 0$, then $\lim_{n\to\infty} 1/x_n = 1/L$.

THEOREM 88. For every $k \in \mathbb{N}$ it holds that $\lim_{n\to\infty} n^{-k} = 0$.

THEOREM 89. $\lim_{n\to\infty}(n^2 - n)/(3n^2 + 1) = 1/3$.

THEOREM 90. If the sequences x and y have limits and if $x_n \leq y_n$ for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$.

THEOREM 91. If the sequences x and y both have limit L and if $x_n \leq z_n \leq y_n$ for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} z_n = L$.

THEOREM 92. If $a > 0$ is a real number, then $\lim_{n\to\infty}(1 + a)^{-n} = 0$. Moreover, if $c \in (-1, 1)$, then $\lim_{n\to\infty} c^n = 0$.

DEFINITION 24. A sequence x of real numbers is called *non-decreasing* (or *non-increasing*) if $x_n \leq x_{n+1}$ (or $x_n \geq x_{n+1}$) for all $n \in \mathbb{N}$. It is called *strictly increasing* (or *strictly decreasing*) if the inequalities are strict.

Warning: Different people use somewhat different notation here: increasing may be used both for non-decreasing or for strictly increasing. One is on the safe side if one uses non-decreasing and strictly increasing as we do here.

THEOREM 93. If the sequence x is non-decreasing and bounded above, then it converges to $\sup\{x_n : n \in \mathbb{N}\}$. Similarly, if x is non-increasing and bounded below, then it converges to $\inf\{x_n : n \in \mathbb{N}\}$.

THEOREM 94. Suppose the sequence x is bounded and define sequences $\hat{x}$ and $\check{x}$ by setting $\hat{x}_n = \sup\{x_k : k \geq n\}$ and $\check{x}_n = \inf\{x_k : k \geq n\}$. Then both $\hat{x}$ and $\check{x}$ are convergent.

DEFINITION 25. Suppose the sequence x is bounded. Then the limit of the sequence $n \mapsto \sup\{x_k : k \geq n\}$ is called the *limit superior* of x and is denoted by $\limsup_{n\to\infty} x_n$, i.e.,

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : \ k \geq n\}.$$

Similarly, the limit of the sequence $n \mapsto \inf\{x_k : k \geq n\}$ is called the *limit inferior* of x and is denoted by $\liminf_{n\to\infty} x_n$, i.e.,

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : \ k \geq n\}.$$

THEOREM 95. Let x be the sequence defined by $x_n = (-1)^n + 1/n$. Then we have $\limsup_{n\to\infty} x_n = 1$ and $\liminf_{n\to\infty} x_n = -1$.

THEOREM 96. Suppose x is a bounded sequence. Then $\liminf_{n\to\infty} x_n \le \limsup_{n\to\infty} x_n$.

THEOREM 97. Let x be a bounded sequence. Then $\limsup_{n\to\infty} -x_n = -\liminf_{n\to\infty} x_n$.

THEOREM 98. Let x and y be bounded sequences such that $x_n \le y_n$ for all $n \in \mathbb{N}$. Then $\limsup_{n\to\infty} x_n \le \limsup_{n\to\infty} y_n$ and $\liminf_{n\to\infty} x_n \le \liminf_{n\to\infty} y_n$.

THEOREM 99. Let x and y be bounded sequences. Then

$$\limsup_{n\to\infty}(x_n + y_n) \le \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n$$

and

$$\liminf_{n\to\infty}(x_n + y_n) \ge \liminf_{n\to\infty} x_n + \liminf_{n\to\infty} y_n.$$

THEOREM 100. The inequalities in Theorem 99 may be strict.

THEOREM 101. Suppose the sequence x is bounded and $\varepsilon > 0$. Then there is an $N \in \mathbb{R}$ such that $-\varepsilon + \liminf_{n\to\infty} x_n \le x_k \le \varepsilon + \limsup_{n\to\infty} x_n$ whenever $k > N$.

DEFINITION 26. If x is a sequence of real numbers and k is a strictly increasing sequence of natural numbers, then the sequence $n \mapsto x_{k_n}$ is called a *subsequence* of x.

For instance, the sequences $n \mapsto 1/(2n)$ and $n \mapsto 1/(n+1)$ are subsequences of $n \mapsto 1/n$.

THEOREM 102. If a sequence of real numbers converges to $L \in \mathbb{R}$ then so does every one of its subsequences.

THEOREM 103.* If the sequence x is bounded, then there exists a subsequence of x which converges to $\limsup_{n\to\infty} x_n$ and one which converges to $\liminf_{n\to\infty} x_n$. In particular, every bounded sequence has convergent subsequences.

The last statement in the previous theorem is called the *Bolzano*[1]*-Weierstrass*[2] *theorem.*

THEOREM 104. Let x be a bounded sequence. Then $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$ if and only if x converges. In this case it holds that

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} x_n = \limsup_{n\to\infty} x_n.$$

DEFINITION 27. A sequence x is called a *Cauchy*[3] *sequence* or *Cauchy* for short if

$$\forall\varepsilon > 0 : \exists N \in \mathbb{R} : \forall n, m > N : |x_n - x_m| < \varepsilon.$$

THEOREM 105. A convergent sequence is Cauchy.

THEOREM 106. Any Cauchy sequence is bounded.

THEOREM 107. If x is a Cauchy sequence and x' a convergent subsequence of x, then x is convergent and x and x' have the same limit.

THEOREM 108. A sequence of real numbers converges if and only if it is a Cauchy sequence.

The fact that every Cauchy sequences converges does not hold in $\mathbb{Q}$. It is a consequence of the least upper bound axiom. In fact, one can show that the least upper bound axiom and the requirement that every Cauchy sequence converges are equivalent. To refer to this essential property easily one says that $\mathbb{R}$ is *complete.*

[1] Bernhard Bolzano (1781 – 1848)
[2] Karl Weierstrass (1815 – 1897)
[3] Augustin-Louis Cauchy (1789 – 1857)

2.2. Sums and the Σ-notation

In Definition 8 we introduced a notation for the sum of finitely many terms using the Σ-symbol. It will convenient to extend this notation to a more general setting.

Definition 28. If m and n are any two integers and f a function from $\mathbb{Z}$ to $\mathbb{R}$, we define

$$\sum_{k=m}^{n} f(k) = \begin{cases} 0 & \text{if } m > n, \\ \sum_{k=1}^{n+1-m} f(m-1+k) & \text{if } m \le n. \end{cases}$$

Note that, if $m \le n$, only the values $f(j)$ with $m \le j \le n$ are needed in the above definition.

The following theorems are useful when dealing with finite sums. In each of them f is a real-valued function with an appropriate domain.

Theorem 109. Suppose $m, n, p \in \mathbb{Z}$. Then $\sum_{k=m}^{n} f(k) = \sum_{k=m+p}^{n+p} f(k-p)$.

Theorem 110. Suppose $m, n \in \mathbb{Z}$. Then $\sum_{k=m}^{n} f(k) = \sum_{k=m}^{n} f(m+n-k)$.

Theorem 111. Suppose $k, m, n \in \mathbb{Z}$ and $k \le m \le n$. Then

$$\sum_{j=k}^{n} f(j) = \sum_{j=k}^{m} f(j) + \sum_{j=m+1}^{n} f(j).$$

Theorem 112. If $m, n \in \mathbb{N} \cup \{0\}$, then $\sum_{k=0}^{m} \sum_{\ell=0}^{n} f(\ell, k) = \sum_{\ell=0}^{n} \sum_{k=0}^{m} f(\ell, k)$.

Theorem 113. Any permutation is a finite composition of transpositions.

Theorems 110 and 112 are special cases of the following one.

Theorem 114 (The generalized commutative law). Suppose $n \in \mathbb{N}$ and $\pi : Z_n \to Z_n$ is bijective. Then $\sum_{k=1}^{n} f(k) = \sum_{k=1}^{n} f(\pi(k))$. Moreover, if A is a finite set with n elements and ϕ and ψ are bijective functions from Z_n to A, then $\sum_{k=1}^{n} f(\phi(k)) = \sum_{k=1}^{n} f(\psi(k))$.

This theorem allows for the following definition.

Definition 29. If A is a finite set and f a function from A to $\mathbb{R}$, then we define $\sum_{a \in A} f(a) = \sum_{k=1}^{n} f(\phi(k))$ where $n = \#A$ and $\phi : Z_n \to A$ is bijective.

2.3. Series

Definition 30. If $x : \mathbb{N} \to \mathbb{R}$ is a sequence of real numbers the sequence

$$s : \mathbb{N} \to \mathbb{R} : n \mapsto s_n = \sum_{k=1}^{n} x_k$$

is a called the *sequence of partial sums* of x or a *series*. We will denote s by $\sum_{k=1}^{\infty} x_k$.

Thus a series is actually a sequence. Conversely, in view of Theorem 33, a sequence may also be cast as a series: given a sequence s defining $x_1 = s_1$ and $x_n = s_n - s_{n-1}$ shows that s is a series, viz., $s = \sum_{k=1}^{\infty} x_k$. We emphasize that (in this context) the symbol $\sum_{k=1}^{\infty} x_k$ is not to be thought of as a real number. It is only a concise notation for a sequence of partial sums. Indeed, if this sequence of partial sums does not converge, it is not sensible to assign any real number to the symbol $\sum_{k=1}^{\infty} x_k$.

Definition 31. Suppose x is a sequence of real numbers. If the sequence s of partial sums of x converges to $L \in \mathbb{R}$, we say the series $\sum_{k=1}^{\infty} x_k$ *converges*. We then write (abusing notation slightly)

$$\sum_{k=1}^{\infty} x_k = \lim_{n\to\infty} s_n = L.$$

Of course, a series (i.e., a sequence of partial sums) may also *diverge*.

We will also use the following notation: If $N \in \mathbb{Z}$ we define $\sum_{k=N}^{\infty} x_k$ to be the sequence of partial sums $n \mapsto \sum_{k=N}^{n+N-1} x_k$. If this is convergent we denote its limit also by $\sum_{k=N}^{\infty} x_k$.

Theorem 115. If $-1 < a < 1$, then the *geometric series* $\sum_{n=0}^{\infty} a^n$ converges to $1/(1-a)$.

Theorem 116. $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$.

Theorem 117. If $c \in \mathbb{R}$ and the series $\sum_{n=1}^{\infty} x_n$ converges, then the series $\sum_{n=1}^{\infty} cx_n$ converges and

$$\sum_{n=1}^{\infty} cx_n = c\sum_{n=1}^{\infty} x_n.$$

Theorem 118. If $\sum_{n=1}^{\infty} x_n$ converges and $\sum_{n=1}^{\infty} y_n$ converges, then $\sum_{n=1}^{\infty}(x_n + y_n)$ converges and

$$\sum_{n=1}^{\infty}(x_n + y_n) = \sum_{n=1}^{\infty} x_n + \sum_{n=1}^{\infty} y_n.$$

Theorem 119. If the series $\sum_{n=1}^{\infty} x_n$ converges, then $\lim_{n\to\infty} x_n = 0$.

