

Mathematical Analysis

HECHEN HU

November 11, 2017

Contents

1	General Mathematical Concepts and Notation	1
1.1	Mathematical Symbols and their meanings	1
1.2	Sets and Operations on Them	1
1.2.1	Naive Set Theory	1
1.2.2	ZFC: Zermelo-Fraenkel Axioms and Axiom of Choice .	2
1.2.3	The <i>Cardinality</i> of a Set(<i>Cardinal Numbers</i>)	2
1.2.4	Operations on Sets	3
1.3	Relations and Functions	4
1.3.1	Definitions of Functions	4
1.3.2	Elementary Classification of Mappings	4
1.3.3	Composition of Functions and Mutually Inverse Map- pings	5
1.3.4	Functions as Relations. The Graph of a Function . . .	6
2	The Real Numbers	9
2.1	The Axiom System and Some General Properties of the Set of Real Numbers	9
2.1.1	The Axiom System of the Real Numbers	9
2.1.2	Some General Algebraic Properties of Real Numbers .	11
2.1.3	The Completeness Axiom and the Existence of a Least Upper(or Greatest Lower) Bound of a Set of Numbers	13
2.2	The Most Important Classes of Real Numbers and Computa- tional Aspects of Operations with Real Numbers	14
2.2.1	The Natural Numbers and the Principle of Mathemat- ical Induction	14
2.2.2	Rational and Irrational Numbers	15
2.2.3	The Principle of Archimedes	17
2.2.4	Miscellaneous	18
2.3	Basic Lemmas Connected with the Completeness of the Real Numbers	20
2.3.1	The Nested Interval Lemma(Cauchy-Cantor Principle)	20
2.3.2	The Finite Covering Lemma(Borel-Lebesgue Princi- ple, or Heine-Borel Theorem)	21

2.3.3	The Limit Point Lemma(Bolzano-Weierstrass Principle)	22
2.4	Countable and Uncountable Sets	22
2.4.1	The Cardinality of the Continuum	24
2.4.2	Miscellaneous	25
3	Limits	27
3.1	The Limit of a Sequence	27
3.1.1	Definitions and Examples	27
3.1.2	Properties of the Limit of a Sequence	28
3.1.3	Questions Involving the Existence of the Limit of a Sequence	30
3.1.4	Elementary Facts about Series	35
3.2	The Limit of a Function	39
3.2.1	Definitions and Examples	39
3.2.2	Properties of the Limit of a Function	42
3.2.3	The General Definition of the Limit of a Function(Limit over a Base)	47
3.2.4	Existence of the Limit of a Function	49
4	Continuous Functions	55
4.1	Basic Definitions and Examples	55
4.1.1	Continuity of a Function at a Point	55
4.1.2	Points of Discontinuity	56
4.2	Properties of Continuous Functions	57
4.2.1	Local Properties	57
4.2.2	Global Properties of Continuous Functions	58
5	Differential Calculus	63
5.1	Differentiable Functions	63
5.1.1	Functions Differentiable at a Point	63
5.1.2	The Tangent Line; Geometric Meaning of the Deriva- tive and Differential	64
5.1.3	Examples	65
5.2	The Basic Rules of Differentiation	66
5.2.1	Differentiation and the Arithmetic Operations	66
5.2.2	Differentiation of a Composite Function(Chain Rule)	67
5.2.3	Differentiation of an Inverse Function	67
5.2.4	Higher-Order Derivatives	68
5.3	The Basic Theorems of Differential Calculus	69
5.3.1	Fermat's Lemma and Rolle's Theorem	69
5.3.2	The Theorems of Lagrange and Cauchy on Finite In- crements	69
5.3.3	Taylor's Formula	70

