MATH360. ADVANCED CALCULUS

A. A. KIRILLOV

1. Glossary

1. Sets:
empty set
intersection, union, difference
direct (Cartesian) product
equivalent
complement
cardinality
denumerable, countable, uncountable

2. Relations:
graph of a relation
reflexive, transitive
symmetric, antisymmetric
equivalence, equivalence classes

3. Sequences:
subsequence
convergence, limit
bounded, — above, — below
upper bound, lower bound
least upper bound (l.u.b.), greatest lower bound (g.l.b.)
increasing, nondecreasing, decreasing, nonincreasing
strictly increasing, strictly decreasing
monotone, strictly monotone
limit superior, limit inferior
cluster point
Cauchy sequence (= fundamental sequence)
complete field, isomorphism of fields

4. Functions:
source (domain), target
injective (1-1), surjective (onto)
bijective, inverse
characteristic
graph of a function
domain, image
preimage

5. Order
partial, linear
well-ordered set

6. Induction principle

Notations and symbols:

$\emptyset$, $\aleph_0$, $\setminus$, $\forall$, $\exists$, $\approx$, $\subset$, $\supset$, $\in$, $\notin$, $:=$, $=:$, $\cap$, $\cup$, $\times$,
&, $\Rightarrow$, $\to$, $\mapsto$, Card, lim, lim sup, lim inf
$\mathbf{N}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

Date: Fall Semester 2012.

2. Necessity of rigorous definitions

In this non-mandatory section I tell you the popular joke which imitates a mathematical proof of evidently wrong statement:

Theorem. *A crocodile is wider than longer.*

The "proof" follows from two lemmas

Lemma 1. *A crocodile is wider than greener.*

Proof. A crok is wide from above and from below, but it is green only from above.

Lemma 2. *A crocodile is greener than longer.*

Proof. A crok is green along and across, but it is long only along.

The defect of the argument above is the absence of accurate definitions of words "wide", "long" and "green" and relations "wider", "longer" and "greener".

The main lesson from this story is:

There is no correct proof without definitions.

3. Sets

The notion of a *set* is the most fundamental in mathematics. In the same time it is very deep and non-trivial notion. Most textbooks say that the word "set" is a synonym of the words "collection", "group", "ensemble", "gathering", "family" etc. It is slightly misleading because it is only a description, not a rigorous definition. Actually, we can not define the notion "set" in terms of some simpler objects (if we do, we should then give the definition of these simpler objects and so on...) Instead, we just agree, in which context we can use the term "set" and what are possible manipulation with this notion.

The main feature of a set is that it contains some elements. The fact that a is an element of the set A is written by the formula $a \in A$. We say also that a *belongs* to A or that A *contains* a. Otherwise, we say that a does not belong to A and write $a \notin A$.

If $a \in A$ implies $a \in B$ (in other words, if every element of A is an element of B), we write $A \subset B$ or $B \supset A$ and say that A is a subset of B. We say that A is equal B and write $A = B$ if $A \subset B$ and $B \subset A$. So, two sets are equal if they have the same amount of elements.

It turns out that if we consider all "collections", "groups", "ensembles", "gatherings", "families" as sets, we come to contradiction. So, we must be cautious and admit that not all "collections" can be acknowleged as sets.

On the other hand, we admit the following constructions of new sets from the sets already acknowleged.

1. The *union* of two sets A and B is denoted by $A \bigcup B$ and is defined by the property:

$$c \in A \bigcup B \iff c \in A \text{ or } c \in B.$$

2. The *intersection* of two sets A and B is denoted by $A \bigcap B$ and is defined by the property:

$$c \in A \bigcap B \iff c \in A \text{ and } c \in B.$$

3. The *product* of two sets A and B is denoted by $A \times B$ and consists of all pairs (a, b) where $a \in A$, $b \in B$.

These construction can be generalized to the case where we have not two, but three, four or even infinitely many sets. More precisely, suppose we have a set I whose elements are themselves sets. It is convenient to denote element $i \in I$, when it is considered as an independent set, by S_i. Then we define the union, the intersection and the product

$$\bigcup_{i \in I} S_i, \qquad \bigcap_{i \in I} S_i, \qquad \prod_{i \in I} S_i$$

of the family of sets $\{S_i\}_{i \in I}$ in the natural way. Namely:

$c \in \bigcup_{i \in I} S_i \iff \exists i \in I$ such that $c \in S_i$;

$c \in \bigcup_{i \in I} S_i \iff c \in S_i \ \forall i \in I$;

$C \in \prod_{i \in I} S_i \iff C$ is a set such that $\forall i \in I$ the intersection $C \bigcap S_i$ is a one-point set (we assume here that the sets S_i are disjoint for different i's).

Note, that the existence of such a set C is garanteed by the Axiom of Choice (which also gives a way to construct new legal sets).

We also introduce the operation of difference: by $A \backslash B$ we denote the set of all elements $a \in A$ which do not belong to B.

Finally, we assume that for any set A there exists a set $P(A)$ whose elements are all subsets of A.

Let us see, what is the minimal amount of sets which must appear in any set theory. If there are no sets at all, the theory is not interesting. So, we suppose that at least one set A exists. Then the difference $A \backslash A$ also exists. It is rather special set: it has no elements at all. This set is called *empty set* and is denoted by the symbol $\emptyset$. It is clear that

$$\emptyset \bigcap \emptyset = \emptyset \bigcup \emptyset = \emptyset \backslash \emptyset = \emptyset \times \emptyset = \emptyset.$$

So, most of our constructions did not produce anything new starting with empty set.

But the operation P does. Indeed, the set $P(\emptyset)$ is not an empty set! It contains exactly one element $\emptyset$. The set $P(P(\emptyset))$ contains two elements: $\emptyset$ and $\{\emptyset\}$.

Definition 1. *We say that a set A' is a* **successor** *of a set A if $A' \supset A$ and $A' \setminus A$ contains exactly one element.*

In particular, $P(\emptyset)$ is a successor of $\emptyset$ and $P(P(\emptyset))$ is a successor of $P(\emptyset)$.

We assume that in our set theory every set has a successor and any non-empty set is a successor of some other set.

Then we postulate the existence of the set $\mathbf{N}$ with the following properties (the Peano axioms for the natural numbers):

• There is an element $0 \in \mathbf{N}$.

• Every element $a \in \mathbf{N}$ has a natural successor, denoted by S(a). (Intuitively, S(a) is a+1).

• There is no element whose successor is 0.

• S is injective, i.e. distinct elements have distinct successors: if $a \neq b$, then $S(a) \neq S(b)$.

• If a property is possessed by 0 and also by the successor of every element which possesses it, then it is possessed by all elements of $\mathbf{N}$.

This postulate ensures that the proof technique of mathematical induction is valid.

3.1. Relations, maps, equivalence of sets, cardinality.

Definition 2. *We call* **relation** *between sets A and B any subset $R \subset A \times B$.*

Instead $(a, b) \in R$ we say: a and b are in the relation R and write aRb. Some special kind of relations are very often used in mathematics. To describe these relation we need some preparations. Let us imagine the product set $A \times B$ as a rectangle (we imagine the first factor A as a segment of x-axis and the second factor B as a segment of y-axis). Then it is natural to call the subset $\{a\} \times B \subset A \times B$ a *vertical segment* and the subset $A \times \{b\} \subset A \times B$ a *horizontal segment*.

Assume that the relation $R \subset A \times B$ has the property: the intersection of R with any vertical segment $\{a\} \times B$ consists of single point. This point must have the form (a, b) where $b \in B$ is uniquely determined by the point a (and relation R). In this case, we say that R is the *graph* of a *function* f from A to B. (Instead of "function" the words "map", "operator" and "transformation" are also used).

Definition 3. *A function $f : A \to B : a \mapsto b = f(a)$ is called:*

injective*, or "1-1" if $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$;*

surjective*, or "onto" if $\forall b \in B \quad \exists a \in A$ s.t. $f(a) = b$;*

bijective *if it is both injective and surjective. In this case the so-called inverse function $\overset{-1}{f}$ is defined by the equation $b = \overset{-1}{f}(a)$.*[1]

---

[1]Here we use the notation $\overset{-1}{f}$ instead of usual f^{-1} which can be mistaken with $\frac{1}{f}$.

Definition 4. *A set A is called* **equivalent** *to a set B if there exists a bijective map $f \to B$; notation: $A \approx B$.*

Definition 5. *The common property of all sets which are equivalent to a given set A is called* **cardinality** *of A and denoted* $\operatorname{Card} A$ *or* $\#A$ *or* $|A|$.

3.2. **Finite sets.** *A set M is called* **finite** *if it is not equivalent to its proper subset.*[2]

Exercise 1. *Show that the empty set $\emptyset$ is finite.*

Hint. *Use the fact: $\emptyset$ has no proper subsets.*

Exercise 2. *Show that a subset of a finite set is finite.*

Hint. *Proof by contradiction: show that if a subset $S \subset M$ is not finite, then M itself is not finite.*

Exercise 3. *Show that a set A is finite iff its successor A' is finite.*

We can now see that the set of cardinalities of finite sets is equivalent to the set $\mathbf{N}$ introduced above. In the same time, the set of cardinalities of non-empty finite sets is equivalent to $\mathbb{N}$. Note also, that a set A can have many different successors, but all of them have the same cardinality, namely, the unique successor of $\operatorname{Card}(A)$.

Some of the elements of $\mathbf{N}$ have standard notations. Namely, cardinality of the empty set is usually denoted by 0 (zero), the cardinality of $P(\emptyset)$ is denoted by 1 (one). It is convenient to identify elements of $\mathbf{N}$ with corresponding cardinalities. Then the operation S of taking the successor in $\mathbf{N}$ will correspond the successor operation $A \to A'$ for sets. In other word, $\operatorname{Card}(A') = S(\operatorname{Card}(A))$. We also say that $m < n$ for two elements of $\mathbf{N}$ if n can be obtained from m by iteration of operation S.

We denote by X_n, $n \in \mathbb{N}$ the set, consisting of 0 and all its successors up to n:

$$X_n = \{0, 1, 2, \dots, n\}.$$

Sometimes it is convenient to use instead of X_n the equivalent set $Y_n = \{x \in \mathbf{N} \mid x < n\}$.

Exercise 4. *Show that a set A is finite iff it is equivalent to one of the sets X_n, $n \in \mathbf{N}$.*

Exercise 5. *All sets X_n, $n \in \mathbf{N}$ are pairwise non-equivalent.*

It can be proved by induction (see below).

Let us call a set A *infinite* if it is not finite, i.e. it is equivalent to one of its proper subsets. E.g. $\mathbb{N}$ is an infinite set because it is equivalent to the subset $2\mathbb{N} \neq \mathbb{N}$ via the map $\alpha : x \mapsto 2x$.

---

[2] A subset S of a set M is called *proper* if $S \neq M$.

3.3. **Comparison of cardinalities.** There is the natural order in the cardinalities. Namely, we say that Card $A \leq$ Card B if one of the following holds:

1. B has a subset C which is equivalent to A.
2. $\exists f : A \to B$ which is injective.
3. $\exists f : B \to A$ which is surjective.

Exercise 6. *Show that the above properties are equivalent.*

Theorem 1. $\aleph_0 :=$ Card $\mathbb{N}$ *is the minimal infinite cardinality.*

Proof. First we show that if A is a finite set, then Card$A \leq \aleph_0$. For this we use a bijective map $f : A \to X_n \subset \mathbb{N}$ which can be defined for any finite set A. For any infinite set A one can construct an injective map $f : \mathbb{N} \to A$ and we are done. (We skip the details which use the Induction principle – see below). □

Theorem 2. *For any non-empty set A we have* Card $\mathcal{P}(A) >$ Card A.