Theorem 120. The series $\sum_{n=1}^{\infty}(-1)^n$ diverges.

Theorem 121 (Cauchy Criterion). The series $\sum_{n=1}^{\infty} x_n$ converges if and only if the following criterion is satisfied: For every $\varepsilon > 0$ there is an $N \in \mathbb{R}$ such that, for all $m \in \mathbb{N}$ and $k \in \mathbb{N}$ with $m > k > N$,

$$\left|\sum_{n=k+1}^{m} x_n\right| < \varepsilon.$$

Theorem 122. The *harmonic series* $\sum_{n=1}^{\infty} 1/n$ diverges.

This example shows that there is no converse to Theorem 119, i.e., the fact that the sequence converges to 0 does not guarantee the convergence of the associated sequence of partial sums.

Theorem 123. The *alternating harmonic series* $\sum_{n=1}^{\infty}(-1)^{n+1}/n$ is convergent.

The following inequality is a generalized version of the triangle inequality.

Theorem 124. Let x_n be a real number for every $n \in \mathbb{N}$. Then

$$\left|\sum_{n=1}^{k} x_n\right| \leq \sum_{n=1}^{k} |x_n|$$

for any $k \in \mathbb{N}$.

Theorem 125 (Comparison Test). Let x, y, and z be sequences and assume there is an $N \in \mathbb{N}$ such that $|x_n| \leq y_n \leq z_n$ for all $n \geq N$. If the series $\sum_{n=1}^{\infty} y_n$ converges, then the series $\sum_{n=1}^{\infty} x_n$ converges, too, and

$$\left| \sum_{n=N}^{\infty} x_n \right| \leq \sum_{n=N}^{\infty} y_n.$$

If the series $\sum_{n=1}^{\infty} y_n$ diverges, then the series $\sum_{n=1}^{\infty} z_n$ diverges, too.

Theorem 126. If $2 \leq k \in \mathbb{N}$, then $\sum_{n=1}^{\infty} n^{-k}$ converges.

Theorem 127 (Raabe's[4] Test). Suppose that there is an $N \in \mathbb{N}$ and a $\beta > 1$ such that

$$\left| \frac{a_{n+1}}{a_n} \right| \leq 1 - \frac{\beta}{n+1}$$

whenever $n \geq N$. Then the series $\sum_{n=1}^{\infty} a_n$ converges.

Theorem 128 (Ratio Test). Suppose that $x_n \neq 0$ for all $n \in \mathbb{N}$ and that

$$\limsup_{n \to \infty} \left| \frac{x_{n+1}}{x_n} \right| < 1.$$

Then the series $\sum_{n=1}^{\infty} x_n$ converges.

Theorem 129. Let a be a fixed real number. Then the series $\sum_{n=0}^{\infty} a^n/n!$ converges.

Definition 32. We set

$$\exp(a) = \sum_{n=0}^{\infty} a^n/n!$$

and

$$\exp(1) = \sum_{n=0}^{\infty} 1/n! = \mathrm{e}.$$

Definition 33. For integers n and k with $n \geq k \geq 0$ let

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

These numbers are called *binomial coefficients*.

Theorem 130. For integers n and k with $1 \leq k \leq n$ it holds that

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

Theorem 131 (Binomial Theorem). Let $a, b \in \mathbb{R}$ and n a non-negative integer. Then

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

Theorem 132. $\limsup_{n\to\infty}(1+1/n)^n \leq \mathrm{e}$.

Theorem 133. $\liminf_{n\to\infty}(1+1/n)^n \geq \mathrm{e}$.

Combining Theorem 132 and Theorem 133 we get

$$\mathrm{e} = \lim_{n \to \infty} (1+1/n)^n.$$

[4] Joseph Ludwig Raabe (1801 – 1859)

THEOREM 134. Let $x \geq 0$ and $n \in \mathbb{N}$. Then there exists a unique $y \geq 0$ such that $y^n = x$.

DEFINITION 34. The unique number y given by Theorem 134 is called the (non-negative) n-th root of $x \geq 0$. We write $y = \sqrt[n]{x}$.

Inspired by Theorem 39 the notation $y = \sqrt[n]{x} = x^{1/n}$ is frequently used.

THEOREM 135. If $p > 0$, then $\lim_{n\to\infty} \sqrt[n]{p} = 1$.

THEOREM 136. $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

THEOREM 137 (Root Test). Let x be a sequence such that $\limsup_{n\to\infty} \sqrt[n]{|x_n|} < 1$. Then the series $\sum_{n=1}^{\infty} x_n$ converges.

The following inequality shows that the root test is more precise than the ratio test. The advantage of the ratio test is that it is usually more easily applicable.

THEOREM 138. Suppose a is a bounded sequence of positive numbers. Then

$$\limsup_{n\to\infty} \sqrt[n]{a_n} \leq \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$$

if the right hand side exists.

DEFINITION 35. The series $\sum_{n=1}^{\infty} x_n$ is said to *converge absolutely* if the series $\sum_{n=1}^{\infty} |x_n|$ converges. If the series $\sum_{n=1}^{\infty} x_n$ converges, but does not converge absolutely, then we say that it *converges conditionally*.

THEOREM 139. If the series $\sum_{n=1}^{\infty} x_n$ converges absolutely, then it converges and

$$\left|\sum_{n=1}^{\infty} x_n\right| \leq \sum_{n=1}^{\infty} |x_n|.$$

DEFINITION 36. Let $\pi : \mathbb{N} \to \mathbb{N}$ be a bijective sequence. Let x be a given sequence and define a sequence y by $y_n = x_{\pi_n}$. Then we say that the series $\sum_{n=1}^{\infty} y_n$ is a *rearrangement* of the series $\sum_{n=1}^{\infty} x_n$.

THEOREM 140.* Suppose that $\sum_{n=1}^{\infty} x_n$ converges absolutely and let $\sum_{n=1}^{\infty} y_n$ be a rearrangement of $\sum_{n=1}^{\infty} x_n$. Then $\sum_{n=1}^{\infty} y_n$ converges absolutely, and

$$\sum_{n=1}^{\infty} y_n = \sum_{n=1}^{\infty} x_n.$$

Let $\sum_{n=1}^{\infty} x_n$ be a convergent series which does not converges absolutely. Then there is a rearrangement of the series such that the sequence of its partial sums is not bounded above or below. Moreover, given any real number s, there is a rearrangement which converges to s. We will not try to prove this statement.

CHAPTER 3

A zoo of functions

Definition 37. Let X be a set and f, g functions from X to $\mathbb{R}$. We define their sum, difference, and product respectively by $(f \pm g)(x) = f(x) \pm g(x)$ and $(fg)(x) = f(x)g(x)$ for all $x \in X$. If $g(x) \neq 0$ for all $x \in X$ we also define the quotient of f and g by $(f/g)(x) = f(x)/g(x)$.

The composition of functions $(g \circ f)(x) = g(f(x))$ is defined in the appendix.

Definition 38. Let $S \subset \mathbb{R}$. A function $f : S \to \mathbb{R}$ is called *non-decreasing* (or *non-increasing*) if $f(x) \leq f(y)$ (or $f(x) \geq f(y)$) whenever $x, y \in S$ and $x \leq y$. It is called *strictly increasing* (or *strictly decreasing*) if the inequalities are strict. A function is called *monotone* if it is non-increasing or non-decreasing. It is called *strictly monotone* if it is strictly increasing or strictly decreasing.

Definition 39. If n is a non-negative integer and a_0, a_1, …, a_n are real numbers, then the function $p : \mathbb{R} \to \mathbb{R}$ defined by

$$p(x) = \sum_{k=0}^{n} a_k x^k$$

is called a *polynomial function* or a *polynomial* for short. The integer n is called the *degree* of p if $a_n \neq 0$. The zero function is also a polynomial but no degree is assigned to it.

Theorem 141. If p and q are polynomials of degree n and k, respectively, then $p+q$ and pq are also polynomials. The degree of $p+q$ is the larger of the numbers n and k unless $n = k$ in which case the degree is at most n. The degree of pq equals $n+k$.

Theorem 142. A polynomial p of degree n has at most n zeros, i.e., there are at most n points x such that $p(x) = 0$.

Definition 40. Let p and q be polynomials and assume $q \neq 0$, i.e., q is not the zero polynomial. Let $S = \{x \in \mathbb{R} : q(x) \neq 0\}$. Then, the function $r : S \to \mathbb{R}$ given by

$$r(x) = \frac{p(x)}{q(x)}$$

is called a *rational function*.

Every polynomial is a rational function (choose $q = 1$).

Definition 41. Let J be an interval in $\mathbb{R}$. A function $f : J \to \mathbb{R}$ is called an *algebraic function* if there is a natural number n and polynomials p_0, p_1, ..., p_n, not all zero, such that

$$p_0(x) + p_1(x)f(x) + \cdots + p_n(x)f(x)^n = 0$$

for all $x \in J$. A function which is not algebraic is called *transcendental*.

Every rational function is algebraic (choose $n = 1$). Other examples of algebraic functions are $f(x) = \sqrt[n]{x}$ for $n \in \mathbb{N}$ and $J = [0, \infty)$ and $f(x) = \sqrt{1 - x^2}$ for $J = [-1, 1]$.

DEFINITION 42. The function $\exp : \mathbb{R} \to \mathbb{R}$ defined by

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

is called the *exponential function*.

The exponential function is transcendental but we will not prove this fact.

THEOREM 143. Let $E_n(x)$ denote the partial sums of the series defining the exponential function, i.e.,

$$E_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

If $|x|, |y| \leq a$ and if $a \geq 1$, then

$$|E_n(x)E_n(y) - E_n(x + y)| \leq \frac{(2a^2)^n}{n!}.$$

THEOREM 144. Suppose $x, y \in \mathbb{R}$. Then

$$\exp(x + y) = \exp(x)\exp(y).$$

THEOREM 145. $\exp(x) > 1$ for all $x > 0$ and $\exp(x) > 0$ for all $x \in \mathbb{R}$.

THEOREM 146. $\exp(x) = 1$ if and only if $x = 0$ and $\exp(x) = \exp(y)$ if and only if $x = y$. The exponential function is strictly increasing.

THEOREM 147. The range of the exponential function is $(0, \infty)$.

Since $\exp : \mathbb{R} \to (0, \infty)$ is bijective it has an inverse function.

DEFINITION 43. The function $\log : (0, \infty) \to \mathbb{R}$ defined by

$$\log(x) = y \text{ if and only if } x = \exp(y)$$

(i.e., the inverse function of exp) is called the *logarithmic function*.

THEOREM 148. The logarithmic function is bijective. $\log(1) = 0$ and $\log(x)$ is positive for $x > 1$ and negative for $0 < x < 1$. Moreover, $\log(xy) = \log(x) + \log(y)$ if $x, y > 0$. Finally, if $x > 0$, then $\log(1/x) = -\log(x)$.