5.4	The Study of Functions Using the Methods of Differential Calculus	73
5.4.1	Conditions for a Function to be Monotonic	73
5.4.2	Conditions for an Interior Extremum of a Function	74
5.4.3	Conditions for a Function to be Convex	74
5.4.4	L'Hôpital's Rule	74
5.4.5	Constructing the Graph of a Function	74
5.5	Complex Numbers and Elementary Functions	74
5.5.1	Complex Numbers	74
5.5.2	Convergence in \mathbb{C} and Series with Complex Terms	74
5.5.3	Euler's Formula and the Connections Among the Elementary Functions	74
5.5.4	Power Series Representation of a Function. Analyticity	74
5.5.5	Algebraic Closedness of the Field \mathbb{C} of Complex Numbers	74
5.6	Primitives	74
5.6.1	The Primitive and the Indefinite Integral	74
5.6.2	The Basic General Methods of Finding a Primitive	74
5.6.3	Primitives of Rational Functions	74
5.6.4	Primitives of the Form $\int R(\cos x, \sin x)dx$	74
5.6.5	Primitives of the Form $\int R(x, y(x))dx$	74
6	Integration	75
6.1	Definition of the Integral and Description of the Set of Integrable Functions	76
6.1.1	Definition of the Riemann Integral	76
6.1.2	The Set of Integrable Functions	76
6.2	Linearity, Additivity and Monotonicity of the Integral	76
6.2.1	The Integral as a Linear Function on the Space $\mathcal{R}[a, b]$	76
6.2.2	The Integral as an Additive Function of the Interval of Integration	76
6.2.3	Estimation of the Integral, Monotonicity of the Integral, and the Mean-Value Theorem	76
6.3	The Integral and the Derivative	76
6.3.1	The Integral and the Primitive	76
6.3.2	The Newton-Leibniz Formula	76
6.3.3	Integration by Parts in the Definite Integral and Taylor's Formula	76
6.3.4	Change of Variable in an Integral	76
6.3.5	Examples	76
6.4	Some Applications of Integration	76
6.4.1	Additive Interval Functions and the Integral	76
6.4.2	Arc Length	76
6.4.3	The Area of a Curvilinear Trapezoid	76
6.4.4	Volume of a Solid Revolution	76

6.4.5	Work and Energy	76
6.5	Improper Integrals	76
6.5.1	Definition, Examples, and Basic Properties of Improper Integrals	76
6.5.2	Convergence of an Improper Integral	76
6.5.3	Improper Integrals with More than One Singularity	76
7	Functions of Several Variables: Their Limits and Continuity	77
7.1	The Space \mathbb{R}^m and the Most Important Classes of Its Subsets	77
7.1.1	The Set \mathbb{R}^m and the Distance in It	77
7.1.2	Open and Closed Sets in \mathbb{R}^m	78
7.1.3	Compact Sets in \mathbb{R}^m	79
7.2	Limits and Continuity of Functions of Several Variables	80
7.2.1	The Limit of a Function	80
7.2.2	Continuity of a Function of Several Variables and Properties of Continuous Functions	82
8	The Differential Calculus of Functions of Several Variables	85
8.1	The Linear Structure on \mathbb{R}^m	86
8.1.1	\mathbb{R}^m as a Vector Space	86
8.1.2	Linear Transformations $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$	86
8.1.3	The Norm in \mathbb{R}^m	86
8.1.4	The Euclidean Structure on \mathbb{R}^m	86
8.2	The Differential of a Function of Several Variables	86
8.2.1	Differentiability and the Differential of a Function at a Point	86
8.2.2	The Differential and Partial Derivatives of a Real-Valued Function	86
8.2.3	Coordinate Representation of the Differential of a Mapping. The Jacob Matrix	86
8.2.4	Continuity, Partial Derivatives, and Differentiability of a Function at a Point	86
8.3	The Basic Laws of Differentiation	86
8.3.1	Linearity of the Operation of Differentiation	86
8.3.2	Differentiation of a Composition of Mappings (Chain Rule)	86
8.3.3	Differentiation of an Inverse Mapping	86
8.4	The Basic Facts of Differential Calculus of Real-Valued Functions of Several Variables	86
8.4.1	The Mean-Value Theorem	86
8.4.2	A Sufficient Condition for Differentiability of a Function of Several Variables	86
8.4.3	Higher-Order Partial Derivatives	86
8.4.4	Taylor's Formula	86

8.4.5	Extrema of Functions of Several Variables	86
8.4.6	Some Geometric Images Connected with Functions of Several Variables	86
8.5	The Implicit Function Theorem	86
8.5.1	An Elementary Version of the Implicit Function The- orem	86
8.5.2	Transition to the Case of a Relation $F(x^1, \dots, x^m, y) =$ 0	86
8.5.3	The Implicit Function Theorem	86
8.6	Some Corollaries of the Implicit Function Theorem	86
8.6.1	The Inverse Function Theorem	86
8.6.2	Local Reduction of a Smooth Mapping to Canonical Form	86
8.6.3	Local Resolution of a Diffeomorphism into a Compo- sition of Elementary Ones	86
8.6.4	Morse's Lemma	86
8.7	Surfaces in \mathbb{R}^n and the Theory of Extrema with Constraint .	86
8.7.1	k -Dimensional Surfaces in \mathbb{R}^n	86
8.7.2	The Tangent Space	86
8.7.3	Extrema with Constraint	86

1

General Mathematical Concepts and Notation

1.1 Mathematical Symbols and their meanings

Notation	Meaning
$L \Rightarrow P$	L implies P
$L \Leftrightarrow P$	L is equivalent to P
$((L \Rightarrow P) \wedge (\neg P)) \Rightarrow (\neg L)$	If P follows from L and P is false, then L is false
$\neg((L \Leftrightarrow G) \vee (P \Leftrightarrow G))$	G is not equivalent either to L or to P
$A := B$	The Definition of A is B (equality by definition)
\square	End of proof

1.2 Sets and Operations on Them

1.2.1 Naive Set Theory

1⁰. A set may consist of any distinguishable objects($x \in A \Rightarrow \exists!x \in A$)

2⁰. A set is unambiguously determined by the collection of objects that comprise it.