Proof. By contradiction. Assume that Card $A \geq$ Card $\mathcal{P}(A)$. Then there is a injective map $f : \mathcal{P}(A) \to A$. In other words, we can label subsets of A by elements of A. Let us call the subset B **normal** if it does not contain its label $f(B)$. Now, consider the set N of all labels of normal subsets. If it is normal, then $f(N) \in N$ which implies that N is not normal. If it is not normal, then $f(N) \notin N$ which implies that N is normal. Contradiction. □

In fact, we can make this argument constructive. We show that for any map $g : A \to \mathcal{P}(A)$ there is some $N \in \mathcal{P}(A)$ which has no label, i.e. is not in the range of g. Namely, define N as the set of those $a \in A$ for which $a \notin g(a)$. Then the assumption that $N = g(b)$ for some $b \in A$ leads to contradiction: $b \in N$ implies $b \notin N$ and $b \notin N$ implies $b \in N$.

Corollary. *There exists uncountable sets, e.g. the set $\mathcal{P}(\mathbb{N})$ or the set $Map\,(\mathbb{N}, \{0, 1\})$ of all binary sequences.*

The statement

$$\text{Card } \mathcal{P}(\mathbb{N}) = \aleph_1 := \text{the minimal uncountable cardinality}$$

is the famous **Continuum Hypothesis**.

Now we discuss the notion of order in more details.

Definition 6. *A* **relation** *in the set A is a subset $R \subset A \times A$.*

A **(partial) order** *on A is a relation R possessing the properties:*

1. Reflexivity: $(x, x) \in R \; \forall x \in A$.
2. Antisymmetry: $(x, y) \in R \;\&\; (y, x) \in R \Rightarrow x = y$.
3. Transitivity: $(x, y) \in R \;\&\; (y, z) \in R \Rightarrow (x, z) \in R$.

Instead of $(x, y) \in R$ one usually says that x follows y (denoted by $x \succeq y$) or that x is greater than y (denoted by $x \geq y$). An order relation R is called **linear** if for any two elements $x, y \in A$ at least one is greater than another;

an ordered set A is called **well-ordered** if any subset B of A has a minimal element b_0 (s.t. $b \geq b_0$ $\forall b \in B$).

Example 1. *1. The set $\mathbb{N}$ of natural numbers with usual order is a well-ordered set.*

2. The set $\mathbb{Z}$ of all integers with usual order is linearly ordered (but not well-ordered).

3. The set $\mathcal{P}(S)$ of all subsets of a set S is partially ordered. Namely $A \succeq B$ means $A \supseteq B$. This order is not linear.

The fact that $\mathbb{N}$ is a well-ordered set (which is one of the axiom of the set theory) is very useful. It is equivalent to the following

Proposition 1 (Induction principle)**.** *If $S \subset \mathbb{N}$ contains 1 and if $k \in S$ implies $k+1 \in S$, then $S = \mathbb{N}$,*

which is just a reformulation of one of Peano axioms (see above).

A lot of statements P_n depending on a natural parameter n are usually proved "by induction" (i.e. one checks that the statement P_0 corresponding to $n = 0$ is true and that P_n implies P_{n+1}).

If we replace the property 2 in the definition of order relation by the

$2'$.Symmetry: $(x, y) \in R \;\Rightarrow\; (y, x) \in R$,

we get the definition of the **equivalence relation**. This notion is often used as follows. Let X be a set with equivalence relation. We construct a new set X/R whose elements are **equivalence classes** in X w.r.t. R. Namely, we say that $x \sim y$ if $(x, y) \in R$ and define the equivalence class $[x] = \{y \in X |\ x \sim y\}$.

Exercise 7. *Show that two equivalence classes coincide or are disjoint.*

4. Ordered fields

Our goal is the rigorous definition of real numbers (which in its turn is a basis for the rigorous construction of the Analysis). We start with the notion of an ordered field.

4.1. Field axioms.

Definition 7. *A* **field** *is a set F with two algebraic operations: addition and multiplication. The first is denoted by $+$, the second by $\cdot$ and both are maps $F \times F \to F$.*

1a) Addition axioms.

(i) Commutativity: $x + y = y + x$, $\forall x, y \in F$.

(ii) Associativity: $(x + y) + z = x + (y + z)$, $\forall x, y, z \in F$.

(iii) (Zero): $\exists 0 \in F$ s.t. $0 + x = x$, $\forall x \in F$.

(iv) (Additive inverse, or negative): $\forall x \in F$ $\exists -x \in F$ s.t. $x + (-x) = 0$.

1b) Multiplication axioms.
(i) Commutativity: $x \cdot y = y \cdot x, \ \forall x, y \in F$.
(ii) Associativity: $(x \cdot y) \cdot z = x \cdot (y \cdot z), \ \forall x, y, z \in F$.
(iii) (Unit): $\exists 1 \in F$ *s.t.* $1 \cdot x = x, \ \forall x \in F$.
(iv) (Multiplicative inverse): $\forall$ *non-zero* $x \in F \ \exists x^{-1} \in F$ *such that* $x \cdot x^{-1} = 1$.

1c) Distributive law (compatibility + and $\cdot$*):*
$(x + y) \cdot z = (x \cdot z) + (y \cdot z), \forall x, y, z \in F$.
1d) (Non-triviality): $1 \neq 0$.

4.2. **Ordered fields.** Now assume that our field F is a linearly ordered set by the relation $>$.[3]
2a) (Compatibility of the order with the addition):
$x > 0 \ \& \ y > 0 \Rightarrow x + y > 0$.
2b) (Compatibility of the order with the multiplication):
$x > 0 \ \& \ y > 0 \Rightarrow x \cdot y > 0$.
In this case we say that F is an **ordered field**.

Example 2. *The field* $\mathbb{Q}$ *of rational numbers is an ordered field.*

All of you already know and constantly use the properties of rational numbers. Here we sketch an accurate definition of this field. We recall that the set $\mathbf{N}$ is defined as the set of cardinalities of finite sets: $0 = \text{Card } \emptyset, \ 1 = \text{Card } \{\emptyset\}$ etc.

The addition and multiplication operations in $\mathbf{N}$ can be defined as follows. If $a = \text{Card } A, \ b = \text{Card } B$ and $A \cap B = \emptyset$, we put $a + b := \text{Card } (A \cup B), \ a \cdot b := \text{Card } (A \times B)$. One can check that both operations are commutative, associative and satisfy the distributive law. But still, $\mathbf{N}$ is not a field: the additive and multiplicative inverses do not always exist. So we have to extend the set $\mathbf{N}$ at first to the set $\mathbb{Z}$ of all integers and afterwards to the set Q of all rationals. These two extensions can be done in the similar way.

First, consider the set $\mathbb{N} \times \mathbb{N}$ and introduce the equivalence relation:

$$(m_1, \ m_2) \sim (n_1, \ n_2) \quad \text{if} \quad m_1 + n_2 = m_2 + n_1.$$

The set of equivalence classes is exactly $\mathbb{Z}$. Namely, the class $\{(m, \ m) | \ m \in \mathbb{N}\}$ represent 0, $\{(m + k, \ m) | \ m \in \mathbb{N}\}$ — the natural number k and the class $\{(m, \ m + k) | \ m \in \mathbb{N}\}$ — the negative k. As usual, we denote by $[x]$ the class of x.

We define the addition and multiplication operations on $\mathbb{Z}$ as follows:

$$[(m_1, \ m_2)] + [(n_1, \ n_2)] := [(m_1 + m_2, \ n_1 + n_2)];$$
$$[(m_1, \ m_2)] \cdot [(n_1, \ n_2)] := [(m_1 \cdot m_2 + n_1 \cdot n_2, \ m_1 \cdot n_2 + m_2 \cdot n_1)].$$

So we managed to construct the set $\mathbb{Z}$ and algebraic operations on it using only elements of $\mathbb{N}$ and operations on it.

[3] Recall that it means that for any two elements x, y of the field only three possibilities exist: $x < y$, $x = y$, $x > y$.

Now, the set $\mathbb{Q}$ is defined as the set of equivalence classes in $\mathbb{Z} \times \mathbb{N}$ w.r.t the relation

$$(m_1,\ n_1) \sim (m_2,\ n_2) \quad \text{if} \quad m_1 \cdot n_2 = m_2 \cdot n_1.$$

The algebraic operations on $\mathbb{Q}$ are defined by

$$[(m_1,\ n_1)] + [(m_2,\ n_2)] := [(m_1 \cdot n_2 + m_2 \cdot n_1,\ n_1 \cdot n_2)]$$

$$[(m_1,\ n_1)] \cdot [(m_2,\ n_2)] := [(m_1 \cdot m_2,\ n_1 \cdot n_2)].$$

Finally, the order relation in $\mathbb{Q}$ is defined by

$$[(m_1,\ n_1)] > [(m_2,\ n_2)] \quad \text{iff} \quad m_1 \cdot n_2 > m_2 \cdot n_1$$

Theorem 3. *The set $\mathbb{Q}$ with the algebraic operations $+$, $\cdot$ and the order relation $>$ defined above is an ordered field.*

The proof is the direct checking of all axioms. Usually it is left to the reader (who never does it).

Definition 8. *Two fields F_1 and F_2 are called* **isomorphic** *if there exists a bijective map $\alpha : F_1 \to F_2$ which respects all algebraic operations. Such a map is called an* **isomorphism** *between F_1 and F_2.*

Two ordered fields F_1 and F_2 are called **isomorphic** *if there exists a bijective map $\alpha : F_1 \to F_2$ which respects all algebraic operations and the order relation.*

Theorem 4. *Any ordered field F contains a subfield isomorphic to $\mathbb{Q}$.*

Proof. Define the map $\alpha : \mathbb{Q} \to F$ as follows. Put $\alpha(0) = 0_F$, $\alpha(1) = 1_F$ and extend this to the set $\mathbb{Z}$ as follows. For $n \in \mathbb{N}$ we define $\alpha(n)$ by induction: $\alpha(n+1) = \alpha(n)+1_F$ and for $n \in -\mathbb{N}$ we put $\alpha(n) = -\alpha(-n)$. Let n_F denotes the image of $n \in \mathbb{Z} \subset \mathbb{Q}$ under α. The properties $n_F + m_F = (n + m)_F$ and $n_F \cdot m_F = (n \cdot m)_F$ follow from the known properties of the addition and multiplication in $\mathbb{Z}$. I claim that the map $\alpha : n \mapsto n_F$ is injective. Indeed, assume that we have $n_F = m_F$ for $n \neq m$ and let for definiteness $n > m$. Then $k = n - m > 0$ and we can prove by induction that $k_F > 0_F$, hence $n_F > m_F$. Contradiction. □

Now, for $r = \frac{p}{q} \in \mathbb{Q}$ we define $\alpha(r) = \frac{p_F}{q_F}$. We leave as an exercise to show that this map is well defined (i.e. does not depend on the choice of the fraction $\frac{p}{q}$ representing the number $r \in \mathbb{Q}$) and injective.

Remark 1. *Usually, there is no reason to distinguish two isomorphic fields. So we can reformulate the last theorem as follows*

Theorem 5. *Any ordered field F contains $\mathbb{Q}$ as a subfield.*

4.3. **Archimedean property.** It was first introduced by Archimedes in the following form:

However big is a number M and however small is a number ε, there is a number $n \in \mathbb{N}$ such that $\underbrace{\varepsilon + \varepsilon + \cdots + \varepsilon}_{n \text{ times}} > M$.

Note that at that time only postive numbers were considered.

This property holds in $\mathbb{Q}$: if $yx^{-1} = \frac{p}{q}$, we put $n = p+1$ and get $n > p \geq \frac{p}{q} = yx^{-1}$, hence $nx > y$.

The importance of the Archimedean property one can see from the following fact.

Theorem 6. *The field* $\mathbb{Q}$ *is* **dense** *in any Archimedean ordered field* F, *i.e. for any interval* $(x, y) = \{z \in F \,|\, x < z < y\}$ *there is a rational number* $r \in (x, y)$.