THEOREM 149. If $0 < x \in \mathbb{R}$, $p, q \in \mathbb{Z}$ and $q \neq 0$, then $x^{p/q} = \exp(p \log(x)/q)$.

DEFINITION 44. Let x be a positive real number and r any real number. Then we define the r-th *power* of x by $x^r = \exp(r \log(x))$.

Since $\exp(1) = \mathrm{e}$ and hence $\log(\mathrm{e}) = 1$ we have, in particular, that $\exp(r) = \mathrm{e}^r$ whenever $r \in \mathbb{R}$.

THEOREM 150. Suppose $x > 0$ and $r, s \in \mathbb{R}$. Then $\log(x^r) = r \log(x)$, $x^r x^s = x^{r+s}$, and $(x^r)^s = x^{rs}$.

THEOREM 151. Suppose $x > 0$. Then the series $\sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$ and $\sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$ are convergent.

Definition 45. The function $\sin : \mathbb{R} \to \mathbb{R}$ defined by

$$\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

is called the *sine function.*

The function $\cos : \mathbb{R} \to \mathbb{R}$ defined by

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

is called the *cosine function.*

The function $\tan = \sin / \cos$ (which is not defined on all of $\mathbb{R}$) is called the *tangent function.*

The functions sin, cos, and tan are called *trigonometric functions.* They are not one-to-one but restrictions to certain intervals are. On these one can define their inverse functions. Before we can study these issues more deeply we need better tools, which will be provided in the next chapters.

CHAPTER 4

Continuity

4.1. Limits of functions

DEFINITION 46. Let $S \subset \mathbb{R}$ and $a \in \mathbb{R}$. We say that a is a *limit point* of S if there exists a sequence x such that $x_n \in S \setminus \{a\}$ for every $n \in \mathbb{N}$ and $\lim_{n\to\infty} x_n = a$. A point $a \in S$ which is not a limit point of S is called an *isolated point* of S.

EXERCISE 14. Find the limit points and the isolated points of $\{1/n : n \in \mathbb{N}\}$ and those of the interval $(0, 1)$.

THEOREM 152. Let S be a subset of $\mathbb{R}$. The point $a \in S$ is an isolated point of S if and only if there exists a positive δ such that $(a - \delta, a + \delta) \cap S = \{a\}$.

DEFINITION 47. Let S be a subset of $\mathbb{R}$, $f : S \to \mathbb{R}$ a function, and $a \in \mathbb{R}$ a limit point of S. We say that f *converges* to the real number b as x tends to a if for every positive ε there is a positive δ such that for all $x \in S$ for which $0 < |x - a| < \delta$ we have that $|f(x) - b| < \varepsilon$ or, more formally, if

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in S : 0 < |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon.$$

It is important to note that the point a does not necessarily have to be an element of S.

THEOREM 153. Suppose $f : S \to \mathbb{R}$ converges to both b_1 and b_2 as x tends to a. Then $b_1 = b_2$.

DEFINITION 48. If $f : S \to \mathbb{R}$ converges to b as x tends to a, we call b the *limit* of f at a and we write $\lim_{x\to a} f(x) = b$.

Thus limits, when they exist, must be unique. We emphasize, however, that the limit of f at a does not have to be equal to $f(a)$ even if $a \in S$. Definition 47 requires that a can be approached by elements of S other than a. We do not define a limit of a function at an isolated point of its domain.

EXERCISE 15. Does the function $x \mapsto (x+2)/x$ have a limit at $a = 4$? Find an interval (c, d) such that $(x+2)/x$ is closer to that limit than $\varepsilon = 0.001$ for all $x \in (c, d)$. What would δ be?

THEOREM 154.* Let S, f, and a be as in Definition 47. Then, f has a limit at a if and only if the sequence $n \mapsto f(u_n)$ is convergent whenever $u : \mathbb{N} \to S \setminus \{a\}$ is a sequence with limit a. In this situation all limits in question are equal to each other.

THEOREM 155. Let $S \subset \mathbb{R}$, $a \in \mathbb{R}$ a limit point of S, and $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ functions which have limits at a. Then $f \pm g$ have limits at a. In fact,

$$\lim_{x\to a}(f \pm g)(x) = \lim_{x\to a} f(x) \pm \lim_{x\to a} g(x).$$

THEOREM 156. Let $S \subset \mathbb{R}$, $a \in \mathbb{R}$ a limit point of S, and $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ functions which have limits at a. Then fg has a limit at a. In fact,

$$\lim_{x \to a} (fg)(x) = \lim_{x \to a} f(x) \lim_{x \to a} g(x).$$

THEOREM 157. Let S, a, f, and g be as in Theorem 156. Suppose that $g(x) \neq 0$ for all $x \in S$ and that $\lim_{x \to a} g(x) \neq 0$. Then f/g has a limit at a. In fact,

$$\lim_{x \to a} (f/g)(x) = \lim_{x \to a} f(x) / \lim_{x \to a} g(x).$$

4.2. Continuous functions

DEFINITION 49. Let $S \subset \mathbb{R}$ and $a \in S$. A function $f : S \to \mathbb{R}$ is called *continuous at a* if and only if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in S$ with $|x - a| < \delta$ it holds that $|f(x) - f(a)| < \varepsilon$. Concisely,

$$\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in S : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

We say that $f : S \to \mathbb{R}$ is *continuous on S* (or simply continuous) if f is continuous at *every* point of S.

THEOREM 158. Suppose $S \subset \mathbb{R}$, $f : S \to \mathbb{R}$, and $a \in S$ is an isolated point of S. Then f is continuous at a.

THEOREM 159. Suppose $S \subset \mathbb{R}$, $f : S \to \mathbb{R}$, and $a \in S$ is not an isolated point of S. Then f is continuous at a if and only if $\lim_{x \to a} f(x)$ exists and equals $f(a)$.

THEOREM 160. Let c a fixed real number.

(1) The constant function $f : \mathbb{R} \to \mathbb{R} : x \mapsto c$ is continuous.
(2) The identity function $f : \mathbb{R} \to \mathbb{R} : x \mapsto x$ is continuous.
(3) The absolute value function $f : \mathbb{R} \to \mathbb{R} : x \mapsto |x|$ is continuous.

THEOREM 161. The Heaviside function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1 & \text{for } x \geq 0, \end{cases}$$

is not continuous at the point $a = 0$ but continuous everywhere else.

THEOREM 162. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x & \text{for } x \in \mathbb{Q}, \\ 0 & \text{for } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is continuous at $a = 0$ but not continuous at any other point in $\mathbb{R}$.

THEOREM 163. Let $S \subset \mathbb{R}$, $a \in S$, and f and g be functions defined on S which are continuous at a. Then the following statements are true.

(1) $f \pm g$ is continuous at a.
(2) fg is continuous at a.
(3) If $g(a) \neq 0$, then f/g is continuous at a.

THEOREM 164. All polynomials are continuous on $\mathbb{R}$.

THEOREM 165. Let p and q be polynomials. Then the rational function $r = p/q$ is continuous on $S = \{x \in \mathbb{R} : q(x) \neq 0\}$, its domain.

THEOREM 166. Let S_1 and S_2 be subsets of $\mathbb{R}$ and consider the functions $f : S_1 \to S_2$ and $g : S_2 \to \mathbb{R}$. If f is continuous at $a \in S_1$ and g is continuous at $f(a)$, then $g \circ f$ is continuous at a.

THEOREM 167. Let $S \subset \mathbb{R}$ and $f : S \to \mathbb{R}$ a continuous function. Then $|f| : S \to \mathbb{R}$ defined by $|f|(x) = |f(x)|$ is a continuous function.

THEOREM 168. The function $x \mapsto \sqrt{x}$ is continuous on $[0, \infty)$ and the function $x \mapsto \sqrt{|x|}$ is continuous on $\mathbb{R}$.

THEOREM 169. The function $x \mapsto \sqrt{1 - x^2}$ is continuous on $[-1, 1]$.

THEOREM 170. Suppose $I \subset \mathbb{R}$ is an interval and $f : I \to \mathbb{R}$ is either strictly increasing or strictly decreasing. If $f(I)$ is also an interval, then f is continuous.

THEOREM 171. If $f : S \to \mathbb{R}$ is strictly increasing, then f is one-to-one and f^{-1} is strictly increasing. The statements where "strictly increasing" is replaced by "strictly decreasing" also holds.

THEOREM 172. The exponential function and the logarithmic function are continuous on their respective domains.

THEOREM 173. Suppose $r \in \mathbb{R}$. Then the power function $x \mapsto x^r$ is continuous on $(0, \infty)$.

4.3. The intermediate value theorem and some of its consequences

The following theorem is very useful in determining the range of a function. We have used the central idea of its proof in showing the existence of roots (Theorems 67 and 134) and in establishing the range of the exponential function (Theorem 147).

THEOREM 174 (Intermediate Value Theorem).* Let $a, b \in \mathbb{R}$ be such that $a < b$. Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and let $A \in \mathbb{R}$ be such that $f(a) \leq A \leq f(b)$ or $f(b) \leq A \leq f(a)$. Then there exists $x \in [a, b]$ such that $f(x) = A$.

THEOREM 175. Let $a, b \in \mathbb{R}$ with $a < b$ and $f : [a, b] \to \mathbb{R}$ a continuous function. Then $f([a, b])$ is bounded.

THEOREM 176. Let $a, b \in \mathbb{R}$ with $a < b$ and $f : [a, b] \to \mathbb{R}$ a continuous function. Then there are points x_1 and x_2 in $[a, b]$ such that $f(x_1) = \sup f([a, b])$ and $f(x_2) = \inf f([a, b])$.

Recall that a supremum is called a maximum and an infimum a minimum, if it is an element of the set in question.

THEOREM 177. Let $a, b \in \mathbb{R}$ with $a < b$ and $f : [a, b] \to \mathbb{R}$ a continuous function. Then $f([a, b])$ is a closed and bounded interval.

4.4. Uniform convergence and continuity

DEFINITION 50. Let $S \subset \mathbb{R}$ and, for each $n \in \mathbb{N}$, let f_n be a function from S to $\mathbb{R}$. The map $n \mapsto f_n$ is called a *sequence* of functions. We say that $n \mapsto f_n$ *converges pointwise* to a function $f : S \to \mathbb{R}$ if for *each* point $x \in S$ the numerical sequence $n \mapsto f_n(x)$ converges to $f(x)$.

THEOREM 178. Define the functions $f_n : [0, 1] \to \mathbb{R}$ by $f_n(x) = x^n$ for $n \in \mathbb{N}$. The sequence $n \mapsto f_n$ converges pointwise to the function which is 0 on $[0, 1)$ but 1 at 1.

Note that each f_n is continuous but that the limit function is not.