3⁰. Any property defines the set of objects having that property($A = \{x|P(x)\} \Rightarrow P(A)$).

However, this will lead to Russell's Paradox:

Let's have $P(M) := M \notin M$

Consider the class $K = \{M|P(M)\}$. If so K is not a set, since whether $P(K)$ is true or false, contradiction arises.

1.2.2 ZFC: Zermelo-Fraenkel Axioms and Axiom of Choice

1⁰. **(Axiom of Extensionality)** Sets A and B are equal iff they have the same elements. $(A = B) \Leftrightarrow (\forall x((x \in A) \Leftrightarrow (x \in B)))$

2⁰. **(Axiom of Separation)** To any set A and any property P there corresponds a set B whose elements are those elements of A , and only those, having property P (if A is a set, then $B = \{x \in A | P(x)\}$ is also a set).

3⁰. **(Union Axiom)** For any set M whose elements are sets there exists a set $\bigcup M$, called the union of M and consisting of those elements and only those that belong to some element of M ($x \in \bigcup M \Leftrightarrow \exists X((X \in M) \wedge (x \in X))$)

Similarly, the intersection of the set M is defined as:

$$\bigcap M := \{x \in \bigcup M | \forall X((X \in M) \Rightarrow (x \in X))\}$$

4⁰ **(Pairing Axiom)** For any sets X and Y there exists a set Z such that X and Y are its only elements.

5⁰ **(Power Set Axiom)** For any set X there exists a set $P(X)$ having each subset of X as an element, and having no other elements.

Definition. The *successor* X^+ of the set X is $X^+ = X \cup \{X\}$.

Definition. An *inductive* set is a set that \emptyset is one of its elements and the successor of each of its elements also belongs to it.

6⁰ **(Axiom of Infinity)** There exist inductive sets (Example: \mathbb{N}_0).

7⁰ **(Axiom of Replacement)** Let $F(x, y)$ be a statement (a formula) such that for every $x_0 \in X$ there exists a unique object y_0 such that $F(x_0, y_0)$ is true. Then the objects y for which there exists an element $x \in X$ such that $F(x, y)$ is true form a set.

And finally, an axiom that is independent of ZF.

Definition. A choice function is a function f , defined on a collection X of nonempty sets, such that for every set A in X , $f(A)$ is an element of A .

8⁰ **(Axiom of Choice/Zermelo's Axiom)** For any set X of nonempty sets, there exists a choice function f defined on X . ($\forall X[\emptyset \notin X \Rightarrow \exists f : X \mapsto \bigcup X \quad \forall A \in X(f(A) \in A)]$)

1.2.3 The Cardinality of a Set (Cardinal Numbers)

Definition. The set X is said to be *equipollent* to the set Y if there exists a bijective mapping of X onto Y (then $X \sim Y$).

Definition. *Cardinality* is a measure of the number of elements of the set. If $X \sim Y$, we write $\text{card}X = \text{card}Y$.

If X is equipollent to some subset of Y , we say $\text{card}X \leq \text{card}Y$, thus

$$(\text{card}X \leq \text{card}Y) := \exists Z \subset Y (\text{card}X = \text{card}Z)$$

A set is called *finite* if it is not equipollent to any proper subset of itself; otherwise it is called *infinite*.

It has the properties below:

$$1^0 \quad (\text{card}X \leq \text{card}Y) \wedge (\text{card}Y \leq \text{card}Z) \Rightarrow (\text{card}X \leq \text{card}Z).$$

$$2^0 \quad (\text{card}X \leq \text{card}Y) \wedge (\text{card}Y \leq \text{card}X) \Rightarrow (\text{card}X = \text{card}Y) \text{ (The Schröder–Bernstein theorem).}$$

$$3^0 \quad \forall X \forall Y (\text{card}X \leq \text{card}Y) \vee (\text{card}Y \leq \text{card}X) \text{ (Cantor's theorem).}$$

We say $\text{card}X < \text{card}Y$ if $(\text{card}X \leq \text{card}Y) \wedge (\text{card}X \neq \text{card}Y)$.

let \emptyset be the empty set and $P(X)$ the set of all subsets (thus, the power set) of the set X . Then:

Theorem 1.2.1. $\text{card}X < \text{card}P(X)$

Proof. The assertion is obvious for the empty set, and we shall assume that $X \neq \emptyset$.