Proof. : Let $n \in \mathbb{N}$ be chosen so that $n > \frac{1}{y-x}$. Then $\frac{1}{n} < (y - x)$. Among the points of the form $\frac{k}{n}$, $k \in \mathbb{Z}$, choose the maximal number $\frac{m}{n}$ which is not bigger than x. Then the number $r = \frac{m+1}{n}$ will be rational, bigger than x and not bigger than $x + \frac{1}{n} < x + (y - x) = y$. (See also the proof in the textbook). □

Remark 2. *Actually, the Archimedean property impies the density of* $\mathbb{Q}$ *in* F *in the following stronger form: any interval* $(x, y) \subset F$ *contains infinitely many rational numbers.*

Indeed, suppose that (x, y) *contains only finitely many rational numbers:* $r_1, r_2, \ldots, r_N$. *We can assume that the numeration is chosen so that* $r_1 < r_2 < \cdots < r_N$. *Then in the intervals* (x, r_1), (r_1, r_2), $\ldots$, (r_N, y) *will be no rational points at all, which contradicts the density theorem.*

Let's call the ordered field F **complete** if every monotone bounded sequence in this field converges. One can show that any complete field has the Archimedean property (see the proof in the textbook), hence contains the field $\mathbb{Q}$.

Theorem 7. *Any two complete ordered fields* F_1 *and* F_2 *are isomorphic. Moreover, the isomorphism* $\alpha : F_1 \to F_2$ *is unique and preserves the order relation.*

Proof. According to theorem $2'$ we assume that both F_1 and F_2 contain $\mathbb{Q}$ and define the map α for rational numbers as identity map: $\alpha(r) = r$. Now for any other element $x \in F_1$ we construct the increasing sequence $\{r_n\} \subset \mathbb{Q}$ which converges to x. Namely, using theorem 3, we choose a rational number $r_0 \in (x-1, x)$. And, supposing we already constructed $r_1, r_2, \cdots, r_n$, choose r_{n+1} from the intersection of two intervals: (r_n, x) and $(x - \frac{1}{2^{n+1}}, x)$. Then, clearly, our sequence possesses the properties:

1) $r_n < r_{n+1}$, i.e. $\{r_n\}$ is strictly increasing.

2) $r_0 < r_n < x < r_0 + 1$, i.e. $\{r_n\}$ is bounded.
3) $|r_n - x| < \frac{1}{2^n}$, hence $\lim\limits_{n\to\infty} r_n = x$.

Now we consider the same sequence $\{r_n\}$ as a sequence in F_2. It is monotone and bounded, hence has a limit y, and we put $\alpha(x) = y$.

We omit the check that the defined map α is an isomorphism. Note only that the preservation of the order relation follows from the preservation of algebraic operations because the set $F_+ = \{x \in F | x \geq 0\}$ coincides with algebraically defined set of all squares $\{x \in F \mid x = y^2 \text{ for some } y \in F\}$. $\square$

Now, finally we can define the field $\mathbb{R}$ of real numbers. Namely, it is the only (up to an isomorphism) complete ordered field.

Practically, we consider real numbers as certain objects which can be approximated by rational numbers with any degree of accuracy and which form a complete ordered field.

So, besides the two algebraic operations, in the set $\mathbb{R}$ we have an additional analytic operation: the limit of a monotone bounded sequence. This is the base for the construction of the whole building of **Analysis**.

5. Limits and cluster points

The standard definition of a **limit** of a sequence is:

Definition 9. *The number a is a limit of a sequence $\{x_n\}$ if for any $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that for all $n > N$ we have $|x_n - a| < \varepsilon$.*

First of all we observe that this definition make sense for sequences with elements in any ordered field F. (It is understood that ε is also an element of F.)

Exercise 8. *There are three strict inequality signs in the definition of a limit. What happens if we replace some of them by the non-strict inequality sign?*

Let's write the definition above in the set-theoretic terms. Fix an ordered field F and let S denote the set $F^{\mathbb{N}}$ of all sequences with elements in F. (Sometimes it is more convenient to consider sequences which start not with the first but with the zero-th term; then we put $S = F^{\mathbf{N}}$.)

For given $a \in F <\ n \in \mathbb{N}$ and a positive $\varepsilon \in F$ we denote by $S_{a,n,\varepsilon}$ the subset of S which consists of sequences satisfying $|x_n - a| < \varepsilon$.

Theorem 8. *Let $S(a)$ be the set of all sequences for which a is a limit. Then*

$$S(a) = \bigcap_{\varepsilon>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n>N} S_{a,n,\varepsilon}. \tag{1}$$

Proof. The property "for all $n > N$ we have $|x_n - a| < \varepsilon$" describes the subset $\bigcap_{n>N} S_{a,n,\varepsilon}$; the property "there is a $N \in \mathbb{N}$ such that for all $n > N$ we have $|x_n - a| < \varepsilon$" defines the set $\bigcup_{N\in\mathbb{N}} \bigcap_{n>N} S_{a,n,\varepsilon}$" and the whole sentence (1) corresponds to the RHS (right hand side) of (2). □

Now we formulate the **duality principle** for union and intersection:

$$\overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i}, \qquad \overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i}.$$

This means that the "bar operation" $A \mapsto \bar{A} = X \backslash A$ interchanges unions and intersections. It gives us the simple way to formulate the negation of any statement formulated in the set-theoretic terms. E.g., the set of sequences for which a **is not a limit** is written as

$$\overline{S(a)} = \bigcup_{\varepsilon>0} \bigcap_{N\in\mathbb{N}} \bigcup_{n>N} \overline{S_{a,n,\varepsilon}}. \tag{2}$$

The backward translation in the standard language is:

The number a is not a limit of a sequence $\{x_n\}$ if there is an $\varepsilon > 0$ such that for any $N \in \mathbb{N}$ there is an $n > N$ such that $|x_n - a| > \varepsilon$.

The next important notion is the **cluster point** of a sequence $\{x_n\}$. The set-theoretic formulation of the property "a is a cluster point for $\{x_n\}$" is

$$C(a) = \bigcap_{\varepsilon>0} \bigcap_{N\in\mathbb{N}} \bigcup_{n>N} S_{a,n,\varepsilon}. \tag{3}$$

Exercise 9. *Translate this definition to the ordinary language.*

To facilitate the understanding of limits and cluster points one can use the following non-official notions.[4]

a) The set A is a **trap** for a sequence $\{x_n\}$ if only finitely many members of the sequence lie outside A.

b) The set A is a **trough**(feeder) for a sequence $\{x_n\}$ if infinitely many members of the sequence lie inside A.

Exercise 10. *Show that each trap is a trough but the converse is not always true.*

Hint. *Consider the example* $x_n = (-1)^n$ *and* $A = \mathbb{N}$.

Recall that a neighborhood (or, more precisely, an ε-neighborhood) of $a \in F$ is the set of the form $\{x \in F \mid |x - a| < \varepsilon\}$ for some $\varepsilon > 0$.

---

[4] I learned these notions from known mathematician E.B.Dynkin, who used them in his lectures at Moscow State University at 1950'ties.

Exercise 11. *a) Show that* $\lim_{n\to\infty} x_n = a$ *iff (if and only if) every neighborhood of* a *is a trap for the sequence* $\{x_n\}$.

b) Show that a *is a cluster point for* $\{x_n\}$ *iff every neighborhood of* a *is a trough for this sequence.*

Remark 3. *The Exercise 3 implies that the limit of a sequence is a cluster point.*

Theorem 9. *In an Archimedean field* F *the number* a *is a cluster point for a sequence* $\{x_n\}$ *iff there exists a subsequence* $\{x_{n_k}\}$ *for which* a *is a limit.*

Proof. The "if" part is valid in any ordered field (it follows from the Remark above). Now assume that a is a cluster point for $\{x_n\}$ and show that there is a subsequence which converges to a. For any $k \in \mathbb{N}$ we can find a member x_{n_k} of our sequence such that $|x_{n_k} - a| < \frac{1}{k}$. We can also arrange that $n_{k+1} > n_k$, so that $\{x_{n_k}\}$ is a subsequence. Now we use the Archimedean property to show that our subsequence converges to a. Namely, for any $\varepsilon > 0$ there is a $m \in \mathbb{N}$ such that $m\varepsilon > 1$. But then for $k > m$ we have $|x_{n_k} - a| < \frac{1}{k} < \frac{1}{m} < \varepsilon$. □

6. Cauchy sequences, l.u.b. , lim sup, etc

Definition 10. *A sequence* $\{x_n\}$ *in an ordered field* F *is called a* **Cauchy sequence** *(or* **fundamental** *sequence) if for all* $\varepsilon > 0$ *there exists* $N \in \mathbb{N}$ *such that for all* $m, n > N$ *we have* $|x_m - x_n| < \varepsilon$.

Non-formal meaning of this definition is: Cauchy sequence is a sequence which could have a limit in some extension of the field F. This is justified by

Theorem 10. *Any convergent sequence is a Cauchy sequence.*

Proof. Let $\lim_{n\to\infty} x_n = a$. Then $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ s.t. $\forall n > N$ we have $|x_n - a| < \frac{\varepsilon}{2}$. Hence, for any $m, n > N$ we get

$$|x_m - x_n| = |(x_m - a) - (x_n - a)| \le |x_m - a| + |x_n - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Definition 11. *A set* $A \subset F$ *is called* **complete** *if every Cauchy sequence in* A *has a limit in* A.

Theorem 11. *The field* $\mathbb{R}$ *of real numbers is complete.*

We shall break the proof into 3 steps.

Lemma 3. *Every Cauchy sequence is bounded.*

Lemma 4. *Every bounded sequence in $\mathbb{R}$ has a cluster point.*

Lemma 5. *For a Cauchy sequence every cluster point is a limit.*

(Hence, a Cauchy sequence can have only one cluster point).

For proofs of Lemmas 1 and 3 we refer to the textbook. (Or better try to do it yourself using non-formal notions of trap and trough introduced in the Lecture 4).

Exercise 12. *Could a sequence have*
a) two disjoint traps? *b) two disjoint troughs?* *c) a trap and a trough which are disjoint?*

Proof of Lemma 4. We use the **bisection method** which is a standard tool in many other cases.

Since our sequence $\{x_n\}$ is bounded, we can assume that it is inside the interval $I_0 = [-M,\ M]$. Bisect this interval into two parts $[-M,\ 0]$ and $[0,\ M]$. At least one of this parts is a trough for our sequence. Denote it by I_1 and bisect it again. Continuing this procedure we get the sequence of intervals possessing the properties:

a) The sequence is **nested** in a sense that $I_{n+1} \subset I_n$.

In other words, if $I_n = [a_n,\ b_n]$, then $a_n \leq a_{n+1},\ b_n \geq b_{n+1}$.

b) The length of I_n is equal to $\frac{\text{length } I_0}{2^n} = \frac{M}{2^{n-1}}$ and goes to 0 as n goes to ∞.

c) Each interval is a trough for $\{x_n\}$, i.e. contains infinitely many members of the sequence.

We see that $\{a_n\}$ is increasing and bounded above, while $\{b_n\}$ is decreasing and bounded below.

Since $\mathbb{R}$ is complete, there exist $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$. Moreover, because $b_n - a_n \to 0$ as $n \to \infty$, we have $a = b$. It is clear that any neighborhood of a contains some interval I_n (indeed, if $n > \frac{M}{\varepsilon} + 1$, then $\varepsilon > \frac{M}{n-1} > \frac{M}{2^{n-1}} =$ length I_n and $(a-\varepsilon,\ a+\varepsilon) \supset I_n$). Hence, in any neighborhood of a there are infinitely many members of our sequence and a is a cluster point. □

Remark 4. *1. You should distinguish between a sequence (which is a map $\mathbb{N} \to \mathbb{R}$ or the set of pairs $(n,\ x_n) \in \mathbb{N} \times \mathbb{R}$) and the set of its points. The former is always infinite while the latter can consist of one element.*

2. When we consider a monotone sequence and ask if it is bounded, it is enough to find a one-side bound (upper bound for increasing sequences and lower bound for decreasing ones). The opposite bound exists automatically.