DEFINITION 51. Let $S \subset \mathbb{R}$ and $n \mapsto f_n$ a sequence of functions on S. We say that $n \mapsto f_n$ *converges uniformly* to a function $f : S \to \mathbb{R}$ if for every $\varepsilon > 0$ there exists $N \in \mathbb{R}$ such that for all $n > N$ and *all* $x \in S$ we have that $|f_n(x) - f(x)| < \varepsilon$.

To reiterate, consider the statements

(1) $\forall x \in S : \forall \varepsilon > 0 : \exists N \in \mathbb{R} : \forall n > N : |f_n(x) - f(x)| < \varepsilon$.
(2) $\forall \varepsilon > 0 : \exists N \in \mathbb{R} : \forall x \in S : \forall n > N : |f_n(x) - f(x)| < \varepsilon$.

A sequence $n \mapsto f_n$ converges to f pointwise if (1) holds and uniformly if (2) holds.

EXERCISE 16. Let f_n be the sequence defined in Theorem 178. Given $\varepsilon > 0$ find N so that $|f_n(x)| < \varepsilon$ for all $n > N$ where x is, in turn, each of the values 1/10, 9/10, and 99/100.

THEOREM 179. Uniform convergence implies pointwise convergence but not vice versa.

THEOREM 180. Let $S \subset \mathbb{R}$ and $n \mapsto f_n$ a sequence of continuous functions on S which converges uniformly to a function $f : S \to \mathbb{R}$. Then f is continuous on S.

THEOREM 181 (Cauchy criterion for uniform convergence). Let $S \subset \mathbb{R}$ and $f_n : S \to \mathbb{R}$ for $n \in \mathbb{N}$. The sequence of functions $n \mapsto f_n$ converges uniformly on S if and only if for all $\varepsilon > 0$ there exists $N \in \mathbb{R}$ such that for all $n, m > N$ and all $x \in S$ it holds that $|f_n(x) - f_m(x)| < \varepsilon$.

A series of functions is defined as the sequence of partial sums of functions with a common domain. Thus the definitions of pointwise and uniform convergence extend also to series of functions. For instance, the exponential function is a series of polynomials (powers to be more precise).

THEOREM 182 (Weierstrass M-test). Let $S \subset \mathbb{R}$ and suppose $n \mapsto g_n$ is a sequence of real-valued functions defined on S. Assume that there are non-negative numbers M_n such that $|g_n(x)| \le M_n$ for all $x \in S$ and that the series $\sum_{n=1}^{\infty} M_n$ converges. Then the following statements hold:

(1) The series $\sum_{n=1}^{\infty} g_n(x)$ converges absolutely for every $x \in S$.
(2) The series $\sum_{n=1}^{\infty} g_n$ converges uniformly in S.

DEFINITION 52. Let $x_0 \in \mathbb{R}$ and $a_n \in \mathbb{R}$ for all non-negative integers n. Then the series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ is called a *power series*.

THEOREM 183. Suppose that $x, t \in \mathbb{R}$ and $|x| < |t|$. If $\sum_{n=0}^{\infty} a_n t^n$ converges, then $\sum_{n=0}^{\infty} a_n x^n$ converges absolutely. If, however, $\sum_{n=0}^{\infty} a_n x^n$ diverges, then $\sum_{n=0}^{\infty} a_n t^n$ diverges also.

DEFINITION 53. Let the series $S = \sum_{n=0}^{\infty} a_n x^n$ and the set

$$A = \{x \ge 0 : \sum_{n=0}^{\infty} |a_n| x^n \text{ converges}\}$$

be given. If A is bounded the number $\sup A$ is called the *radius of convergence* of S. If A is unbounded, we say that S has infinite radius of convergence.

The following theorem explain the choice of the name "radius of convergence".

THEOREM 184. Let S and A be as in Definition 53. If A is bounded, then S converges absolutely when $|x| < \sup A$ and diverges when $|x| > \sup A$. If A is unbounded, then S converges absolutely for all $x \in \mathbb{R}$.

THEOREM 185. Suppose the power series $\sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $R > 0$ and that $0 < \varepsilon < R$. Then the power series converges uniformly on $S = \{x \in \mathbb{R} : |x| \leq R - \varepsilon\}$. If it has infinite radius of convergence and $r > 0$, then it converges uniformly on $S = \{x \in \mathbb{R} : |x| \leq r\}$.

THEOREM 186. The power series defining the exponential function or the sine or cosine functions have infinite radius of convergence. Each of these functions is continuous on $\mathbb{R}$.

CHAPTER 5

Differentiation

Throughout this chapter S denotes a set without isolated points. The most important case is when S is an interval (finite or infinite; open, closed, or half-closed).

5.1. Derivatives

DEFINITION 54. Let $f : S \to \mathbb{R}$ be a function, and $a \in S$. We say that f is *differentiable* at a if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists. The limit is called the *derivative* of f at a and is commonly denoted by $f'(a)$. If $f : S \to \mathbb{R}$ is differentiable at every point $a \in S$, we say that f is differentiable on S.

The expression

$$\frac{f(x) - f(a)}{x - a},$$

defined for $x \in S \setminus \{a\}$ is called a difference quotient.

The difference quotient has an obvious geometric interpretation as the slope of a secant passing through two points on the graph of f. The notation f' for the derivative of f was introduced by Newton[1]. We will not use Leibniz's[2] notation involving differentials (df/dx).

THEOREM 187. Let $f : S \to \mathbb{R}$ be differentiable at $a \in S$. Then f is continuous at a.

THEOREM 188. The constant function and the identity function are differentiable on $\mathbb{R}$. Their derivatives at any point are 0 and 1, respectively.

THEOREM 189. The function $h : \mathbb{R} \to [0, \infty) : x \mapsto |x|$ is differentiable on $\mathbb{R} \setminus \{0\}$ but not at zero.

THEOREM 190. The function $f : [0, \infty) \to [0, \infty) : x \mapsto \sqrt{x}$ is differentiable on $(0, \infty)$ but not at zero.

THEOREM 191. Let $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ be functions which are differentiable at $a \in S$ and let $c \in \mathbb{R}$ be a constant. Then

(1) cf is differentiable at a and $(cf)'(a) = cf'(a)$ and
(2) $f \pm g$ are differentiable at a, and $(f \pm g)'(a) = f'(a) \pm g'(a)$.

THEOREM 192 (Product or Leibniz rule for derivatives). Let $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ be functions which are differentiable at $a \in S$. Then, fg is differentiable at a and

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

[1]Isaac Newton (1643 – 1727)
[2]Gottfried Wilhelm Leibniz (1646 – 1716)

THEOREM 193 (Quotient rule for derivatives). Let $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ be functions which are differentiable at $a \in S$. Also assume that $g(a) \neq 0$. Then, f/g is differentiable at a and

$$(f/g)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{[g(a)]^2}.$$

THEOREM 194. All polynomials are differentiable on $\mathbb{R}$. In particular, if $n \in \mathbb{N}$ and $f(x) = x^n$, then

$$f'(x) = nx^{n-1}.$$

THEOREM 195. Any rational function is differentiable on its domain. In particular, if $n \in \mathbb{N}$ and $g(x) = x^{-n}$, then

$$g'(x) = -nx^{-n-1} \text{ for } x \in \mathbb{R} \setminus \{0\}.$$

Note that the rule for the differentiation of the power function with an integer power works for all $n \in \mathbb{Z}$.

THEOREM 196. Suppose $a \in S$, $f : S \to \mathbb{R}$, and $m \in \mathbb{R}$. Then f is differentiable at a with derivative m if and only if there exists a function $h : S \to \mathbb{R}$ which is continuous at a with $h(a) = 0$ such that

$$f(x) = f(a) + m(x-a) + (x-a)h(x).$$

Since, in a vicinity of a, the term $(x-a)h(x)$ is generally much smaller than $f(a) + m(x-a)$, the theorem says that $f(x)$ may be well approximated by $f(a) + m(x-a)$ if and only if it is differentiable at a. The function $x \mapsto f(a) + m(x-a)$ is called the tangent line of f at the point $(a, f(a))$.

THEOREM 197 (Chain Rule). Suppose S_1 and S_2 are intervals, $f : S_1 \to \mathbb{R}$, $g : S_2 \to \mathbb{R}$ and $f(S_1) \subset S_2$. If f is differentiable at a and g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a, and

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Let $f : S \to \mathbb{R}$ be a differentiable function on S. Then $x \mapsto f'(x)$ defines another function $f' : S \to \mathbb{R}$.

DEFINITION 55 (Higher order derivatives). Let $f : S \to \mathbb{R}$ be a differentiable function on S. If f' is differentiable at some point $a \in S$, then we denote its derivative at a by $f''(a)$, and we call it the *second order derivative* of f at a. The function f is then called *twice differentiable* at a.

Higher order derivatives may now be defined recursively. Assuming, for some $n \in \mathbb{N}$, that f has derivatives up to order n on S (denoted by $f^{(k)}$, $k = 1, ..., n$) and that the n-th order derivative is differentiable at a, one defines the $(n+1)$-st order derivative $f^{(n+1)}(a)$ as the derivative of $f^{(n)}$ at a. It is also customary to use $f^{(0)}$ instead of f.

THEOREM 198. Polynomials have derivatives of all orders on all of $\mathbb{R}$. In fact, if p is a polynomial of degree d, then $p^{(n)}(x) = 0$ for all $x \in \mathbb{R}$ and all integers $n > d$.

THEOREM 199. The exponential functions has derivatives of all orders on all of $\mathbb{R}$, in fact $\exp^{(n)}(x) = \exp(x)$ for all $n \in \mathbb{N}$ and all $x \in \mathbb{R}$.

5.2. The mean value theorem and Taylor's theorem

In this section a and b always denote two real numbers such that $a < b$.

Definition 56 (Local extrema of a function). Let $f : S \to \mathbb{R}$ be a function. We say that f has a *local minimum* (or a *local maximum*) at $x_0 \in S$, if there exists a number $\delta > 0$ such that $f(x) \geq f(x_0)$ (or $f(x) \leq f(x_0)$) for every $x \in (x_0 - \delta, x_0 + \delta) \cap S$. The point x_0 is called a *local extremum* if it is either a local minimum or a local maximum.

Theorem 200. Let $f : (a,b) \to \mathbb{R}$ be differentiable at $x_0 \in (a,b)$. If f has a local maximum or a local minimum at x_0, then $f'(x_0) = 0$.

Theorem 201 (Rolle's[3] Theorem). Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Suppose that $f(a) = f(b)$. Then, there exists $c \in (a,b)$ such that $f'(c) = 0$.

Theorem 202 (The mean value theorem). Let $f, g : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then there exists $c \in (a,b)$ such that

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

In particular, there is a $c \in (a,b)$ such that

$$f(b) - f(a) = f'(c)(b - a).$$

Theorem 203. Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Suppose that $f'(x) = 0$ for all $x \in (a,b)$. Then there exists $C \in \mathbb{R}$ such that $f(x) = C$ for all $x \in [a,b]$.