Since $P(X)$ contains all the one-element subsets of X , $\text{card}X \leq \text{card}P(X)$. Suppose, contrary to the assertion, that there exists a bijective mapping $f : X \rightarrow P(X)$. Let set $A = \{x \in X : x \notin f(x)\}$ consisting of the elements $x \in X$ that do not belong to the set $f(x) \in P(X)$ assigned to them by the bijection. Because $A \in P(X)$, there exists $a \in X$ such that $f(a) = A$. For the element a the relation $a \in A$ or $a \notin A$ is impossible by the definition of A (Similar to Russell's Paradox). \square

1.2.4 Operations on Sets

Notation	Meaning	Definition
$A \subset B$	A is a subset of B	$\forall x ((x \in A) \Rightarrow (x \in B))$
$A = B$	A equals to B	$(A \subset B) \wedge (B \subset A)$
\emptyset	Empty Set	$\{x x \neq x\}$
$A \cup B$	The union of A and B	$\{x x \in A \vee x \in B\}$
$A \cap B$	The intersection of A and B	$\{x x \in A \wedge x \in B\}$
$A \setminus B$	The difference between A and B	$\{x x \in A \wedge x \notin B\}$
$C_M A$	The complement of A in M	$\{x x \in M \wedge x \notin A\}$ where $A \subset M$
$A \times B$	The Cartesian Product of A and B	$\{(x, y) x \in A \wedge y \in B\}$
A^2	$A \times A$	

In the ordered pair $z = (x_1, x_2)$ where $Z = X_1 \times X_2, z \in Z, x_1 \in X_1, x_2 \in X_2$, x_1 is called the *first projection* of the pair z and denoted $\text{pr}_1 z$ while x_2 is called the *second projection* of the pair z and denoted $\text{pr}_2 z$.

1.3 Relations and Functions

1.3.1 Definitions of Functions

Definition. We say that there is a *function* defined on X with values in Y if, by virtue of some rule f , to each element $x \in X$ there corresponds an element $y \in Y$.

$$f(X) := \{y \in Y \mid \exists x((x \in X) \wedge (y = f(x)))\}$$

X is called the *domain of definition* and Y is called *set of values* or *range* of the function.

Definition. If $A \subset X$ and $f : X \rightarrow Y$ is a function. We denote by $f|_A$ the function $\varphi : A \rightarrow Y$ that agrees with f on A . More precisely, $f|_A(x) := \varphi(x)$ if $x \in A$. The function $f|_A$ is called the *restriction* of f to A , and the function $f : X \rightarrow Y$ is called an *extension* or a *continuation* of φ to X .

We use the term *domain of departure* of the function to denote any set X containing the domain of a function, and *domain of arrival* to denote any subset of Y containing its range.

1.3.2 Elementary Classification of Mappings

Definition. When a function $f : X \rightarrow Y$ is called a *mapping*, the value $f(x) \in Y$ that it assumes at the element $x \in X$ is usually called the *image* of x .

The *image* of a set $A \subset X$ under the mapping $f : X \rightarrow Y$ is defined as the set

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

consisting of the elements of Y that are images of elements of A .

The set

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}$$

consisting of the elements of X whose images belong to B is called the *pre-image* (or *complete pre-image*) of the set $B \subset Y$.

Definition. A mapping $f : X \rightarrow Y$ is said to be

surjective if $f(X) = Y$

injective if $\forall x_1, x_2 \in X, (f(x_1) = f(x_2)) \Rightarrow (x_1 = x_2)$ holds.

bijective if it's both surjective and injective.

Definition. The inverse mapping of a bijective f is denoted as

$$f^{-1} : Y \rightarrow X$$

and defined as follows: if $f(x) = y$, then $f^{-1}(y) = x$.

Note that the pre-image of a set is defined for any mapping $f : X \rightarrow Y$, even if it is not bijective and hence has no inverse.

1.3.3 Composition of Functions and Mutually Inverse Mappings

Definition. For two mappings $f : X \rightarrow Y$ and $g : Y \rightarrow Z$,

$$g \circ f : X \rightarrow Z, \quad (g \circ f)(x) := g(f(x))$$

is called the *composition* of the mapping f and the mapping g .

If all the terms of a composition $f_n \circ \dots \circ f_1$ are equal to the same function f , we abbreviate it to f^n .

Function composition is associative, that is,

$$h \circ (g \circ f) = (h \circ g) \circ f$$

But in general, $g \circ f \neq f \circ g$.