3. We shall constantly use the following properties of limits (known from the Calculus course) which show that the algebraic operations and the order relation are preserved:

$$(4)\quad \lim_{n\to\infty}(x_n \pm y_n) = \lim_{n\to\infty} x_n \pm \lim_{n\to\infty} y_n, \qquad \lim_{n\to\infty}(x_n \cdot y_n) = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n,$$

$$(5)\quad \lim_{n\to\infty}\frac{x_n}{y_n}=\frac{\lim\limits_{n\to\infty}x_n}{\lim\limits_{n\to\infty}y_n}\ \textit{if } y_n\neq 0\ \&\ \lim_{n\to\infty}y_n\neq 0,\qquad \lim_{n\to\infty}x_n\geq 0\ \textit{if } x_n\geq 0.$$

4. The set $\mathbb{R}$ of real numbers with the operation $+$ and the set $\mathbb{R}^+$ of positive real numbers with the operation $\cdot$ are isomorphic abelian groups. There are many isomorphisms between them; namely, for each positive number $a\neq 1$ the maps

$$(6)\qquad \mathbb{R}\to\mathbb{R}^+ : x\mapsto a^x \quad \textit{and} \quad \mathbb{R}^+\to\mathbb{R} : x\mapsto \log_a(x)$$

are the mutually inverse isomorphisms.

Observe, that groups $(\mathbb{Q},\ +)$ and $\mathbb{Q}^+,\ \cdot)$ **are not** *isomorphic.*

Definition 12. *A real number b is called an* **upper bound** *for a set $A\subset\mathbb{R}$ if $a\leq b\quad \forall a\in A$.*

Theorem 12. *For any non-empty set $A\subset\mathbb{R}$ the set B of all upper bounds for A either is empty or contains a minimal element $b_{\min}$ which is called the* **least upper bound** *(l.u.b. for short) for A. It is also denoted by $\sup A$ (pronounced: supremum).*

Proof. Assume that $B\neq\emptyset$. Choose elements $a\in A$ and $b\in B$. □

Exercise 13. *If $a=b$, then $b=b_{\min}$.*

So, we can suppose that $a\neq b$ and then $a<b$. Put $I_0=(a,\ b)$ and construct by the bisection method the nested sequence of intervals $\{I_n=(a_n,\ b_n)\}$ with the properties:

i) $b_n-a_n=\frac{b-a}{2^n}$;

ii) $b_n\in B$ while $a_n\notin B$.

We do this by induction. If $I_0,\ I_1,\ \cdots,I_n$ are already constructed, we consider the middle point $c_n=\frac{a_n+b_n}{2}$ of I_n and put

$$I_{n+1}=\begin{cases}(a_n,\ c_n) & \text{if}\quad c_n\in B\\ (c_n,\ b_n) & \text{if}\quad c_n\notin B\end{cases}$$

As above, we conclude that there exist $\lim\limits_{n\to\infty}a_n=\lim\limits_{n\to\infty}b_n$. Denote this common limit by $b_{\min}$ and check that it is l.u.b. for A. Indeed, $\forall a\in A$ we have $a\leq b_n$, hence $a\leq\lim\limits_{n\to\infty}b_n=b_{\min}$ and $b_{\min}\in B$. On the other side, $\forall b\in B$ we have $a_n<b$ (because a_n is not in B and therefore there is $a\in A$ s.t. $a_n<a\leq b$).

So, $b_{\min}=\lim\limits_{n\to\infty}a_n\leq b$.

The **greatest lower bound** (g.l.b. for short) is defined similarly to l.u.b. and is also denoted by $\inf A$ (pronounced: infimum).

Exercise 14. *Let $-A$ denote the set $\{x \in \mathbb{R} \mid -x \in A\}$. Show that* $\sup(-A) = -\inf A$ *and* $\inf(-A) = -\sup A$.

Definition 13. *The* **upper limit** $\limsup$ *and* **lower limit** $\liminf$ *of a bounded sequence $\{x_n\}$ are defined as follows. Let C be the set of all cluster points for a sequence $\{x_n\}$. Put*

$$\limsup_{n\to\infty} x_n = \sup C, \qquad \liminf_{n\to\infty} x_n = \inf C \tag{7}$$

If the sequence $\{x_n\}$ is unbounded from above (resp. from below), we put

$$\limsup_{n\to\infty} x_n = +\infty \quad (\text{resp.} \quad \liminf_{n\to\infty} x_n = -\infty). \tag{8}$$

The agreements (8) would be a particular case of (7) if we add to $\mathbb{R}$ two ideal points $\pm\infty$ and consider $+\infty$ (resp. $-\infty$) as a cluster point of all sequences unbounded from above (resp. from below). But we should keep in mind that the extended set $\{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$ is **not** a field.

Exercise 15. *Show that if* $\limsup\limits_{n\to\infty} x_n = \liminf\limits_{n\to\infty} x_n$*, then* $\lim\limits_{n\to\infty} x_n$ *exists.*

(In other words, if a bounded sequence has only one cluster point, then this point is a limit of the sequence.)

Definition 14. *A sequence $\{x_n\}$ has the limit $+\infty$ (notation: $\lim\limits_{n\to\infty} x_n = +\infty$) if*

$$\forall m \in \mathbb{R} \quad \exists n \in \mathbb{N} \quad s.t. \quad \forall n > N \quad \text{we have} \quad x_n > M.$$

Warning. Such a sequence is **not** a Cauchy sequence. So, we do not call it convergent.

Exercise 16. *Give the definition of a sequence which has the limit $-\infty$.*

The art of estimation. Here we collect some methods and notations which are useful for computing limits and for proofs of existence of a limit.

1. To estimate the absolute value of the difference between two numbers it is useful to consider it as a distance and use the triangle inequality.

Example 3. *Let $\lim\limits_{n\to\infty} x_n = a$ and $\lim\limits_{n\to\infty} y_n = b$. How to show that $\lim\limits_{n\to\infty}(x_n + y_n)$ exists and is equal to $a + b$?*

We have to estimate the difference between $(x_n + y_n)$ and $a + b$. Introduce the "intermediate quantity" $a + y_n$ and write:

$d(x_n{+}y_n,\ a{+}b) \le d(x_n{+}y_n,\ a{+}y_n){+}d(a{+}y_n,\ a{+}b) = d(x_n,\ a){+}d(y_n,\ b)$.

If both summands in the last sum are $< \frac{\varepsilon}{2}$, then the distance under estimation is $< \varepsilon$.

Now, we know that $\exists N_1 \in \mathbb{N}$, s.t. $|x_n - a| < \frac{\varepsilon}{2}$ and $\exists N_2 \in \mathbb{N}$, s.t. $|y_n - b| < \frac{\varepsilon}{2}$. Let $N = \max\,(N_1,\ N_2)$. For $n > N$ both inequalities above are satisfied and we have the desired estimation $d(x_n + y_n,\ a + b) < \varepsilon$.

2. Symbols O ("big Oh") and o ("little oh").

Given two sequences $\{x_n\}$ and $\{y_n\}$, we write $x_n = O(y_n)$ if there exists a positive constant C s.t.

$$|x_n| \leq C \cdot |y_n| \qquad \text{for all } n.$$

We write $x_n = o(y_n)$ if there exists a sequence $\{\varepsilon_n\}$ of positive numbers s.t.

$$\lim_{n\to\infty} \varepsilon_n = 0 \quad \text{and} \quad |x_n| \leq \varepsilon_n \cdot |y_n| \qquad \text{for all } n.$$

(When $y_n \neq 0$, we can replace this definitions by simpler ones:

$$\frac{|x_n|}{|y_n|} \leq C \qquad \text{for all } n, \qquad \text{i.e. the sequence } \left\{\frac{|x_n|}{|y_n|}\right\} \text{ is bounded.}$$

and

$$\frac{|x_n|}{|y_n|} \leq \varepsilon_n \qquad \text{for all } n, \qquad \text{i.e. the sequence } \left\{\frac{|x_n|}{|y_n|}\right\} \text{ goes to 0.}$$

In particular, $x_n = O(1)$ means that $\{x_n\}$ is bounded and $x_n = o(1)$ means that $\lim_{n\to\infty} x_n = 0$.

Lemma 1 If $x_n = O(y_n)$ and $\lim_{n\to\infty} y_n = 0$, then $\lim_{n\to\infty} x_n = 0$.

Proof. For any $\varepsilon > 0 \quad \exists N \in \mathbb{N}$ s.t. $\forall n > N$ we have $|y_n| < \frac{\varepsilon}{C}$. Then for the same n we get $|x_n| \leq C \cdot |y_n| < C \cdot \frac{\varepsilon}{C} = \varepsilon$.

Example 2. Under the hypothesis of Exercise 1 show that $\lim_{n\to\infty}(x_n \cdot y_n)$ exists and is equal to $a \cdot b$.

Here we have to estimate $d(x_n \cdot y_n,\ a \cdot b)$. Introduce the intermediate quantity $a \cdot y_n$ and write:

$d(x_n \cdot y_n,\ a \cdot b) \leq d(x_n \cdot y_n,\ a \cdot y_n) + d(a \cdot y_n,\ a \cdot b) = d(x_n,\ a) \cdot |y_n| + d(y_n,\ b) \cdot |a|$.

We know that any convergent sequence is bounded. So, we can write

$d(x_n \cdot y_n,\ a \cdot b) = O(d(x_n,\ a)) + O(d(y_n,\ b)) = o(1) + o(1) = o(1)$.

7. Euclidean n-space $\mathbb{R}^n$

By definition, the Euclidean n-space $\mathbb{R}^n$ consists of n-tuples of real numbers $\vec{x} = (x_1, \cdots, x_n)$. Two operations are defined in this space:

1. Addition of two elements x and y:

$(x_1, \cdots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_1, \cdots, x_n + y_n)$.

2. Multiplication of an element by a real number a:

$a \cdot (x_1, \cdots, x_n) = (ax_1, \cdots, ax_n)$.

The first operation satisfies the axioms of an abelian group: associativity, commutativity, existence of zero and additive inverse (see Lecture 3)

The second one satisfies the associativity axiom

i) $a \cdot (b \cdot \vec{x}) = (ab) \cdot \vec{x}$

it is related to the addition by two kinds of the distributivity axiom

ii) $a \cdot (\vec{x} + \vec{y}) = a \cdot \vec{x} + a \cdot \vec{y}, \qquad (a + b) \cdot \vec{x} = a \cdot \vec{x} + b \cdot \vec{y}$

and satisfies the multiplicative identity axiom:

iii) $1 \cdot \vec{x} = \vec{x}$.

The verification of these axioms is straitforward but the abstract notion defined by them is very important. It is the notion of a **real vector space**.

Definition 1. A real vector space is a set V (whose elements are called **vectors**) with two operations: addition of vectors and multiplication of a vector by a real number, satisfying the axioms listed above.

In fact, there is another important operation in $\mathbb{R}^n$: the **scalar multiplication** of two vectors (called also **inner**, or **dot** multiplication) defined by

$$(\vec{x},\ \vec{y}) = \sum_{k=1}^{n} x_k y_k.$$

The following properties of the scalar (inner) product are taken as the axioms for the corresponding abstract notion:

i) $(a \cdot \vec{x} + b \cdot \vec{y},\ \vec{z}) = a(\vec{x},\ \vec{z}) + b(\vec{y},\ \vec{z})$ (linearity)
ii) $(\vec{x},\ \vec{y}) = (\vec{y},\ \vec{x})$ (symmetry)
iii) $(\vec{x},\ \vec{x}) \geq 0$ and moreover $(\vec{x},\ \vec{x}) = 0$ iff $\vec{x} = \vec{0}$ (positivity).