Theorem 204. Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then the following statements hold:

(1) If $f'(x) \geq 0$ for all $x \in (a,b)$, then f is non-decreasing on $[a,b]$.
(2) If $f'(x) > 0$ for all $x \in (a,b)$, then f is strictly increasing on $[a,b]$.

Analogous results hold, of course, if $f'(x) \leq 0$ or $f'(x) < 0$ for all $x \in (a,b)$.

Theorem 205. Let $f : (a,b) \to \mathbb{R}$ be differentiable and assume that $f'(x) > 0$ (or $f'(x) < 0$) for all $x \in (a,b)$. Then the inverse function g of f is differentiable on its domain and $g' = 1/(f' \circ g)$.

Theorem 206. The logarithmic function is differentiable on $(0,\infty)$ and $\log'(x) = 1/x$.

Theorem 207. Suppose $r \in \mathbb{R}$. Then the power function $p_r : x \mapsto x^r$ is differentiable on $(0,\infty)$ and $p_r'(x) = rx^{r-1}$.

Theorem 208 (Taylor's[4] theorem).* Let $f : (a,b) \to \mathbb{R}$ have $n+1$ derivatives on (a,b) and let x and x_0 be two distinct points in (a,b). Then there exists c between x and x_0 such that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

Theorem 209. Let $f : (a,b) \to \mathbb{R}$ be a function which together with its derivative is differentiable on (a,b) and whose second derivative is continuous on (a,b). Let $x_0 \in (a,b)$ and suppose $f'(x_0) = 0$. Then the following two statement are true.

(1) If $f''(x_0) < 0$, then f has a local maximum at x_0.
(2) If $f''(x_0) > 0$, then f has a local minimum at x_0.

[3]Michel Rolle (1652 – 1719)
[4]Brook Taylor (1685 – 1731)

5.3. Uniform convergence and differentiation

THEOREM 210. Define the functions $f_n : [0,1] \to \mathbb{R}$ by $f_n(x) = x^n/n$ for $n \in \mathbb{N}$. The sequence $n \mapsto f_n$ converges uniformly to the function $f = 0$, but the sequence $n \mapsto f_n'$ does not converge to $f' = 0$.

Note that the operations of taking a limit and taking a derivative do not necessarily commute.

THEOREM 211. Suppose that the functions $f_n : [a,b] \to \mathbb{R}$ are differentiable, that the sequence of functions $n \mapsto f_n'$ converges uniformly on $[a,b]$ and that, for some $x_0 \in [a,b]$, the numerical sequence $n \mapsto f_n(x_0)$ is also convergent. Then $n \mapsto f_n$ converges uniformly on $[a,b]$.

THEOREM 212. Suppose functions f_n with the properties of Theorem 211 are given. For fixed $x \in [a,b]$ and $n \in \mathbb{N}$ define $H_n : [a,b] \to \mathbb{R}$ by

$$H_n(t) = \begin{cases} (f_n(t) - f_n(x))/(t-x) & \text{if } t \neq x \\ f_n'(x) & \text{if } t = x. \end{cases}$$

Then each H_n is continuous and the sequence $n \mapsto H_n$ converges uniformly on $[a,b]$.

THEOREM 213. Suppose that the functions $f_n : [a,b] \to \mathbb{R}$ are differentiable, that the sequence of functions $n \mapsto f_n'$ converges uniformly on $[a,b]$ and that, for some $x_0 \in [a,b]$, the numerical sequence $n \mapsto f_n(x_0)$ is also convergent. Then $n \mapsto f_n$ converges uniformly on $[a,b]$ to a differentiable function on $[a,b]$ and, for each $x \in [a,b]$,

$$(\lim_{n\to\infty} f_n)'(x) = \lim_{n\to\infty} f_n'(x).$$

THEOREM 214. Suppose that the functions $g_k : [a,b] \to \mathbb{R}$ are differentiable, that the series of functions $\sum_{k=0}^{\infty} g_k'$ converges uniformly on $[a,b]$ and that, for some $x_0 \in [a,b]$, the numerical series $\sum_{k=0}^{\infty} g_k(x_0)$ also converges. Then $\sum_{k=0}^{\infty} g_k$ converges uniformly to a differentiable function on $[a,b]$ and, for each $x \in [a,b]$,

$$\left(\sum_{k=0}^{\infty} g_k\right)'(x) = \sum_{k=0}^{\infty} g_k'(x).$$

THEOREM 215. The power series $\sum_{n=0}^{\infty} c_n x^n$ and $\sum_{n=1}^{\infty} n c_n x^{n-1}$ have the same radius of convergence.

THEOREM 216. Let $a > 0$. Suppose the function $f : (-a,a) \to \mathbb{R}$ is defined by the power series $\sum_{n=0}^{\infty} c_n x^n$ whose radius of convergence is at least a. Then f has derivatives of all orders on $(-a,a)$.

THEOREM 217. sin and cos have derivatives of all orders on $\mathbb{R}$. In fact, $\sin' = \cos$ and $\cos' = -\sin$.

DEFINITION 57. Suppose $f : (a,b) \to \mathbb{R}$ has derivatives of all orders at $x_0 \in (a,b)$. Consider the power series

$$T_f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.$$

If there is a positive r such that $f(x) = T_f(x)$ for all $x \in (x_0 - r, x_0 + r)$ then f is called *analytic* at x_0 and T_f is called the *Taylor series* of f at x_0.

Analytic functions are the realm of Complex Analysis, perhaps the most beautiful subject in Mathematics.

THEOREM 218. The function $x \mapsto 1/(1-x)$ defined on $(-\infty, 1)$ is analytic at zero.

CHAPTER 6

Integration

Throughout this chapter a and b denote two real numbers such that $a < b$.

6.1. Existence and uniqueness of integrals

DEFINITION 58. A *partition* P of $[a, b]$ is a finite subset of $[a, b]$ which contains both a and b. If the number of elements in P is $n + 1$ we will label them so that

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b.$$

A partition P' is called a *refinement* of P if $P \subset P'$. $P \cup Q$ is called the *common refinement* of the partitions P and Q.

DEFINITION 59. If P is a partition of $[a, b]$ with $n + 1$ elements and $f : [a, b] \to \mathbb{R}$ is a bounded function we define the *lower Riemann sum* $L(f, P)$ and the *upper Riemann sum* $U(f, P)$ by

$$L(f, P) = \sum_{j=1}^{n} m_j (x_j - x_{j-1})$$

and

$$U(f, P) = \sum_{j=1}^{n} M_j (x_j - x_{j-1})$$

where $m_j = \inf\{f(x) : x_{j-1} \le x \le x_j\}$ and $M_j = \sup\{f(x) : x_{j-1} \le x \le x_j\}$.

Integration was put first on firm footing by Cauchy in 1823. The above definition is due to Darboux[1]. Riemann[2] sums are actually such where m_j or M_j is replaced by $f(t_j)$ for some $t_j \in [x_{j-1}, x_j]$.

THEOREM 219. If P is a partition of $[a, b]$ and $f : [a, b] \to \mathbb{R}$ a bounded function, then $L(f, P) \le U(f, P)$.

THEOREM 220. If P is a partition of $[a, b]$, P' a refinement of P, and $f : [a, b] \to \mathbb{R}$ a bounded function, then $L(f, P) \le L(f, P') \le U(f, P') \le U(f, P)$.

THEOREM 221. If P and Q are partitions of $[a, b]$ and $f : [a, b] \to \mathbb{R}$ a bounded function, then $L(f, P) \le U(f, Q)$.

THEOREM 222. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. Then the set $\{L(f, P) : P \text{ is a partition of } [a, b]\}$ is bounded above and thus has a supremum. Likewise, the set $\{U(f, P) : P \text{ is a partition of } [a, b]\}$ is bounded below and has an infimum.

---

[1] Jean-Gaston Darboux (1842 – 1917)
[2] Bernhard Riemann (1826 – 1866)

DEFINITION 60. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. The number

$$I_-(f, a, b) = \sup\{L(f, P) : P \text{ is a partition of } [a, b]\}$$

is called the *lower Riemann integral* of f over $[a, b]$ and the number

$$I_+(f, a, b) = \inf\{U(f, P) : P \text{ is a partition of } [a, b]\}$$

is called the *upper Riemann integral* of f over $[a, b]$.

THEOREM 223. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. Then $I_-(f, a, b) \leq I_+(f, a, b)$.

DEFINITION 61. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. If $I_-(f, a, b) = I_+(f, a, b)$ we say that f is *Riemann integrable* (or just integrable) over $[a, b]$. It is customary to denote the common value of $I_-(f, a, b)$ and $I_+(f, a, b)$ by

$$\int_a^b f(x)dx.$$

Leibniz thought of the integral sign $\int$ as an elongated s, the initial letter of sum. He thought of the integral as a sum of infinitely many "infinitesimally" narrow rectangles of width dx and height $f(x)$. Today we might as well use the notation $\int_a^b f$ but Leibniz's notation still has the following two advantages: (1) In physics, where quantities have units, it becomes clear what the unit of the integral is, e.g., if x denotes time and f velocity the integral will be a distance. (2) For the integrand x^r the use of dx indicates that one integrates $x \mapsto x^r$ rather than $r \mapsto x^r$.

THEOREM 224. Suppose $f : [a, b] \to \mathbb{R}$ is a bounded function. Then f is Riemann integrable if and only if, for every positive ε, there is a partition P such that

$$U(f, P) - L(f, P) < \varepsilon.$$

THEOREM 225. Let $c \in [0, 1]$ and suppose $f : [0, 1] \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} 0 & \text{for } 0 \leq x \leq c, \\ 1 & \text{for } c < x \leq 1. \end{cases}$$

Then f is Riemann integrable and $\int_0^1 f = 1 - c$.

THEOREM 226 (Dirichlet's function). The function $f : [0, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{for } x \in [0, 1] \cap \mathbb{Q}, \\ 0 & \text{for } x \in [0, 1] \cap (\mathbb{R} \setminus \mathbb{Q}) \end{cases}$$

is not Riemann integrable.

EXERCISE 17. Compute $I_\pm$ for the function given in Theorem 226.

THEOREM 227. The identity function on $[a, b]$ is Riemann integrable.

EXERCISE 18. Compute $\int_a^b x\, dx$.

THEOREM 228. Monotone functions on finite closed intervals are Riemann integrable.

DEFINITION 62. Let S be a subset of $\mathbb{R}$. Then $f : S \to \mathbb{R}$ is called *uniformly continuous* on S if for every $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x, y \in S$ for which $|x - y| < \delta$ it is true that $|f(x) - f(y)| < \varepsilon$.

THEOREM 229. Uniformly continuous functions on finite closed intervals are Riemann integrable.

As a matter of fact, the qualifier "uniformly" is not necessary in the previous theorem, because any function which is continuous on a finite closed interval is uniformly continuous there. However, to prove this fact one needs the Heine-Borel theorem, which we have not available. Instead, we will try for a different proof for the integrability of continuous functions in Theorem 241 following an idea advocated by C. Bennewitz[3].