Definition. The mapping $f : X \rightarrow X$ that assigns each element in X to itself is called the *identity mapping* on X and denoted e_X .

Lemma. $(g \circ f = e_X) \Rightarrow (g \text{ is surjective}) \wedge (f \text{ is injective})$.

Proof. If $f : X \rightarrow Y$, $g : Y \rightarrow X$, and $g \circ f = e_X : X \rightarrow X$, then

$$X = e_X(X) = (g \circ f)(X) = g(f(X)) \subset g(Y)$$

and hence g is surjective.

Further, if $x_1 \in X$ and $x_2 \in X$, then

$$\begin{aligned} (x_1 \neq x_2) &\Rightarrow (e_X(x_1) \neq e_X(x_2)) \Rightarrow ((g \circ f)(x_1) \neq (g \circ f)(x_2)) \Rightarrow \\ &\Rightarrow (g(f(x_1)) \neq g(f(x_2))) \Rightarrow (f(x_1) \neq f(x_2)) \end{aligned}$$

and therefore f is injective. □

Proposition. The mappings $f : X \rightarrow Y$ and $g : Y \rightarrow X$ are bijective and mutually inverse to each other if and only if $g \circ f = e_X$ and $f \circ g = e_Y$.

1.3.4 Functions as Relations. The Graph of a Function

Relations

Definition. A Relation R is any set of ordered pairs (x, y) .

The set X is called the *domain of definition* of R , and the set Y is the *range of values*.

Any set containing the domain of definition of a relation is called a *domain of departure* for that relation, and *domain of arrival* is a set that contains the range of values of the relation.

Instead writing $(x, y) \in R$, we write xRy and say that x is *connected with* y *by the relation* R .

If $R \subset X^2$, then we say that the relation R is defined on X .

Classification of Relations

Definition. An *equivalence relation* is a relation that satisfy the following properties:

- aRa (Reflexivity);
- $aRb \Rightarrow bRa$ (Symmetry);
- $(aRb) \wedge (bRc) \Rightarrow aRc$ (Transitivity).

An equivalence relation is denoted by the special symbol \sim . $a \sim b$ means a is *equivalent* to b .

Definition. A *partial ordering* on a set X^2 is a relation R that have the following properties:

- aRa (Reflexivity);
- $(aRb) \wedge (bRc) \Rightarrow aRc$ (Transitivity).
- $(aRb) \wedge (bRa) \Rightarrow (a = b)$ (Anti-symmetry);

We often write $a \preceq b$ and say that b *follows* a . If the condition

$$\forall a \forall b ((aRb) \vee (bRa))$$

holds in addition to transitivity and anti-symmetry defining a partial ordering relation (this means any two elements of X is comparable), the relation R is called an *ordering*, and the set X is said to be *linearly ordered*.

Functions and Their Graphs

Definition. A relation R is said to be functional if

$$(xRy_1) \wedge (xRy_2) \Rightarrow (y_1 = y_2)$$

and it is called a *function*.

$R \subset X \times Y$ is a *mapping from X into Y* , or a *function from X into Y* .

Definition. The *graph* of a function $f : X \rightarrow Y$, is the subset Γ of $X \times Y$.

$$\Gamma := \{(x, y) \in X \times Y | y = f(x)\}$$

2

The Real Numbers

2.1 The Axiom System and Some General Properties of the Set of Real Numbers

2.1.1 The Axiom System of the Real Numbers

Axioms for Addition:

An operation

$$+ : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$$

is defined, assigning to each ordered pair (x, y) of elements x, y of \mathbb{R} a certain element $x + y \in \mathbb{R}$.

1₊ There exists a neutral, or identity element 0 (called zero) such that

$$x + 0 = 0 + x = x$$

for every $x \in \mathbb{R}$.

2₊ For every element $x \in \mathbb{R}$ there exists an element $-x \in \mathbb{R}$ called the *negative of x*.

$$\forall x \in \mathbb{R} \Rightarrow (\exists!(-x) \in \mathbb{R}) \wedge (x + (-x) = (-x) + x = 0)$$

3₊ The operation $+$ is associative.

$$\forall x \forall y \forall z \in \mathbb{R} \Rightarrow x + (y + z) = (x + y) + z$$

4₊ The operation $+$ is commutative.

$$\forall x \forall y \in \mathbb{R} \Rightarrow (x + y = y + x)$$

Definition. A group structure is defined on the set G – or G is a group – if Axioms 1_+ , 2_+ , and 3_+ holds for an operation defined on this set. The group is called *additive* group if the operation is called addition. When the operation is also commutative, that is, Axiom 4_+ holds, the group is also called a *commutative group* or an *Abelian group*.