We define the **length** or **norm** $||\vec{x}||$ of a vector $\vec{x}$ as

$$||\vec{x}|| = \sqrt{(\vec{x},\ \vec{x})}$$

and the **distance** $d(\vec{x}, \vec{y})$ between two elements $\vec{x}, \vec{y}$ as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Exercise 1. The n-dimensional unit cube is a subset $I^n \subset \mathbb{R}^n$ defined as $I_n = \{x \in \mathbb{R}^n \mid |x_k| \leq 1,\ 1 \leq k \leq n\}$. Find the diameter of I^n, i.e. the maximal distance between its points.

Theorem 13 (The Cauchy-Schwarz-(Bunyakovsky-Gauss etc...) inequality)**.** *:*

$$|(\vec{x},\ \vec{y})| \leq ||\vec{x}|| \cdot ||\vec{y}||.$$

Several proofs of it are given in the textbook and I refer you to them. The geometric reason for the inequality is the fact that in the usual 3-space the scalar product of two vectors can be defined as the product of its lengths times the cosine of the angle between these two vectors.

In general Euclidean space we can **define** the angle ϕ between $\vec{x}$ and $\vec{y}$ by

$$\cos\phi = \frac{(\vec{x},\ \vec{y})}{||\vec{x}|| \cdot ||\vec{y}||}, \quad 0 \leq \phi \leq \pi.$$

Exercise 17. *Find the angles of the triangle ABC, where $A = (1, 0, 0)$, $B = (0, 1, 0)$, $C = (0, 0, 1)$.*

I omit the definition of a complex vector space but highly recommend to test yourself by solving the exercises for $1.8 in the textbook.

Exercise 18. *Identify the Euclidean plane* $\mathbb{R}^2$ *with the complex field* $\mathbb{C}$ *by*

$$\mathbb{R}^2 \ni (x_1, x_2) \quad \longleftrightarrow \quad z = x_1 + ix_2 \in \mathbb{C}.$$

Show that $(z, w) = Re\ z\bar{w}$ *(in particular,* $||z|| = \sqrt{z\bar{z}}$*).*

Very important observation is that the definition of a limit can be formulated in terms of distances:

$$\lim_{n\to\infty} \vec{x}_n = \vec{a} \Longleftrightarrow d(\vec{x}_n,\ \vec{a}) \to 0 \text{ as } \quad n \to \infty.$$

This allows to define the limit of a sequence of vectors in any Euclidean space.

Exercise 19. *Let* $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^2$ *be arbitrary vectors. Define the sequence* $\{\vec{x}_n\}$ *of vectors by the initial condition* $\vec{x}_1 = \vec{c}$ *and recursion relations:*

$\vec{x}_{2n} = \frac{1}{2}(\vec{x}_{2n-1} + \vec{a}), \quad \vec{x}_{2n+1} = \frac{1}{2}(\vec{x}_{2n} + \vec{b}).$

a) Does $\lim_{n\to\infty} \vec{x}_n$ *exist?*

b) What about $\lim\limits_{n\to\infty} \vec{x}_{2n}$ *and* $\lim\limits_{n\to\infty} \vec{x}_{2n+1}$*?*

7.1. **Lecture 7.**

Topology of Euclidean space $\mathbb{R}^n$. To simplify the notations from now on we omit the arrows over letters which denote vectors.

We start with the observation that the distance between two points in $\mathbb{R}^n$ introduced in Lecture 6 as $d(x, y) = ||x - y||$ satisfies the same conditions as the distance on real line given by $d(x, y) = |x - y|$, namely

i) $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = y$. (Positivity)

ii) $d(x, y) = d(y, x)$. (Symmetry)

iii) $d(x, y) \leq d(x, z) + d(z, y)$. (Triangle inequality)

The two first properties are obvious from the definition of norm. The third can be derived from the Cauchy-Schwarz inequality (consult the textbook).

Definition 15. *The* ε-**neighborhood** *of a point* $a \in \mathbb{R}^n$ *is the set*

$$U_\varepsilon(a) = \{x \in \mathbb{R}^n | d(a, x) < \varepsilon\}$$

which is also called ε-**disc** *or* ε-**ball** *centered in* a*.*

Now we introduce the notions of interior, exterior and boundary points of a set A

Definition 16. *i) A point* a *is an* **interior** *point of a set* $A \in \mathbb{R}^n$ *if there exists an* ε*-neighborhood of the point which is entirely inside* A*.*

ii) A point a *is an* **exterior** *point of a set* A *if it is an interior point of the complement* $\bar{A} = \mathbb{R}^n \backslash A$ *(i.e. there exists an* ε*-neighborhood of the point which is entirely inside* $\bar{A}$*).*

iii) A point a *is a* **boundary** *point of a set* A *if any neighborhood of* a *contains points of* A *as well as points of* $\bar{A}$*.*

The set of all interior points of A is denoted by $\mathrm{int}(A)$ and the set of all boundary points by $\mathrm{bd}(A)$.

Definition 17. *i) A set A is called* **open** [5] *if it contains no boundary points (i.e. all its points are interior).*

ii) A set A is called **closed** *if it contains all its boundary points.*

Exercise 20. *Show that a set A is closed iff its complement $\bar{A}$ is open.*

Usually, the subsets of $\mathbb{R}^n$ which are given by strict inequalities are open and those which are given by equalities and non-strict inequalities are closed.

Exercise 21. *Prove that $U_\varepsilon(a)$ is open in $\mathbb{R}^n$.*

Warning. The same set considered as a subset of $\mathbb{R}^{n+1}$ is **not** open!

Theorem 14. *i) The union of any family of open sets is open.*

ii) The intersection of a finite family of open sets is open.

iii) The whole set $\mathbb{R}^n$ and the empty set $\emptyset$ are open.

Proof. a) Let $\{A_i\}_{i\in I}$ be a family of open sets and let $A = \cup_{i\in I} A_i$. Then if $a \in A$, then a belongs to some $A_i,\ i \in I$. Since A_i is open, there exists an $\varepsilon > 0$ such that $D(a,\varepsilon) \subset A_i$. But then $D(a,\varepsilon) \subset A$ and a has a neighborhood entirely in A.

b) Let $A_1, \cdots, A_N$ be open sets and $A = \cap_{k=1}^{N}$. Any point $a \in A$ belongs to all A_k and since A_k is open, there exists an $\varepsilon_k > 0$ such that $U_{\varepsilon_k}(a) \subset A_k$. Put $\varepsilon = \underset{1\le k\le N}{\rightarrow} \min \varepsilon_k$. Then $U_\varepsilon(a) \subset A$ and we are done.

c) The first statement is obvious and the second is a matter of definition.

□

Corollary. *i) The intersection of any family of closed sets is closed.*

ii) The union of a finite family of closed sets is closed.

iii) The whole set $\mathbb{R}^n$ and the empty set $\emptyset$ are closed.

The proof is a consequence of Theorem 1, Exercise 1 and the general duality principle.

Remark 5. *We do not discuss here the axiomatic definition of open sets as a family of subsets satisfying conditions i), ii), iii) of the Theorem 1. Note only that such a family is called* **topology** *and a set with a topology is called* **topological space***.*

[5] More precisely, open in $\mathbb{R}^n$. The notion of openness is a relative one: it depends on the bigger set X in which our set lies. But if this bigger set is fixed, it is traditionally not mentioned.

Exercise 22. *Show that there exists the maximal open subset in A and that it coincides with int(A).*

Hint. *Use the property i) of open sets.*

Definition 18. *A point b is called an* **accumulation** *point of a set $B \in \mathbb{R}^n$ if any neighborhood of b contains infinitely many points of B.*

(In other words, any neighborhood of b is a trough for a set B.)

Exercise 23. *Show that a point b is an accumulation point of a set $B \in \mathbb{R}^n$ iff there exists a sequence $\{b_n\}$ of* **different** *points in B such that $\lim_{n\to\infty} b_n = b$.*

Warning. An accumulation point of A not necessarily lies in A. For example, all real numbers are accumulation points of the set $\mathbb{Q}$ of rational numbers.

Theorem 15. *A set $B \in \mathbb{R}^n$ is closed iff it contains all its accumulation points.*

Proof. Suppose, $B \in \mathbb{R}^n$ is closed and $b = \lim_{n\to\infty} b_n$, $b_n \in B$, is its accumulation point. We have to show that $b \in B$. Assume the contrary: $b \in \bar{B}$. Since $\bar{B}$ is open, b possesses a neighborhood which lies in $\bar{B}$, hence contains no points from B. It contradicts to the definition of limit.

Suppose now that B contains all its accumulation points and show that B is closed, i.e. $A = \bar{B}$ is open. Take the point $a \in A$. It is not an accumulation point of B. According to Exercise 3, there is a neighborhood of b which contains no points from B, hence lies in A. □

Warning. It is not true that any set is either open or closed. Most of sets are neither open nor closed. And only two of them are open and closed simultaneously. Which ones?

Definition 19. *The smallest closed sets which contain a given set $B \in \mathbb{R}^n$ is called the* **closure** *of B and is denoted by cl(B).*

Exercise 24. *Show that cl(B) can be obtained either as a union of B and all its accumulation points or as int(B)$\cup$bd(B).*

The notion of closure is the dual to the notion of interior:

$$\overline{\mathrm{cl}(B)} = \mathrm{int}(\bar{B}); \qquad \overline{\mathrm{int}(A)} = \mathrm{cl}(\bar{A}).$$

Exercise 25. *Determine interior, exterior and boundary points of the unit ball $D(0,1) \subset \mathbb{R}^n$.*

7.2. **Connected sets.** Now we come to the definition of connected and disconnected sets.

Roughly speaking, connected sets consists of one piece while disconnected ones split into several pieces. Of course, one have to make precise the notion of a "piece" and this is done in the definition given in the textbook.

In fact, this definition looks much simpler if we introduce the notion of an open subset in any $A \subset \mathbb{R}^n$ (not only in $\mathbb{R}^n$ itself). Namely, let us define an ε-neighborhood of a in A as

$$D_A(a, \varepsilon) = \{x \in A | \ d(a, x) < \varepsilon\}$$

and call the subset $B \subset A$ **open in** A if for any point $b \in B$ there exists an $\varepsilon > 0$ such that $D_A(a, \varepsilon) \subset A$.

Definition 20. *A set A is called* **disconnected** *if it is a union of two disjoint non-empty subsets U and V which are open in A. Otherwise it is called* **connected**.

Exercise 26. *Show that the interval $(a, b) \in \mathbb{R}$ is connected but the same interval with one point c deleted is disconnected.*

Hint. *To prove the first statement, use the bisection method. To prove the second one put $U = (a, c), \ V = (c, b)$.*

Remark 6. *The property of a set $A \subset \mathbb{R}^n$ to be connected is an* **absolute** *one. It means that it does not change if we consider A as a subset of any other $\mathbb{R}^m$. E.g. an interval remains connected when considered as a subset of a plane or of 3-space.*

8. Path-connected sets

There is another mathematical notion which corresponds to our intuitive idea of connectedness.

Definition 21. *A* **continuous path** *in a set $A \subset \mathbb{R}^n$ is a continuous map $p : [0, 1] \to A$. The point $p_0 = p(0)$ is called the beginning of p and the point $p_1 = p(1)$ is called the end of p We say also that p joins the points p_0 and p_1.*

Definition 22. *A set A is called* **path-connected** *if for any two points $a\, b \in A$ there exists a continuous path p in A with $p_0 = a$ and $p_1 = b$.*

Theorem 16. *Any path-connected set is connected.*

9. Compact sets and continuous maps

9.1. Compact sets.

Definition 23. *A subset $A \subset \mathbb{R}^m$ is called* **compact** *if every sequence in A has a cluster point in A.*

Later we shall see that this property is equivalent to many other simply formulated conditions which shows that it is a natural notion. For example, a set A is compact iff any sequence $\{x_n\}$ has a convergent subsequence.