The concept of the Riemann integral has proved to be too weak and too cumbersome for many applications in differential equations, probability, and elsewhere. It is too weak since many functions one might want to integrate are not Riemann integrable, e.g., Dirichlet's function. It is too cumbersome since it is sometimes difficult to show when limit processes can be interchanged with integration (despite what we show in Section 6.5). To remedy this Lebesgue[4] introduced a more powerful integral called the Lebesgue integral. This is the topic of more advanced courses in Real Analysis.

6.2. Properties of integrals

THEOREM 230. Suppose f is Riemann integrable over $[a,b]$ and $[r,s] \subset [a,b]$. Then f is Riemann integrable over $[r,s]$.

THEOREM 231. Suppose f is bounded on $[a,b]$ and $c \in (a,b)$. Then

$$I_\pm(f,a,c) + I_\pm(f,c,b) = I_\pm(f,a,b).$$

This theorem suggest we define $I_\pm(f,r,s)$ and $\int_r^s f$ also for the case where $r \geq s$.

DEFINITION 63. Suppose $f : [a,b] \to \mathbb{R}$ is a bounded function and $r,s \in [a,b]$. We set $\int_r^s f = 0$ for $r = s$ and $\int_r^s f = -\int_s^r f$ for $r > s$. Similarly, $I_\pm(f,r,r) = 0$, and $I_\pm(f,r,s) = -I_\pm(f,s,r)$.

THEOREM 232. If f and g are Riemann integrable over $[a,b]$, then so is $f+g$ and $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

THEOREM 233. If f is Riemann integrable over $[a,b]$ and $\alpha \in \mathbb{R}$, then αf is Riemann integrable over $[a,b]$ and $\int_a^b \alpha f = \alpha \int_a^b f$.

The properties in Theorems 232 and 233 combine to establish *linearity* of the Riemann integral: $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g$.

THEOREM 234 (Monotonicity). Suppose $f,g : [a,b] \to \mathbb{R}$ are Riemann integrable. If $f(x) \leq g(x)$ for all $x \in [a,b]$, then $\int_a^b f \leq \int_a^b g$.

DEFINITION 64. Given a real-valued function f we define its *positive* and *negative* parts f^+ and f^- by $f^+(x) = \max\{f(x),0\}$ and $f^-(x) = -\min\{f(x),0\}$.

Note that both f^+ and f^- are non-negative functions.

THEOREM 235. Suppose f is a real-valued function. Then $f^- = (-f)^+$, $f = f^+ - f^-$, and $|f| = f^+ + f^-$.

THEOREM 236. If f is Riemann integrable over $[a,b]$, then so are its positive and negative parts.

[3]Private communication and unpublished lecture notes

[4]Henri Lebesgue (1875 – 1941)

Theorem 237 (Triangle Inequality). Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable on $[a, b]$. Then $|f|$ is Riemann integrable over $[a, b]$ and

$$\left| \int_a^b f \right| \leq \int_a^b |f|.$$

6.3. The fundamental theorem of calculus

Theorem 238. Suppose f is bounded on $[a, b]$. Then the functions $F_\pm : [a, b] \to \mathbb{R} : x \mapsto I_\pm(f, a, x)$ are continuous on $[a, b]$.

Theorem 239. Suppose f is bounded on $[a, b]$. Then the functions $F_\pm : [a, b] \to \mathbb{R} : x \mapsto I_\pm(f, a, x)$ are differentiable at any point c where f is continuous in which case $F'_\pm(c) = f(c)$.

Definition 65. A function $F : [a, b] \to \mathbb{R}$ is called a *primitive* or an *antiderivative* of $f : [a, b] \to \mathbb{R}$ if it is continuous and if $F' = f$ on (a, b). The function $F : [a, b] \to \mathbb{R}$ is called a *piecewise primitive* or a *piecewise antiderivative*, if it is continuous and if there is a finite subset D of $[a, b]$ such that $F' = f$ on $[a, b] \setminus D$.

Theorem 240. Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable over $[a, b]$ and $F : [a, b] \to \mathbb{R}$ is a piecewise antiderivative of f. Then

$$\int_a^b f = F(b) - F(a).$$

The combination of Theorems 239 (for the case where $F_\pm$ coincide) and 240 is known as the *Fundamental Theorem of Calculus*. Its essence is that the operations of integration and differentiation are in some sense inverses of each other. But note that Theorem 239 makes no statement as to the size of the set where F' does not exist while Theorem 240 requires that set to be finite. This imbalance is a shortcoming of the Riemann integral and is only overcome by the more powerful Lebesgue integral.

6.4. Integration of piecewise continuous functions

Definition 66. A function $f : [a, b] \to \mathbb{R}$ is called piecewise continuous if there exists a finite subset D of $[a, b]$ such that f is continuous on $[a, b] \setminus D$.

Theorem 241. Bounded, piecewise continuous functions on finite closed intervals have piecewise antiderivatives and are Riemann integrable.

Theorem 242 (Integration by parts). Suppose $f, g : [a, b] \to \mathbb{R}$ are bounded and piecewise continuous on $[a, b]$ and let F and G, respectively, be piecewise antiderivatives. Then

$$\int_a^b Fg + \int_a^b fG = F(b)G(b) - F(a)G(a).$$

Theorem 243 (Substitution rule). Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and that $g : [a, b] \to \mathbb{R}$ is differentiable with a continuous derivative g'. Then

$$\int_{g(a)}^{g(b)} f = \int_a^b (f \circ g) g'.$$

6.5. Uniform convergence and integration

THEOREM 244. For $n \in \mathbb{N}$ define the functions $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 2n^2 x & \text{if } 0 \le x \le 1/n \\ 0 & \text{if } 1/n < x \le 1. \end{cases}$$

The sequence $n \mapsto f_n$ converges pointwise to the function $f = 0$, but the sequence $n \mapsto \int_0^1 f_n$ converges to 1.

The operations of integration and taking limits may not always be interchanged.

THEOREM 245. Suppose that $n \mapsto f_n : [a,b] \to \mathbb{R}$ is a sequence of functions which are Riemann integrable over $[a,b]$ and that this sequence converges uniformly to a function $f : [a,b] \to \mathbb{R}$. Then f is Riemann integrable and

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n.$$

THEOREM 246. Suppose that $n \mapsto g_n : [a,b] \to \mathbb{R}$ is a sequence of functions which are Riemann integrable over $[a,b]$ and that the series $\sum_{n=0}^{\infty} g_n$ converges uniformly to a function $g : [a,b] \to \mathbb{R}$. Then g is Riemann integrable and

$$\int_a^b g = \sum_{n=0}^{\infty} \int_a^b g_n.$$

THEOREM 247. For all x for which $|x| < 1$ we have

$$\log(1+x) = -\sum_{n=1}^{\infty} \frac{(-x)^n}{n}.$$

In particular, the function $x \mapsto \log(1+x)$ defined on $(-1, \infty)$ is analytic at zero.

CHAPTER 7

Special topics

7.1. Generalized limits

DEFINITION 67. Let a be a real number. If $r > 0$, the set $(a-r, a+r) = \{x : |x-a| < r\}$ is called a *neighborhood* of a. The sets (c, ∞), where $c \in \mathbb{R}$, are called neighborhoods of ∞ while the sets $(-\infty, c)$ are called neighborhoods of $-\infty$. The set of all neighborhoods of $a \in \mathbb{R} \cup \{-\infty, \infty\}$ is denoted by $\mathcal{N}(a)$.

DEFINITION 68. We say that ∞ is a limit point of $S \subset \mathbb{R}$ if S is not bounded above. Similarly, if S is not bounded below, we say that $-\infty$ is a limit point of S.

THEOREM 248. The point $a \in \mathbb{R} \cup \{-\infty, \infty\}$ is a limit point of $S \subset \mathbb{R}$ if and only if for every neighborhood V of a we have $V \cap S \setminus \{a\} \neq \emptyset$.

THEOREM 249. Suppose $S \subset \mathbb{R}$, $f : S \to \mathbb{R}$, $a \in \mathbb{R}$ is a limit point of S, and $b \in \mathbb{R}$. Then the following statement is true: If

$$\forall U \in \mathcal{N}(b) : \exists V \in \mathcal{N}(a) : \forall x \in S : x \in V \setminus \{a\} \Rightarrow f(x) \in U,$$

then f converges to b as x tends to a.

The same conclusion holds when $S = \mathbb{N}$ and $a = \infty$ (in which case f is a sequence).

Note that in place of $\forall x \in S : x \in V \setminus \{a\} \Rightarrow f(x) \in U$ we might as well have written $\forall x \in V \cap S \setminus \{a\} : f(x) \in U$. Because of the previous theorem we may now extend the definition of convergence to include cases where a and/or b are in $\{-\infty, \infty\}$.

DEFINITION 69. Let S be a subset of $\mathbb{R}$, $f : S \to \mathbb{R}$ a function, and $a \in \mathbb{R} \cup \{-\infty, \infty\}$ a limit point of S. We say that f converges to $b \in \mathbb{R} \cup \{-\infty, \infty\}$ as x tends to a if

$$\forall U \in \mathcal{N}(b) : \exists V \in \mathcal{N}(a) : \forall x \in V \cap S \setminus \{a\} : f(x) \in U.$$

THEOREM 250. Suppose $f : S \to \mathbb{R}$ converges to both b_1 and b_2 as x tends to a. Then $b_1 = b_2$.

DEFINITION 70. If $f : S \to \mathbb{R}$ converges to b as x tends to a, we call b the limit of f at a and we write $\lim_{x\to a} f(x) = b$.

THEOREM 251. $\lim_{x\to\infty} \mathrm{e}^{-x} = 0$.

THEOREM 252. The function $f : \mathbb{R} \to [0, \infty)$ defined by

$$f(x) = \begin{cases} \mathrm{e}^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is continuous on all of $\mathbb{R}$.

THEOREM 253. Let (a, b) be an open interval (allowing for $a = -\infty$ and $b = \infty$) and let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions. Assume $g'(x) \neq 0$ for all $x \in (a, b)$ and

$\lim_{x\to a} f'(x)/g'(x) = L$ for some real number L. Then g is one-to-one on (a,b) and, for every $\varepsilon > 0$, there is a point $c \in (a,b)$ such that for all $x, y \in (a,c)$ we have $g(x) \neq 0$ and

$$L - \varepsilon < \frac{f(x) - f(y)}{g(x) - g(y)} < L + \varepsilon$$

as long as $x \neq y$.

THEOREM 254 (L'Hôpital's[1] rule I). Let (a,b) be an open interval (allowing for $a = -\infty$ and $b = \infty$) and let $f, g : (a,b) \to \mathbb{R}$ be differentiable functions. Assume $g'(x) \neq 0$ for all $x \in (a,b)$, $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, and $\lim_{x\to a} f'(x)/g'(x) = L$ where $L \in \mathbb{R} \cup \{-\infty, \infty\}$. Then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

THEOREM 255 (L'Hôpital's rule II). Let (a,b) be an open interval (allowing for $a = -\infty$ and $b = \infty$) and let $f, g : (a,b) \to \mathbb{R}$ be differentiable functions. Assume $g'(x) \neq 0$ and $g(x) > 0$ for all $x \in (a,b)$, $\lim_{x\to a} g(x) = \infty$, and $\lim_{x\to a} f'(x)/g'(x) = L$ where $L \in \mathbb{R} \cup \{-\infty, \infty\}$. Then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

There are analogous theorems for the case where x tends to b.