According to Axioms 1_+ – 4_+ , \mathbb{R} is an additive Abelian group.

Axioms for Multiplication:

An operation

$$\bullet : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$$

is defined, assigning to each ordered pair (x, y) of elements $x, y \in \mathbb{R}$ a certain element $x \cdot y \in \mathbb{R}$, called the product of x and y .

1. There exists a neutral, or identity element $1 \in \mathbb{R} \setminus 0$ such that

$$\forall x \in \mathbb{R} \Rightarrow (x \cdot 1 = 1 \cdot x = x)$$

2. For every element $x \in \mathbb{R} \setminus 0$ there exists an element $x^{-1} \in \mathbb{R}$, called the *inverse* or *reciprocal* of x .

$$\forall x \in \mathbb{R} \Rightarrow (x \cdot x^{-1} = x^{-1} \cdot x = 1)$$

3. The operation \bullet is associative.

$$\forall x \forall y \forall z \in \mathbb{R} \Rightarrow x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

4. The operation \bullet is commutative.

$$\forall x \forall y \in \mathbb{R} \Rightarrow x \cdot y = y \cdot x$$

The set $\mathbb{R} \setminus 0$ is a *multiplicative* group.

Multiplication is distributive with respect to addition.

$$\forall x \forall y \forall z \in \mathbb{R} \Rightarrow (x + y)z = xz + yz$$

Definition. If two operations satisfying the Axioms of Addition and Multiplication are defined on a set G , then G is called a *field*.

Order Axioms

Between elements of \mathbb{R} there is a relation \leq defined, and :

- 0 \leq $\forall x \in \mathbb{R} (x \leq x)$ (Reflexivity)
- 1 \leq $(x \leq y) \wedge (y \leq x) \Rightarrow (x = y)$ (Anti-Symmetry)
- 2 \leq $(x \leq y) \wedge (y \leq z) \Rightarrow (x \leq z)$ (Transitivity)
- 3 \leq $\forall x \in \mathbb{R} \forall y \in \mathbb{R} (x \leq y) \vee (y \leq x)$

Thus, the relation \leq is an ordering, and \mathbb{R} is linearly ordered.

2.1. THE AXIOM SYSTEM AND SOME GENERAL PROPERTIES OF THE SET OF REAL NUMBERS

The Connection between Addition and Order on \mathbb{R}

$$\forall x \forall y \forall z \in \mathbb{R} \Rightarrow ((x \leq y) \Rightarrow (x + z \leq y + z))$$

The Connection between Multiplication and Order on \mathbb{R}

$$\forall x \forall y \in \mathbb{R} \Rightarrow ((0 \leq x) \wedge (0 \leq y) \Rightarrow (0 \leq x \cdot y))$$

The Axiom of Completeness(Continuity)

If X and Y are nonempty subsets of \mathbb{R} having the property that $x \leq y$ for every $x \in X$ and every $y \in Y$, then there exists $c \in \mathbb{R}$ such that $x \leq c \leq y$ for all $x \in X$ and $y \in Y$.

$$(\forall X \forall Y \subset \mathbb{R}) \wedge (X, Y \neq \emptyset) \wedge (\forall x \in X \forall y \in Y \Rightarrow x \leq y) \Rightarrow (\exists c \in \mathbb{R} \forall x \in X \forall y \in Y (x \leq c \leq y))$$

Any set on which these axioms hold can be considered a *model* of the real numbers.

Definition. An axiom system is said to be *categorical* if it determines an unique mathematical object.

Definition. If there are two models of independent number systems \mathbb{R}_A and \mathbb{R}_B that satisfying all the axioms, then a bijective correspondence can be established between these two systems, say $f : \mathbb{R}_A \rightarrow \mathbb{R}_B$, preserving the arithmetic operations and the order, that is,

$$\begin{aligned} f(x + y) &= f(x) + f(y) \\ f(x \cdot y) &= f(x) \cdot f(y) \\ x \leq y &\Leftrightarrow f(x) \leq f(y) \end{aligned}$$

and we can say that \mathbb{R}_A and \mathbb{R}_B are *isomorphic* and the mapping f is called an *isomorphism*.

Theorem 2.1.1. *The Axiom System of The Real Numbers is categorical.*

2.1.2 Some General Algebraic Properties of Real Numbers

Consequences of the Addition Axioms

- 1⁰ There is only one zero in the set of real numbers.
- 2⁰ Each element of the set of real numbers has a unique negative.
- 3⁰ In \mathbb{R} the equation

$$a + x = b$$

has the unique solution

$$x = b + (-a)$$

Consequences of the Multiplication Axioms

- 1⁰ There is only one multiplicative unit in the real numbers.
 2⁰ For each $x \neq 0$ there is only one reciprocal x^{-1} .
 3⁰ For $a \in \mathbb{R} \setminus 0$, the equation $a \cdot x = b$ has the unique solution $x = b \cdot a^{-1}$.