When dealing with sequences in $\mathbb{R}^m$ we shall use the notation $x_n^{(i)}$ for the i-th coordinate of the vector $\vec{x}_n \in \mathbb{R}^m$.

Lemma 6. *For a sequence $\{\vec{x}_n\}$ in $\mathbb{R}^m$ the following are equivalent:*
i) $\underset{n\to\infty}{\to} \lim \vec{x}_n = \vec{a}$.
ii) $\underset{n\to\infty}{\to} \lim x_n^{(i)} = a^{(i)}$ *for all* $i,\ 1 \le i \le m$.

Proof. Suppose that $\lim_{n\to\infty} x_n = a$. It means that $\lim_{n\to\infty} |x_n - a| = 0$. By the definition of the norm we have $|\vec{x}_n - \vec{a}| = \sqrt{\sum_{i=1}^m (x_n^{(i)} - a^{(i)})^2} \ge |x_n^{(i)} - a^{(i)}|$ for $1 \le i \le m$. Hence, $\underset{n\to\infty}{\to} \lim |x_n^{(i)} - a^{(i)}| = 0$ for $1 \le i \le m$.

Conversely, if $\underset{n\to\infty}{\to} \lim |x_n^{(i)} - a^{(i)}| = 0$ for $1 \le i \le m$, then for any $\varepsilon > 0$ we can find such an $N_i \in \mathbb{N}$ that $|x_n^{(i)} - a^{(i)}| < \frac{\varepsilon}{\sqrt{m}}$ for $n > N_i$. Put $N = \max\ (N_1,\ \cdots,\ N_m)$. Then, for $n > N$ we get $|\vec{x}_n - \vec{a}| = \sqrt{\sum_{i=1}^m (x_n^{(i)} - a^{(i)})^2} \le \sqrt{m \cdot \frac{\varepsilon^2}{m}} = \varepsilon$, i.e. $\underset{n\to\infty}{\to} \lim(\vec{x}_n - \vec{a}) = 0$. □

Definition 24. *A set $S \subset \mathbb{R}^m$ is called* **an ε-net** *for a set $A \subset \mathbb{R}^m$ if*

$$A \subset \underset{s\in S}{\to} \bigcup D(s, \varepsilon).$$

Definition 25. *A set A is called* **totally bounded** *if for any $\varepsilon > 0$ there exists a finite ε-net for A.*

Lemma 7. *For a set $A \subset \mathbb{R}^m$ the properties "bounded" and "totally bounded" are equivalent.*

Proof. If A is totally bounded, it is contained in the finite union of unit balls $D(\vec{s}_i, 1),\ 1 \le i \le n$. Let $M = \underset{1\le i\le n}{\to} \max\ |\vec{s}_i| + 1$. Then $|\vec{a}| \le M$ for all $\vec{a} \in A$.

Conversely, if A is bounded, say $|\vec{a}| \le M$, then $|a^{(i)}| \le M$ for $1 \le i \le m$. So, A is contained the m-dimensional cube with the edge of length $2M$ centered in the origin:

$$A \subset I^m(M) = \{\vec{a} \in \mathbb{R}^m | \ -M \le \ a^{(i)} \ \le M \text{ for } 1 \le i \le m\}$$

Let's divide $I^m(M)$ into N^m equal parts which are small cubes with the edges of length $\frac{2M}{N}$. They are centered in points $(M \cdot \frac{2k_1-N-1}{N}, \ \cdots, \ M \cdot \frac{2k_m-N-1}{N})$, $1 \le k_i \le N$. Every small cube is contained in a closed ball of radius $\frac{M}{N}\sqrt{m}$. (See the exercise 1 from the Lecture 6). This closed ball is contained in an open ε-ball for any $\varepsilon > \frac{M}{N}\sqrt{m}$ and the N^m centers of small cubes form a finite ε-net for A. Since the quantity $\frac{M}{N}\sqrt{m}$ goes to 0 when N goes to infinity, A is totally bounded. □

Definition 26. *A collection* $\{V_i\}_{i\in I}$ *of open sets is called an* **open covering** *for a set* A *if* $A \subset \bigcup_{i\in I} V_i$. *In this case we say that* $\{V_i\}_{i\in I}$ **covers** A. *We call* **subcovering** *any subcollection of* $\{V_i\}_{i\in I}$ *which also covers* A.

Theorem 17. *For a set* $A \subset \mathbb{R}^m$ *the following properties are equivalent:*

i) A is compact.

ii) A is closed and bounded.

iii) Any open covering of A contains a finite subcovering.

Proof. We shall prove that iii) $\Longrightarrow$ii) $\Longrightarrow$ i) $\Longrightarrow$ iii).

iii) $\Longrightarrow$ii). Let a is a boundary point of A. Consider the family of open sets $V_n = \{x \in \mathbb{R}^m | d(x,a) > \frac{1}{n}\}$, $n \in \mathbb{N}$. It is clear that $\bigcup_{n\in\mathbb{N}} V_n = \mathbb{R}^m \backslash \{a\}$. So, if $a \notin A$, the family is an open cover of A. Let $V_{n_1}, \ \cdots, \ V_{n_k}$ be a finite subcover and $n = \underset{1}{\le} i \le k \to \max \ n_i$. Then V_n alone covers A which contradicts to the fact that $a \in \text{bd}(A)$.

Consider the family $V_n = D(0,n)$, $n \in \mathbb{N}$. Since $\bigcup_{n\in\mathbb{N}} V_n = \mathbb{R}^m$, it is an open cover for A. Let $V_{n_1}, \ \cdots, \ V_{n_k}$ be a finite subcover and $n = \underset{1}{\le} i \le k \to \max \ n_i$. Then V_n alone covers A which implies A is bounded.

ii) $\Longrightarrow$ i). The proof is the modification of the Lemma 2 of Lecture 5 (one have to replace "bounded" by "totally bounded" which is possible according to Lemma 2 above).

i)$\Longrightarrow$ iii). See pp 165 -166 of the textbook. □

9.2. **Continuous maps.** Let $A \subset \mathbb{R}^m$, $b \subset \mathbb{R}^n$ and $f : A \to B$ be a map from A to B.

Definition 27. *A map $f : A \to B$ is called* **continuous at a point** $a \in A$ *if for any $\varepsilon > 0$ there is a $\delta > 0$ (which can depend on ε, f and a) such that $x \in A$ & $d(x, a) < \delta$ implies $d(f(x), f(a)) < \varepsilon$.*

A map $f : A \to B$ is called **continuous** *if it is continuous at all points of A.*

Remark 7. *1. The first condition is empty (i.e. is automatically satisfied) if a is an* **isolated** *point of A (i.e. not an accumulation point).*

2. There are several equivalent definitions of a continuous map. We mention the following two.

Theorem 18. *A map $f : A \to B$ is called* **continuous** *if it "permutable" with the limits, i.e.*

$$f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n).$$

Theorem 19. *A map $f : A \to B$ is called* **continuous** *if for any subset $V \in B$, open in B, its preimage $f^{-1}(V) = \{a \in A |\ f(a) \in V\}$ is open in A.*

9.3. **Properties of continuous maps.**

Theorem 20. *Let $f : A \to B$ and $g : B \to C$ are continuous maps. Denote by $h = g \circ f$ the composition of f and g, i.e. the map of A to C given by $h(a) = g(f(a))$. Then h is continuous.*

In short: The composition of two continuous maps is continuous.

Proof. Let us use the theorem (19). For any open subset $V \subset C$ we see that $U = g^{-1}(V)$ is open in B since g is continuous. Then $W = f^{-1}(U)$ is open in Asince f is continuous. But $h^{-1} = f^{-1} \circ g^{-1}$ and $h^{-1}(V) = W$. Hence, h is continuous. □

Theorem 21. *Let $A \subset \mathbb{R}^n$ be a compact set and $f : A \to \mathbb{R}^m$ is a continuous map. Then the image $B = f(A)$ is compact.*

Proof. Let $\{b_n\}$ be a sequence of points of B. For any b_n choose a point $a_n \in f^{-1}(b_n) \subset A$. The sequence $\{a_n\}$ has a cluster point $a \in A$. I claim that $b = f(a)$ is a cluster point of $\{b_n\}$. Indeed, let $U(b, \varepsilon) \subset B$ is a ε-neighborhood of b. Then, since f is continuous at a, there exists $\delta > 0$ such that $f(U(a, \delta)) \subset U(b, \varepsilon)$. But a is a cluster point for $\{a_n\}$, so $U(a, \delta)$ contains infinitely many members of $\{a_n\}$. The images of these members belong to $U(b, \varepsilon)$. □

Theorem 22. *Let $A \subset \mathbb{R}^n$ be a connected set and $f : A \to \mathbb{R}^m$ is a continuous map. Then the image $B = f(A)$ is connected.*

Proof. Assume the contrary: B is disconnected. Then $B = U\bigcup V$, $U\bigcap V = \emptyset$ and U, V are open in B. Put $U' = f^{-1}(U)$, $V' = f^{-1}(V)$. Then U', V' are open in A according to theorem (19). Also $U'\bigcap V' = \emptyset$ since $U\bigcap V = \emptyset$. Hence, A is disconnected. A contradiction. □

10. Differentiation and integration of functions of one variable

The goal of this lecture is to recall the basic notions and rules from Calculus and to give them rigorous definitions and proofs.

10.1. Differentiation.

We shall consider real valued functions defined on some open subset in $\mathbb{R}$.

Definition 28. *We say that a function $f : A \to \mathbb{R}$ is* **differentiable** *at a point $a \in A$ if the limit*

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

exists. This limit is denoted by $f'(a)$ and is called the **derivative of f at a**. *A function f is called* **differentiable on** *A if the derivative $f'(a)$ exists at all points of A.*

The first application of this notion is related with extremal (i.e. maximal and minimal) values of a function.

Definition 29. *We say that a function $f : A \to \mathbb{R}$ has a* **local maximum** *(resp.* **local minimum***) at a point $a \in A$ if there exists a neighborhood $D(a, \varepsilon)$ such that $f(a) \geq f(x)$ (resp. $f(a) \leq f(x)$) for all $x \in D(a, \varepsilon)$. Both local maximum and minimum are called* **local extremum**.

Theorem 23. *If f has a local extremum at a, then $f'(a) = 0$.*

Proof. Suppose f has a local maximum at a. Let $D(a, \varepsilon)$ be a neighborhood in which $f(a)$ is the maximal value of f. Then if $\varepsilon > h > 0$, then the LHS (left hand side) of (1) is non-positive because $f(a+h) \leq f(a)$. We conclude that $f'(a) \leq 0$ as a limit of a non-positive sequence. Considering now the case $-\varepsilon < h < 0$ we get $f'(a) \geq 0$. Hence, $f'(a) = 0$. □

So, the problem of finding the extremal values of a differentiable function f is reduced to the computation of the derivative f' and to the solution of the equation $f'(x) = 0$. Of course, both problems could be difficult, but for many important examples this approach is very successful.

As for calculating the derivatives, one can use the known properties of limits which imply the following rules:

$i)(cf)' = cf'$; $\qquad$ $iii)(fg)' = f'g + fg'$;

$ii) (f+g)' = f' + g'$; $\qquad iv) (\frac{f}{g})' = \frac{f'g - fg'}{g^2}$.