The next two theorems compare the growth of positive powers with that of the logarithmic and exponential functions.

THEOREM 256. Suppose r is a positive real number. Then $\lim_{x\to\infty} \log(x)/x^r = 0$ and $\lim_{x\to 0} x^r \log(x) = 0$.

THEOREM 257. Suppose r is a real number. Then $\lim_{x\to\infty} x^r/\mathrm{e}^x = 0$.

THEOREM 258. The function f defined in Theorem 252 has derivatives of all orders on all of $\mathbb{R}$. In particular, $f^{(n)}(0) = 0$ for every non-negative integer n.

THEOREM 259. The function f defined in Theorem 252 has a Taylor series T_f at $x_0 = 0$ with infinite radius of convergence. $T_f(x) = f(x)$ holds for $x = 0$ but nowhere else.

This shows that analyticity of a function is an even stronger property than having derivatives of any order.

7.2. Trigonometric functions and their inverses

THEOREM 260 (Pythagoras[2]). $(\sin x)^2 + (\cos x)^2 = 1$ for all $x \in \mathbb{R}$.

It follows immediately from this that $|\sin x| \leq 1$ and $|\cos x| \leq 1$.

THEOREM 261 (Addition theorems). Let $x, y \in \mathbb{R}$. Then

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$$

and

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y).$$

THEOREM 262. cos has a unique zero in $(0,2)$.

[1] Guillaume François Antoine, Marquis de l'Hôpital (1661 – 1704)
[2] Pythagoras of Samos (ca. 570 – ca. 495 BC)

Definition 71. The unique zero of cos in $(0,2)$ is denoted by $\pi/2$, i.e., π is twice that zero.

Theorem 263. $\pi > 0$, $\cos(\pi/2) = 0$, and $\sin(\pi/2) = 1$.

Theorem 264. $\sin(x + \pi/2) = \cos(x)$, $\sin(x + \pi) = -\sin(x)$, and $\sin(x + 2\pi) = \sin(x)$ for all $x \in \mathbb{R}$.

Theorem 265. $\sin(x) = 0$ if and only if $x = m\pi$ with $m \in \mathbb{Z}$.

Theorem 266. Suppose $\sin(x+a) = \sin(x)$ for all $x \in \mathbb{R}$. Then a is an integer multiple of 2π.

Theorem 267. sin is strictly increasing on $[-\pi/2, \pi/2]$. Also, cos is strictly decreasing on $[0, \pi]$.

Definition 72. A function $f : (a, b) \to \mathbb{R}$ is called *convex* on (a, b), if

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

for all $x, y \in (a, b)$ and all $t \in [0, 1]$. It is called *concave* on (a, b), if $-f$ is convex on (a, b).

Note that the graph of the function $t \mapsto (1-t)f(x) + tf(y)$ describes a straight line segment joining $f(x)$ and $f(y)$ in the coordinate plane. Convexity means that the graph of a function lies below any such segment.

Theorem 268. f is convex on (a, b) if and only if

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(v)}{w - v}$$

whenever $a < u < v < w < b$.

Theorem 269. If $f : (a, b) \to \mathbb{R}$ is twice differentiable and $f'' \geq 0$ on (a, b), then f is convex on (a, b).

Theorem 270. The sine function is positive in $(0, \pi)$ and negative in $(\pi, 2\pi)$. It is strictly increasing in $(0, \pi/2)$ and $(3\pi/2, 2\pi)$ and strictly decreasing in $(\pi/2, 3\pi/2)$. Finally, it is concave in $(0, \pi)$ and convex in $(\pi, 2\pi)$.

This theorem allows to draw a sketch of the sine function on $[0, 2\pi]$ and by periodicity on all of $\mathbb{R}$. The graph of the cosine function is simply obtained by shifting the graph of the sine function.

Theorem 271. tan is defined on $\mathbb{R} \setminus \{(2m+1)\pi/2 : m \in \mathbb{Z}\}$ and has range $\mathbb{R}$. Moreover, $\tan(x + \pi) = \tan(x)$ for all x in its domain.

Theorem 272. The tangent function is negative and concave in $(-\pi/2, 0)$ and positive and convex in $(0, \pi/2)$. It is strictly increasing in the entire interval $(-\pi/2, \pi/2)$.

Definition 73 (Inverse trigonometric functions). arcsin, arccos, and arctan are the inverse functions of $\sin|_{[-\pi/2,\pi/2]}$, $\cos|_{[0,\pi]}$, and $\tan|_{(-\pi/2,\pi/2)}$, respectively.

Theorem 273. $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ is strictly increasing and continuous. It is differentiable on $(-1, 1)$. In fact $\arcsin'(x) = 1/\sqrt{1 - x^2}$.

Theorem 274. $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$ is strictly increasing and differentiable. In fact $\arctan'(x) = 1/(1 + x^2)$.

Theorem 275. If $|x| < 1$, then $\arctan(x) = \sum_{n=0}^{\infty} (-1)^n x^{2n+1}/(2n + 1)$.

Theorem 276. $\pi/4 = \arctan(1) = \arctan(1/2) + \arctan(1/3)$.

Theorem 277. $\frac{505}{162} \leq \pi \leq \frac{6115}{1944}$.

7.3. Analytic geometry

THEOREM 278. The function

$$d : \mathbb{R}^2 \times \mathbb{R}^2 \to [0,\infty) : ((x_1, y_1), (x_2, y_2)) \mapsto \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

is a metric on $\mathbb{R}^2$.

DEFINITION 74. The function d defined in Theorem 278 is called the *Euclidean metric* or Euclidean distance in $\mathbb{R}^2$.

DEFINITION 75. The set of all points in $\mathbb{R}^2$ which have a given distance $r > 0$ from a given point $p_0 = (x_0, y_0)$ is called a *circle* of *radius* r about p_0. The set of all points whose distance from p_0 is less than r is called the (open) *disk* of radius r about p_0. The circle of radius 1 about the point $(0, 0)$ is called the *unit circle*. The *unit disk* is the set of all points whose distance from $(0, 0)$ is less than 1.

THEOREM 279. The unit circle is the set

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} = \{(\cos(t), \sin(t)) : t \in [0, 2\pi]\}.$$

DEFINITION 76. A *plane curve* is an ordered pair (x, y) of continuous functions defined on some interval $[a, b]$.

DEFINITION 77. The curve $p_2 : [a_2, b_2] \to \mathbb{R}^2$ is called equivalent to the curve $p_1 : [a_1, b_1] \to \mathbb{R}^2$ if there exists a strictly increasing, continuous, and surjective function $\varphi : [a_1, b_1] \to [a_2, b_2]$ such that $p_2 \circ \varphi = p_1$.

THEOREM 280. Equivalence of curves is an equivalence relation, i.e., it has the following three properties:

(1) Any curve is equivalent to itself (reflexivity).
(2) If the curve p_2 is equivalent to the curve p_1, then p_1 is equivalent to p_2 (symmetry).
(3) If p_2 is equivalent to p_1 and p_3 is equivalent to p_2, then p_3 is equivalent to p_1 (transitivity).

DEFINITION 78. The curve $(x, y) : [a, b] \to \mathbb{R}^2$ is called continuously differentiable if the derivatives of x and y (the components of the curve) exist and are continuous. The pair (x', y') is then called the *velocity* of the curve and the function $t \mapsto \sqrt{x'(t)^2 + y'(t)^2}$ is called the *speed* of the curve.

EXERCISE 19. Note that, in view of Theorem 279 or a proper generalization of it, a circle may be viewed as a continuously differentiable curve.

DEFINITION 79. The *length* of the continuously differentiable curve $(x, y) : [a, b] \to \mathbb{R}^2$ is defined as

$$\int_a^b \sqrt{x'^2 + y'^2}.$$

Note that for constant speed v the length of a curve is $v(b - a)$ (speed times time).

THEOREM 281. Suppose two curves are continuously differentiable, equivalent to each other, and that the function φ establishing the equivalence has a continuous derivative. Then the curves have the same length.

The importance of this observation is that it shows length to be a property of (certain) geometric objects in the plane rather than of functions with values in the plane.

THEOREM 282. The length of the unit circle (its circumference) is 2π.

DEFINITION 80. Let $f, F : [a, b] \to \mathbb{R}$ be two continuous functions with the property $f(x) \le F(x)$. Suppose $S = \{(x, y) \in \mathbb{R}^2 : a \le x \le b \text{ and } f(x) \le y \le F(x)\}$. Then the *area* of S is given by

$$A = \int_a^b (F - f).$$

EXERCISE 20. Suppose $f(x) = 0$ and $F(x) = h$. What kind of set is S and what is its area?

THEOREM 283. The area of the unit disk is π.

7.4. Sets of measure zero and some consequences

DEFINITION 81. A subset A of $\mathbb{R}$ is said to have *measure zero* if for every $\varepsilon > 0$ there exist bounded open intervals (a_n, b_n), $n \in \mathbb{N}$, such that $A \subset \bigcup_{n=1}^{\infty}(a_n, b_n)$ and $\sum_{n=1}^{\infty}(b_n - a_n) < \varepsilon$.

THEOREM 284. The set of rational numbers has measure zero.

THEOREM 285. A countable union of sets of measure zero has itself measure zero.

DEFINITION 82. For any function $f : [a, b] \to \mathbb{R}$ and any $n \in \mathbb{N}$ let

$$\mathcal{D}(f, n) = \{x \in [a, b] : \forall\delta > 0 : \exists y \in [a, b] : |y - x| < \delta \text{ and } |f(y) - f(x)| > 1/n\}.$$

THEOREM 286. If $f : [a, b] \to \mathbb{R}$, then $\bigcup_{n=1}^{\infty} \mathcal{D}(f, n)$ is the set of all points for which f fails to be continuous.

THEOREM 287. Suppose f is a bounded real-valued function on $[a, b]$, $(c, d) \subset [a, b]$, and $\mathcal{D}(f, n) \cap (c, d) \neq \emptyset$. Then $\sup\{f(x) : c \le x \le d\} - \inf\{f(x) : c \le x \le d\} > 1/n$.

THEOREM 288. If f is Riemann integrable on $[a, b]$, then $\mathcal{D}(f, n)$ is a set of measure zero.

THEOREM 289. If f is Riemann integrable on $[a, b]$ the set of all points where f is not continuous is a set of measure zero.

This theorem shows that, in some sense, the class of Riemann integrable functions is quite small, emphasizing again the need for a different notion of integral. The converse of this theorem is also true but the proof of this requires a little more preparation so we will skip it.