Consequences of the Axiom Connecting Addition and Multiplication

- 1⁰ For any $x \in \mathbb{R}$

$$x \cdot 0 = 0 \cdot x = 0$$

- 2⁰ $(x \cdot y = 0) \Rightarrow (x = 0) \vee (y = 0)$.

- 3⁰ For any $x \in \mathbb{R}$

$$-x = (-1) \cdot x$$

- 4⁰ For any $x \in \mathbb{R}$

$$(-1) \cdot (-x) = x$$

- 5⁰ For any $x \in \mathbb{R}$

$$(-x) \cdot (-x) = x \cdot x$$

Consequences of the Order Axioms

- 1⁰ For any x and y in \mathbb{R} precisely one of the following relations holds:

$$x < y, \quad x = y, \quad x > y$$

- 2⁰ For any $x, y, z \in \mathbb{R}$

$$(x < y) \wedge (y \leq z) \Rightarrow (x < z)$$

$$(x \leq y) \wedge (y < z) \Rightarrow (x < z)$$

Consequences of the Axiom Connecting Order with Addition and Multiplication

- 1⁰ For any $x, y, z, w \in \mathbb{R}$

$$(x < y) \Rightarrow (x + z) < (y + z)$$

$$(0 < x) \Rightarrow (-x < 0)$$

$$(x \leq y) \wedge (z \leq w) \Rightarrow (x + z) \leq (y + w)$$

$$(x < y) \wedge (z < w) \Rightarrow (x + z) < (y + w)$$

2.1. THE AXIOM SYSTEM AND SOME GENERAL PROPERTIES OF THE SET OF REAL NUMBERS

2⁰ If $x, y, z \in \mathbb{R}$, then

$$\begin{aligned}(0 < x) \wedge (0 < y) &\Rightarrow (0 < xy) \\ (x < 0) \wedge (y < 0) &\Rightarrow (0 < xy) \\ (x < 0) \wedge (0 < y) &\Rightarrow (xy < 0) \\ (x < y) \wedge (0 < z) &\Rightarrow (xz < yz) \\ (x < y) \wedge (z < 0) &\Rightarrow (yz < xz)\end{aligned}$$

3⁰ $0 < 1$.

4⁰ $(0 < x) \Rightarrow (0 < x^{-1})$ and $(0 < x) \wedge (x < y) \Rightarrow (0 < y^{-1}) \wedge (y^{-1} < x^{-1})$.

2.1.3 The Completeness Axiom and the Existence of a Least Upper(or Greatest Lower) Bound of a Set of Numbers

Definition. A set $X \subset \mathbb{R}$ is said to be *bounded above* (resp. *bounded below*) if there exists a number $c \in \mathbb{R}$ such that $x \leq c$ (resp. $c \leq x$) for all $x \in X$.

Definition (Maximal and Minimal Elements).

$$\begin{aligned}(a = \max X) &:= (a \in X \wedge \forall x \in X (x \leq a)) \\ (a = \min X) &:= (a \in X \wedge \forall x \in X (a \leq x))\end{aligned}$$

It follows from the order Axiom 1_< that if there is a maximal (resp. minimal) element in a set of numbers, it is the only one.

However, not every set, not even every bounded set, has a maximal or minimal element (e.g. $X = \{x \in \mathbb{R} | 0 \leq x < 1\}$).

Definition. The smallest number that bounds a set $X \subset \mathbb{R}$ from above is called the *least upper bound* (or the *exact upper bound*) of X and denoted $\sup X$ ("the supremum of X ").

$$(s = \sup X) := \forall x \in X ((x \leq s) \wedge (\forall s' < s \exists x' \in X (s' < x')))$$

Similarly, the greatest lower bound of X , $\inf X$ ("the infimum of X ") can be defined as:

$$(i = \inf X) := \forall x \in X ((i \leq x) \wedge (\forall i' > i \exists x' \in X (x' < i')))$$

Lemma. (*The least upper bound principle*) Every nonempty set of real numbers that is bounded from above has a unique least upper bound.

Proof. Since we already know that the minimal element of a set of numbers is unique (the relation \leq is Anti-Symmetric), we need only verify that the least upper bound exists.