The important complement is the **chain rule** and its corollary for an inverse function:

$v) (f \circ g)' = (f' \circ g) \cdot g'$; $\qquad vi) (f^{-1})' = \frac{1}{f' \circ f^{-1}}$

Proof of the chain rule. :

$$(f \circ g)'(x) = \lim_{n\to\infty} \frac{f(g(x+h)) - f(g(x))}{h} =$$

$$\lim_{n\to\infty} \left(\frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \frac{g(x+h) - g(x)}{h} \right) =$$

$$\lim_{n\to\infty} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)} \cdot \lim_{n\to\infty} \frac{g(x+h) - g(x)}{h} = f'(g(x)) \cdot g'(x).$$ □

Proof of the corollary. : if $\phi = f^{-1}$, then $(f \circ \phi)(x) = x$ and $(f' \circ \phi) \cdot \phi' = 1$. Hence, $\phi'(x) = \frac{1}{f'(\phi(x))}$. □

These rules allow us to calculate the derivatives of all so called **elementary functions** if the know only the following fundamental facts.

Theorem 24. $(e^x)' = e^x$; $\qquad (\sin x)' = \cos x$; $\qquad (\log x)' = \frac{1}{x}$.

Remark 8. *The Theorem* (24) *is equivalent to the calculation of the three remarkable limits:*

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1; \qquad \lim_{h\to 0} \frac{\sin h}{h} = 1; \qquad \lim_{h\to 0} \frac{\log\,(1+h)}{h} = 1. \tag{9}$$

Remark 9. *In fact, the three statements of the Theorem* (24) *are equivalent to the single statement about the complex exponent:* $(e^z)' = e^z$*. Here the derivative of a complex-valued function* f *of a complex variable* z *is defined as before by (1) where* x *should be replaced by* z *and* h *is allowed to take complex values. The complex exponent is defined by the* **Euler identity***:*

$$e^{x+iy} = e^x(\cos y + i\,\sin y) \quad \textit{for } x, y \in \mathbb{R}$$

or by the formula

$$e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n. \tag{10}$$

We shall not prove here the theorem (24) *because we have not given the details of the rigorous definitions of the functions* e^x, $\sin x$, $\cos x$*. Note only that the most economical way to do it is to define these functions as the unique solutions to the differential equations*

$$(e^x)' = e^x, \qquad (\sin x)' = \cos x, \qquad (\cos x)' = -\sin x \tag{11}$$

with the initial conditions $e^0 = 1,\ \sin 0 = 0,\ \cos 0 = 1$.

Note also, that the relations (11) *are equivalent to the functional equations*

$$e^{x+y} = e^x \cdot e^y, \qquad \sin(x+y) = \sin x \cos y + \cos x \sin y,$$

$$\cos(x+y) = \cos x \cos y - \sin x \sin y. \tag{12}$$

Exercise 27. *Calculate the derivatives of the following functions:*

a) $x^3 - 3x$; *b)* $\sin(x^2)$; *c)* $\log(\frac{1-x}{1+x})$;

d)$\arcsin(2x)$; *e)* $\log(\tan x)$; *f)* x^x.

Second application of derivatives is the possibility to approximate any differentiable function $f(x)$ in a small neighborhood of a point x_0 by a linear function:

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + o(|x - x_0|).$$

For practical calculations the following fact is more important

Theorem 25 (Mean Value Theorem)**.** *If* $f[a,\ b] \to \mathbb{R}$ *is a continuous function which is differentiable on* $(a,\ b)$*, then there exists* $c \in (a,\ b)$ *such that*

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

We refer to the textbook for the geometric interpretation and proof of this theorem. Here we mention the important

Corollary. *If* $A \subset \mathbb{R}$ *is connected and* $f'(x) = 0$ *on* A*, then* $f(x)$ *is constant on* A *(i.e. takes the same value in all points of* A*).*

10.2. **Integration.** .

We shall consider real valued functions on closed intervals $[a,\ b] \in \mathbb{R}$.

The **partition** of such an interval is a finite collection P of points $a = x_0 < x_1 < \cdots x_{n-1} < x_n = b$ so that $[a,\ b] = \bigcup_{i=1}^n [x_{i-1},\ x_i]$. Choose a sample of representatives $c = \{c_i\}$ from each interval $\Delta_i := [x_{i-1},\ x_i]$ and also put $M_i := \sup_{x \in \Delta_i} f(x),\ m_i := \inf_{x \in \Delta_i} f(x)$.

Definition 30. *The* **integral sum** I*, the* **upper integral sum** U *and the* **lower integral sum** L *are defined by:*

$$I(f,\ P,\ c) := \sum_{i=1}^n f(c_i) \cdot |\Delta_i|, \quad U(f,\ P) := \sum_{i=1}^n M_i \cdot |\Delta_i|, \quad L(f,\ P) := \sum_{i=1}^n m_i \cdot |\Delta_i|$$

where $|\Delta_i| := x_i - x_{i-1}$ *is the length of* Δ_i.

It is clear from this definition that

$$L(f,\ P) \leq I(f,\ P,\ c) \leq U(f,\ P)$$

for any partition P and any choice of $c = \{c_i\}$.

Definition 31. *Put*

$$S(f) = \inf_P U(f,\ P), \qquad\qquad s(f) = \sup_P L(f,\ P)$$

and call $S(f)$ *(resp* $s(f)$*) the* **upper** *(resp.* **lower***)* **integral** *of the function* f *over the interval* $[a,\ b]$. *A function* f *is called* **integrable on** $[a,\ b]$ *when these two quantities coincide. Their common value is called the* **integral** *of* f *over* $[a,\ b]$ *and is denoted* $\int_a^b f(x)dx$.

The existence of lower and upper integrals and also inequality $s(f) \le S(f)$ follow from the known properties of real numbers.

Theorem 26. *Suppose that a function* f *is continuous on* $[a,\ b]$, *then*
i) f *is integrable on* $[a,\ b]$.
ii) If $F(t) := \int_a^t f(x)dx$, *then* $F' = f$ *on* $[a,\ b]$.

Definition 32. *The function* F *is called* **antiderivative** *of* f *if* $F' = f$.

The theorem shows that any continuous function has at least one antiderivative and the Corollary to the Mean Value Theorem tells that it is defined uniquely up to a constant summand.

So we get the following result on which all the practical calculations of integrals are based.

Theorem 27 (Fundamental Theorem of Calculus)**.** *For any continuous function* f *on* $[a,\ b]$ *and for any of its antiderivative* F *we have*

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

The connection between integration and differentiation given by (8) and rules i) –vi) above allow to compute explicitly a lot of integrals. (But not all of them: many elementary functions have non-elementary antiderivatives).

The proof of the theorem (26) is based on the following general fact.

Definition 33. *A function* f *on a set* $A \subset \mathbb{R}^m$ *with values in* $\mathbb{R}^n$ *is called* **uniformly continuous on** A *if for any* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that for any* $x, y \in A$ *with* $d(x, y) < \delta$ *we have* $d(f(x), f(y)) < \varepsilon$.

Theorem 28. *If a set* $A \subset \mathbb{R}^m$ *is compact and a function* $f : A \to \mathbb{R}^n$ *is continuous, then it is uniformly continuous on* A.

Proof. Assume that f is not uniformly continuous. Then $\exists\ \varepsilon > 0$ such that $\forall\ \delta > 0\ \exists x, y \in A$ with $d(x, y) < \delta$ and $d(f(x), f(y)) \ge \varepsilon$. Put $\delta = \frac{1}{k}$ and let x_k, y_k be such points of A for which $d(f(x_k), f(y_k)) \ge \varepsilon$. Since A is compact, the sequence $\{x_k\}$ contains a subsequence which converges to some point $a \in A$. Since f is continuous at a, there exists a neighborhood $D(a, r)$

such that $d(f(x), f(a)) < \frac{\varepsilon}{2}$ for all $x \in D(a,r)$. Now, there are infinitely many members x_k in $D(a, \frac{r}{2})$. Choose one of them with $k > \frac{2}{r}$. Then we have $d(a, y_k) \le d(a, x_k) + d(x_k, y_k) < \frac{r}{2} + \frac{1}{k} < r$. Hence, both x_k and y_k are in $D(a,r)$ and $d(f(x_k), f(y_k)) \le d(f(x_k), f(a)) + d(f(a), f(y_k)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. A contradiction. □

We use this theorem to estimate the difference between the upper and lower integral sums for f on $[a,\ b]$. Since $[a,\ b]$ is compact and f is continuous, it is uniformly continuous. So, for any $\varepsilon > 0$ we can find a $\delta > 0$ such that if $|\Delta_i| < \delta$, then $M_i - m_i < \varepsilon$. It follows, that there is a partition P for which we have $S(f,P) - s(f,P) < \varepsilon \cdot (b-a)$. Hence, $0 \le S(f) - s(f) < S(f,P) - s(f,P) < \varepsilon \cdot (b-a)$ for any $\varepsilon > 0$ which implies $S(f) = s(f)$.

The second statement of the Theorem (26) follows from the following estimation:

$$\min_{x\in[a,\ b]} f(x) \le \frac{1}{b-a}\int_a^b f(x)dx \le \max_{x\in[a,\ b]} f(x).$$

Exercise 28. *Calculate the following integrals:*

a) $\int_0^1 x^{2012}dx$; *b)* $\int_a^b \frac{dx}{x}$; *c)* $\int_{-1}^1 \frac{dx}{1+x^2}$;

d) $\int_0^{2\pi} \cos^2 x dx$; *e)* $\int_{\pi/2}^{3\pi/2} \frac{dx}{\sin x}$; *f)* $\int_0^\infty x^{1997}e^{-x}dx$.

10.3. **Change of variables.** Here we describe the very important property of integrals often used in practice. Note that this fact is missing in the textbook, though the more general (and much more difficult) theorem for mutiple integrals is contained in Chapter 9.

Theorem 29. *Let φ be a monotone differentiable function establishing a bijection between segments $[a,\ b]$ and $[\alpha,\ \beta]$. Let f be an integrable function on $[\alpha,\ \beta]$. Then the function $f(\varphi(x))\varphi'(x)$ is integrable on $[a,\ b]$ and*

$$\int_a^b f(s)ds = \int_\alpha^\beta f(\varphi(x))\varphi'(x)dx.$$

Proof. Let $P = \{x_k\}_{0\le k\le n}$ be a partition of the segment $[\alpha,\ \beta]$. Choose the special sample c of representatives $c_k \in [x_{k-1},\ x_k]$, so that $\varphi(x_k) - \varphi(x_{k-1}) = \varphi'(c_k)(x_k - x_{k-1})$. Then the corresponding integral sum for $f(\varphi(x))\varphi'(x)$ we have

$$I(f\circ\varphi\cdot\varphi',\ P,\ c) = \sum_{k=1}^n f(\varphi(c_k))\varphi'(c_k)(x_k - x_{k-1}) =$$
$$\sum_{k=1}^n f(\varphi(c_k))(\varphi(x_k) - \varphi(x_{k-1})) = I(f,\ \varphi(P),\ \varphi(c)).$$

When the diameter of P goes to zero, the diameter of $\varphi(P)$ aso goes to zero (since φ is uniformly continuous) and both integral sums approach to the corresponding integrals.

□

11. Uniform convergence

11.1. **The spaces $B(X)$ and $C(X)$.** There are several ways to define the convergence for a sequence of functions. The simplest way is given by

Definition 34. *Let $\{f_n\}$ be a sequence of functions $f_n : X \to M$ where X is any set and M is a metric space. We say that $\{f_n\}$ converges pointwise to a function f and write*

$$f_n(x) \to f(x) \qquad \text{(pointwise)}$$

if for any $x \in X$ we have

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad (\text{i.e.} \lim_{n\to\infty} d(f_n(x),\, f(x)) = 0).$$

Another way is to take some collection $\mathcal{F}$ of functions $f : X \to Y$ and introduce a metric in $\mathcal{F}$. (Usually $\mathcal{F}$ is a real or complex vector space and the metric is given by a norm in it: $d(f,\, g) = ||f - g||$.)