THEOREM 290.* Suppose that there are intervals (a_n, b_n), $n \in \mathbb{N}$, such that $[0, 1] \subset \bigcup_{n=1}^{\infty}(a_n, b_n)$. Then there is an $N \in \mathbb{N}$ such that $[0, 1] \subset \bigcup_{n=1}^{N}(a_n, b_n)$.

THEOREM 291. The interval $[0, 1]$ does not have measure zero.

THEOREM 292. The set of real numbers is uncountable.

APPENDIX A

Some set theory and logic

In this section we briefly recall some notation and a few facts from set theory and logic[1]. Some of what is said below is intentionally vague lest we should write a book on the matter.

A.1. Elements of logic

A *statement* is a sentence for which it can (in principle) be determined whether it is true or false ("true" and "false" are called *truth values*). Statements may be connected to form new statements. If p and q are statements we may form the statement "p and q", the statement "p or q", and the statement "not p" (*negation*). Very important is also the connective "implies": "p implies q" is an abbreviation for "q or not p" (this entails no logical implication). Propositional logic deals with determining the truth values of such compound statements given the status of their components.

Matters become more interesting when the internal structure of statements is also considered. The simplest statements with internal structure are those where something (the *predicate*) is said about something else (the *subject*). For instance, the sentences "London is a city." and "Rome is the capital of Georgia." are such statements. Next one introduces *variables*, i.e., symbols meant to represent something unspecified. A *formula* is a sentence containing a variable, which becomes a statement after specifying the variable. For instance, in the sentence "$x^2 + 16 = 25$." x represents a number but we do not specify which one. At this point logic becomes interwoven with set theory since, in general, we need to make clear what kind of thing a variable is.

There are two more important ways of turning formulas into statements. If $F(x)$ is a formula involving the variable x we write $\forall x : F(x)$ to mean that $F(x)$ becomes a true statement whenever x is replaced by any specific value. We write $\exists x : F(x)$ to mean that there is some specific value, say v, such that $F(v)$ is a true statement. The negation of the statement "$\forall x : F(x)$" is the statement "$\exists x :$ not $F(x)$" and the negation of "$\exists x : F(x)$' is "$\forall x :$ not $F(x)$".

Both $\forall x : F(x)$ and $\exists x : F(x)$ are statements despite the occurrence of a variable. Such variables are called *bound*. For example, the statement " Every state has a capital." would be formalized as $\forall x : \exists y : y$ is the capital of x. The variables are there only for convenience; notice that the plain English version makes no explicit use of them. A variable which is not bound is called *free*. A sentence containing a free variable is a formula not a statement. To give another example consider the formula $\sum_{k=1}^n k^2 = n^2 + n/2$. Here k is a bound variable but n is a free variable. One may want to determine the truth value of "$\sum_{k=1}^n k^2 = n^2 + n/2$" for a specific n; in contrast, it makes no sense to specify k.

[1]For more details you may check my lecture notes on Algebra (to be found on my website) and the books referenced there.

A.2. Basics of set theory

We will not formally define the terms set and element but we think of a *set* (or a collection or a family) as a single entity which collects a variety of other entities, the *elements* (or members) of the set. A little more precisely, we assume that a set determines its elements and vice versa. In particular, two sets A and B are equal if and only if they contain precisely the same elements. We will use the following phrases: "x is an element (or a member) of (the set or collection) A", "x belongs to A", or "A contains x" and this relationship is denoted by $x \in A$. Otherwise, if x does not belong to A, we write $x \notin A$. The set with no elements is called the *empty set* and is denoted by $\emptyset$.

There are essentially two ways to specify a set. Firstly, one can list all the elements of a set, e.g., $\{2, 3, 5\}$ is the set containing the numbers 2, 3, and 5. Secondly, a set might collect elements which all share a certain property or certain properties. For instance, $\{x : x \text{ is a city}\}$ denotes the set of all cities.

If A and B are sets and every element of B is an element of A, we say that B is a *subset* of A or that A *includes* B, and we write $B \subset A$ or $A \supset B$. So, to check if $B \subset A$ one picks an *arbitrary* element $x \in B$ and shows that it is in A. The meaning of 'arbitrary' here is simply that the only fact we use about x is that it is in B. In this way we have checked the inclusion simultaneously for all $x \in B$. A subset B of A is called a *proper* subset of A if it is different from A.

Let $\mathcal{F}$ be a non-empty collection of sets. Then we define the *intersection* of all sets in $\mathcal{F}$ by

$$\bigcap_{A \in \mathcal{F}} A = \{x : x \in A \text{ for every } A \in \mathcal{F}\} = \{x : (\forall A \in \mathcal{F} : x \in A)\}.$$

If $\mathcal{F} = \{A, B\}$ we denote their intersection simply by $A \cap B$. If $A \cap B = \emptyset$ we say that A and B are *disjoint*. Similarly, the *union* of all sets in $\mathcal{F}$ is defined by

$$\bigcup_{A \in \mathcal{F}} A = \{x : x \in A \text{ for some } A \in \mathcal{F}\} = \{x : (\exists A \in \mathcal{F} : x \in A)\}.$$

The union of two sets A and B is denoted by $A \cup B$.

THEOREM A.1. Let A, B and C be sets and $\mathcal{F}$ a non-empty collection of sets. Then the following statements hold:

(1) Commutative laws:

$$A \cup B = B \cup A \quad \text{and} \quad A \cap B = B \cap A.$$

(2) Associative laws:

$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{and} \quad (A \cap B) \cap C = A \cap (B \cap C).$$

(3) Distributive laws:

$$A \cup \Big(\bigcap_{X \in \mathcal{F}} X\Big) = \bigcap_{X \in \mathcal{F}} (A \cup X) \quad \text{and}$$
$$A \cap \Big(\bigcup_{X \in \mathcal{F}} X\Big) = \bigcup_{X \in \mathcal{F}} (A \cap X).$$

The set $X \setminus A$ consists of all the elements of X which do not belong to A, i.e., $x \in X \setminus A$ if and only if $x \in X$ and $x \notin A$. If A is a subset of X, then $X \setminus A$ is called the *complement* of A in X. If it is clear what X is, then the complement is also denoted by A^c.

THEOREM A.2 (De Morgan's laws). Let E be a set and suppose that $\mathcal{F}$ is a non-empty collection of subsets of E. Then De Morgan's laws hold:

$$\left(\bigcup_{X\in\mathcal{F}} X\right)^c = \bigcap_{X\in\mathcal{F}} X^c,$$

$$\left(\bigcap_{X\in\mathcal{F}} X\right)^c = \bigcup_{X\in\mathcal{F}} X^c.$$

These formulas can be expressed concisely, if not precisely, as follows: The complement of a union is the intersection of the complements and the complement of an intersection is the union of the complements.

A.3. Functions

Let X and Y be sets. An *ordered pair* is an object of the form (x, y) where $x \in X$ and $y \in Y$. The set of all such ordered pairs (x, y) is called the *Cartesian product* of X and Y, and is written $X \times Y$. A subset f of $X \times Y$ is a *function*, if for each $x \in X$ there is a unique $y \in Y$ such that $(x, y) \in f$. As the concept of function is extremely important we repeat the preceding statement in a more formal way: $f \subset X \times Y$ is a function if and only if

(i) $\forall x \in X : \exists y \in Y$ such that $(x, y) \in f$ and
(ii) if $(x, y_1) \in f$ and $(x, y_2) \in f$, then $y_1 = y_2$.

The more conventional notation for functions is $y = f(x)$ instead of $(x, y) \in f$. If $f \subset X \times Y$ is a function, we write $f\colon X \to Y$ or, to define it precisely, $f : X \to Y : x \mapsto y = f(x)$, e.g., $f : \mathbb{R} \to [0, 1] : x \mapsto 1/(1 + x^2)$. We call X the *domain* of f, Y the *codomain* or *target* of f, and say that f maps X to Y. The element $f(x) \in Y$ is called the *image* of x under f. The set of those $y \in Y$ for which there is an $x \in X$ such that $y = f(x)$ is called the *range* or *image* of f, and will be written as $\operatorname{ran}(f)$. Let $f\colon X \to Y$. If $A \subset X$, we define the function $f|_A\colon A \to Y$, called the *restriction* of f to A, by setting $(f|_A)(x) = f(x)$ for every $x \in A$. We define the *image* of A under f as the set

$$f(A) = \{y \in Y : \exists x \in A : f(x) = y\} = \{f(x) :\ x \in A\}.$$

In other words $f(A) = \operatorname{ran}(f|_A)$.

THEOREM A.3. Let $f\colon X \to Y$ and $A, B \subset X$. Then $f(A \cup B) = f(A) \cup f(B)$ and $f(A \cap B) \subset f(A) \cap f(B)$.

Equality does not always hold in the second case. Can you think of an example where it does not?

If $A \subset Y$, we define the *pre-image* of A under f, denoted by $f^{-1}(A) \subset X$, by:

$$f^{-1}(A) = \{x \in X :\ f(x) \in A\}.$$

THEOREM A.4. Let $f\colon X \to Y$ and $A, B \subset Y$. Then $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$ and $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

A function $f\colon X \to Y$ is *onto* (or *surjective*) if $\operatorname{ran}(f) = Y$. A function $f\colon X \to Y$ is *one-to-one* (or *injective*) if $f(x_1) = f(x_2)$ implies that $x_1 = x_2$. If f is both one-to-one and onto, we say that f is *bijective*. A bijective function from a finite set X to itself is called a *permutation* of X. A permutation π is called a *transposition* if $\pi(x) = x$ for all but two elements x of the domain of π.

Suppose now that $f : X \to Y$ is an injective function. Then there exists a unique function $g : f(X) \to X$ such that $g(f(x)) = x$ for all $x \in X$, and $f(g(y)) = y$ for all

$y \in f(X)$. The function g is denoted f^{-1}, and is called the *inverse function* of f. Note that notation is abused here in the sense that f^{-1} is used to describe two different things. Only the context tells which of the two meanings is used. In any case, caution is advised. These two different meanings are exemplified in

$$f^{-1}(\{y\}) = \{f^{-1}(y)\}$$

which holds for all $y \in Y$ if $f \colon X \to Y$ is bijective.

Given a function $f : X \to Y$ and a function $g : Y \to Z$ their *composition* $g \circ f$ is the function from X to Z defined by

$$(g \circ f)(x) = g(f(x)).$$

THEOREM A.5. The inverse of an injective function is injective. The composition of injective functions is injective and the composition of surjective functions is surjective.

A.4. The recursion theorem

The following theorem states that it is ok to make recursive definitions. It is not strictly a part of set theory but requires also a proper definition of the set of natural numbers.

THEOREM A.6 (The Recursion Theorem). Let X be a non-empty set, f a function from X to X, and x_1 an element of X. Then there is one and only one function $u : \mathbb{N} \to X$ such that $u(1) = x_1$ and $u(n+1) = f(u(n))$ for every $n \in \mathbb{N}$.

Index