Definition 35. *We say that a sequence $\{f_n\}$ converges to f in $\mathcal{F}$ and write $f_n(x) \xrightarrow{\mathcal{F}} f(x)$ if $d(f_n,\, f) \to 0$ when $n \to \infty$.*

The important example. Let X be any set and let $\mathcal{B}(X)$ be the set of all bounded real or complex functions on X. We introduce the norm in $\mathcal{B}(X)$ by the formula

$$||f|| = \sup_{x\in X} |f(x)|.$$

The convergence in $\mathcal{B}(X)$ is called *uniform* convergence. It is a stronger property than the pointwise convergence S.

Assume now that X is a compact set and let $\mathcal{C}(X)$ denotes the space of all continuous real or complex functions on X. We know that any continuous function on a compact set is bounded, so $\mathcal{C}(X) \subset \mathcal{B}(X)$. Moreover, for every $f \in \mathcal{C}(X)$ the function $|f|$ attains its maximum value on X, so

$$||f|| = \max_{x\in X} |f(x)|.$$

Theorem 30. *The normed spaces $\mathcal{B}(X)$ and $\mathcal{C}(X)$ are complete. Moreover, $\mathcal{C}(X)$ is closed in $\mathcal{B}(X)$.*

The proof follows from theorems in the textbook.

Example 4. *Consider the sequence of real-valued functions $f_n(x) = x^{\frac{1}{2n-1}}$. Using the properties of logarithm, one can show that*

$$\lim_{n\to\infty} f_n(x) = \operatorname{sgn}(x) := \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

This function sgn *is discontinuous. It shows that f_n converges to* sgn *pointwise but not uniformly. Indeed, one can check that for any segment $[-a,\, a]$*

in the space $\mathcal{B}([-a,\, a])$ the norm $||f_n - \mathrm{sgn}|| = 1$ and does not go to zero when $n \to \infty$.

11.2. **Power series for elementary functions.** Every elementary function in a neighborhood of a non-singular point (see below) can be written in a form of convergent power series. We start from the function

$$\ln x := \int_1^x \frac{dt}{t}.$$

It is defined for all real $x > 0$. Consider a neighborhood of the point $x = 1$. Introduce the local parameter y so that $x = 1 + y$. We have $\ln(1 + y) = \int_0^y \frac{dt}{1+t}$. But for the function $\varphi(t) = \frac{1}{1+t}$ we have an expression as a power series

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \ldots = \sum_{k=0}^{\infty} (-t)^k.$$

This series is convergent on the open interval $(-1,\, 1)$ and the convergence is uniform on any segment $[-a,\, a]$, $0 < a < 1$. We can integrate this series term-by-term and get the equality

$$\ln(1 + y) = y - \frac{y^2}{2} + \frac{y^3}{3} - \ldots = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{y^{k+1}}{k+1}$$

which is valid on $(-1,\,)$. (Actually, the equality holds also for $y = 1$, but it needs a special consideration).

Let now $exp(t)$ denotes the inverse function for $\ln(x)$, so that the equalities $x = \exp(t)$ and $t = \ln(x)$ are equivalent. From the properties of $\ln(x)$ we derive the following properties of $\exp(t)$. This function is defined on the whole line $\mathbb{R}$ and gives a strictly monotone bijection of it to the half-line $\mathbb{R}_{>0}$. Moreover, we have $\exp(t + s) = \exp(t)\exp(s)$. Denote by e the real number $\exp(1)$ and compare two functions: $\exp(t)$ and e^t. We check consecutively that these two functions coincide:

a) for t=1; b) for $t \in \mathbb{N}$; c) for $t \in \mathbb{Z}$; d) for $t \in \mathbb{Q}$ and finally for $t \in \mathbb{R}$.

It follows that $\ln(x)$ coincides with the function $\log_e(x)$. Further, we can identify c^x with $e^{x\ln(c)}$ and x^c with $e^{c\ln(x)}$ so that the function of two variables $f(x,\, y) = x^y$ is defined for all positive x and all real y and satisfies the equations

$$x^{(y_1+y_2)} = x^{y_1} \cdot x^{y_1}; \qquad (x_1 \cdot x_2)^y = x_1^y \cdot x_2^y.$$

Consider now the power series

$$E(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \quad \text{where} \quad k! := 1 \cdot 2 \cdot 3 \cdots k.$$

Theorem 31. *The series $E(t)$ converges for all real t (actually, it is true also for all complex numbers) and the sum is exactly e^t.*

There are several ways to prove the theorem. We indicate two of them.

First, one can check that the series converges uniformly on any segment $[-a, a] \subset \mathbb{R}$, hence can be integrated term-by-term. It follows that

$$\int_0^x E(t)dt = \sum_{k=0}^{\infty} \int_0^x \frac{t^k}{k!} = \sum_{k=0}^{\infty} \frac{t^{k+1}}{(k+1)!}\Bigg|_0^x = E(x) - 1.$$

which gives $E'(t) = E(t)$. Then the function $E(t)e^{-t}$ has the zero derivative, hence is a constant. Since $E(0) = 1$, this constant is 1 and we are done.

Second, one can check directly that

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{k=0}^{\infty} \frac{s^k}{k!} = \sum_{k=0}^{\infty} \frac{(s+t)^k}{k!}$$

which gives $E(t)E(s) = E(s+t)$. Then we check that $E(t) = e^t$ for $t = 1$, for $t \in \mathbb{N}$, etc as above.

The function $E(z)$ of a complex argument z is defined in the same way and possesses the similar properties. In particular, $E(x+iy) = E(x)\cdot E(iy)$. Denote by $\cos(y)$ and $\sin(y)$ the real and imaginary parts of the number $E(iy)$. Then we obtain $E(x+iy) = e^x(\cos(y) + i\sin(y))$. Note, that by definition

$$\cos(y) = \sum_{k=0}^{\infty} \frac{(-1)^k y^{2k}}{(2k)!} \qquad \sin(y) = \sum_{k=0}^{\infty} \frac{(-1)^k y^{2k+1}}{(2k+1)!}.$$

It follows that $\cos'(y) = -\sin(y)$, $\sin'(y) = \cos(y)$.

Consider now in more detail the map $f : \mathbb{R} \to \mathbb{C}$ given by $f(t) = e^{it} = \cos(t) + i\sin(t)$. First we note that $\overline{f(t)} = \overline{e^{it}} = e^{-it} = f(-t)$. Therefore, $|f(t)|^2 = f(t)f(-t) = 1$. Hence, $\cos^2(t) + \sin^2(t) = 1$. We see that the point $f(t)$ is moving along the unit circle on the complex plane $\mathbb{C}$. Moreover, the velocity $f'(t) = \frac{d}{dt}e^{it} = if(t)$ is perpendicular to the radius and has absolute value 1. So, we can identify the parameter t with the length of the arc and the functions $\cos(t)$ and $\sin(t)$ with analogous functions in ordinary trigonometry.

Finally, we introduce one more power series, the so called binomial series:

$$(1+x)^\alpha = 1 + \frac{\alpha}{1}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Here the binomial coefficient $\binom{\alpha}{k}$ is defined for any number α by

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

12. Homeworks

12.1. **HW1. Due Sept 19.** The number of points is indicated for every exercise.

1. Simplify the expression:

$$(A\backslash(A \cap B)) \cup (B\backslash(B\backslash A)) .$$

(2pts)

2. Show that the following sets are countable:

a) The set $\mathbb{Z}$ of all integers. (3pts)

b) The set $\mathbb{Z}[x]$ of all polynomials in x with integer coefficients. (6pts)

3. How many different sets one can construct from 10 given subsets of a set X using the operations $\cap$, $\cup$ and $^-$ ($\bar{A} := X\backslash A$)?

(E.g. from one set A one can construct 4 sets: A itself, $\bar{A}, \emptyset = A \cap \bar{A}$ and $X = A \cup \bar{A}$). (5pts)

4. Prove by induction:

a) $1^3 + 2^3 + \cdots + n^3 = (1 + 2 + \cdots + n)^2$. (2pts)

b)The last digit of N^5 is the same as the last digit of N. (2pts)

5. Which relations are transitive:

a) being relative , b) being acquaintance, c) being ancestor? (2pts)

12.2. **HW2. Due Oct 8.** The number of points is indicated for every exercise.

1. Show that 1 is not a limit of the sequence $x_n = (-1)^n$. (5 pts)

2. The sequence $\{x_n\}$ is defined by the recurrence $x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$ and the initial condition $x_1 = 2$.

a) Show that it converges. (4 pts)

b) Find its limit. (2 pts)

3. Find all cluster points for the sequences:

a) $x_n = n$; (1pt) b)$x_n = \frac{1}{n}$; (1pt)

c) $x_n = \sin\frac{\pi n}{6}$; (2pts) d) $x_n = n$-th rational number. (5pts)

In the last problem a labelling of rational numbers by positive integers is used. (Such labellings do exist because $\mathbb{Q}$ is denumerable and we fix one of them; the answer does not depend on this choice.)

12.3. **HW3. Due Nov 6.** 1. Which of the following sets are compact and which are connected? (3 pts each)

a) $\mathbb{Z}$; b) $\mathbb{R}$; c) $[0, \infty)$; d) $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$; e) $\mathbb{C}\backslash S^1$.

2. Are the following functions continuous and are they uniformly continuous? (2 pts each)

a) $f(x) = x^2$ on $\mathbb{R}$

b) $f(x) = \frac{1}{x}$ on $[1, \infty)$.

c) $f(x) = \frac{x^{20}+12x}{x^4+x^2+1}$ on $[0, 1]$.

d) $f(x) = \cos^3 x$ on $\mathbb{R}$.

3. Show that there is no continuous surjective map $f : [0, 1] \to (0, 1)$. (2pts)

12.4. **HW4. Due Nov 20.** 1. Can we differentate the series

$$x = \sum_{k=1}^{\infty} \left(\frac{x^k}{k} - \frac{x^{k+1}}{k+1} \right), \qquad 0 \le x \le 1,$$

term by term? 4 pts

2. Evaluate the following limits: 4 pts each

a) $\lim_{x\to 0} \frac{1-\cos x}{3^x - 2^x}$

b) $\lim_{x\to 0^+} (1 + 2\sin 2x)^{1/x}$

c) $\lim_{x\to 0^+} \left(\frac{1}{\sin x} - \frac{1}{x}\right)$.

3. Test the following series for convergence ordivergence: 4 pts each

a) $\sum_{k=1}^{\infty} \frac{\sqrt{k}\log k}{k^2+2k+3}$.

b) $\sum_{k=1}^{\infty} \frac{k!\,3^k}{k^k}$.

c) $\sum_{k=1}^{\infty} \frac{(k!)^2}{(2k)!}$.

12.5. **HW5. Due Dec 4.** (5 pts for each problem)

1. Test the folowing series for absolute and conditional convergence:

a) $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^\alpha}$, α real; b) $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^\alpha + (-1)^n}$, $\alpha > 0$.

2. Find $\lim_{n\to\infty} \frac{nx}{x^2+n^2}$. Is the convergence uniform on $\mathbb{R}$?

3. Compute the domain of convergence for the series

$$\sqrt{1-x} = 1 - \frac{x}{2} - \sum_{k\ge 1} \frac{(2k-1)!!}{(2k+2)!!} x^{k+1}$$

where $(2k-1)!! = 1\cdot 3\cdot 5 \cdots (2k-1)$ and $(2k+2)!! = 2\cdot 4\cdot 6\cdots(2k+2)$.

4. Find the distances between any two members of the sequence $f_k(x) = \cos(2^k x)$ in the space $\mathcal{C}([-\pi, \pi])$. Use the result to show that the unit ball in $\mathcal{C}([-\pi, \pi])$ is not compact.

E-mail address: kirillov@math.upenn.edu