Let $X \subset \mathbb{R}$ be a given set and $Y = \{y \in \mathbb{R} | \forall x \in X (x \leq y)\}$. We know that $X \neq \emptyset$ and $Y \neq \emptyset$. Then, by the completeness axiom there exists $c \in \mathbb{R}$ such that $\forall x \in X \forall y \in Y (x \leq c \leq y)$. Because c is greater than all the elements in X and smaller than all the elements in Y , we can see that $c \in Y$ and $c = \min Y$. $\forall c' < c$ we have $c' \in X$, and according to the completeness axiom of real numbers, there exists some $x' \in X$ such that $c' \leq x' \leq c$. Thus c is $\sup X$. \square

The existence of greatest lower bound is analogous with the existence of least upper bound, so

Lemma. $(X \text{ is nonempty and bounded below}) \Rightarrow (\exists! \inf X)$.

2.2 The Most Important Classes of Real Numbers and Computational Aspects of Operations with Real Numbers

2.2.1 The Natural Numbers and the Principle of Mathematical Induction

Definition of the Set of Natural Numbers

Definition. A set $X \subset \mathbb{R}$ is *inductive* if for each number $x \in X$, it also contains $x + 1$.

Definition. The set of *natural numbers* is the smallest inductive set (has the cardinality of \aleph_0) containing 1, that is, the intersection of all inductive sets that contain 1.

The Principle of Mathematical Induction

Definition. If a subset E of the set of natural numbers \mathbb{N} is such that $1 \in E$ and together with each number $x \in E$, the number $x + 1$ also belongs to E , then $E = \mathbb{N}$.

$$(E \subset \mathbb{N}) \wedge (1 \in E) \wedge (x \in E \Rightarrow (x + 1) \in E) \Rightarrow E = \mathbb{N}$$

Some properties of the natural numbers:

1^0 The sum and product of natural numbers are natural numbers.

2^0 $(n \in \mathbb{N}) \wedge (n \neq 1) \Rightarrow ((n - 1) \in \mathbb{N})$.

3^0 For any $n \in \mathbb{N}$ the set $\{x \in \mathbb{N} | n < x\}$ contains a minimal element, namely

$$\min\{x \in \mathbb{N} | n < x\} = n + 1$$

2.2. THE MOST IMPORTANT CLASSES OF REAL NUMBERS AND COMPUTATIONAL ASPECTS

4^0 $(n \in \mathbb{N}) \wedge (m \in \mathbb{N}) \wedge (n < m) \Rightarrow (n + 1 \leq m)$.

5^0 The number $(n+1) \in \mathbb{N}$ is the immediate successor of the number $n \in \mathbb{N}$; that is, if $n \in \mathbb{N}$, there are no natural numbers x satisfying $n - 1 < x < n$.

6^0 If $n \in \mathbb{N}$ and $n \neq 1$, then $(n - 1) \in \mathbb{N}$ and $(n - 1)$ is the immediate predecessor of $n \in \mathbb{N}$; that is, if $n \in \mathbb{N}$, there are no natural numbers x satisfying $n - 1 < x < n$.

7^0 In any nonempty subset of \mathbb{N} there is a minimal element.

Proof. Let $M \subset \mathbb{N}$.

Case 1: For $1 \in M$, We'll have $\min M = 1$, since $\forall n \in \mathbb{N} (1 \leq n)$.

Case 2: For $1 \notin M$, we find a set E such that $1 \in E = \mathbb{N} \setminus M$. If n is $\max E$, then $\forall e \in E \Rightarrow e \leq n$. However, $(n + 1) \notin E$ because $\forall n \in \mathbb{N} \Rightarrow (n + 1) > n$, and thus $(n + 1) \in M$ (If such n do not exist, then E which contains 1 is not bounded from above, and we can see that $(n \in E) \Rightarrow ((n + 1) \in E)$. By the principle of induction, $E = \mathbb{N}$. But this is impossible, since $\mathbb{N} \setminus E = M \neq \emptyset$). Therefore, $(\forall x \in M) \Rightarrow (n \not\leq x \not\leq (n + 1)) \Rightarrow (\min M = (n + 1))$. \square

2.2.2 Rational and Irrational Numbers

The Integers

Definition. The union of the set of natural numbers, the set of negatives of natural numbers, and zero is called the set of *integers* and is denoted \mathbb{Z} .

The addition and multiplication of integers do not lead outside of \mathbb{Z} . Thus, \mathbb{Z} is an Abelian group with respect to addition, but \mathbb{Z} nor $\mathbb{Z} \setminus 0$ is a group with respect to multiplication.

Theorem 2.2.1 (The fundamental theorem of arithmetic). *Each natural number admits a representation as a product*

$$n = p_1 \cdots p_k$$

where p_1, \dots, p_k are prime numbers. This representation is unique except for the order of the factors.

Corollary. 1 is not a prime number, since the product representation can contain infinite numbers of 1, which makes the representation not unique.

The Rational Numbers

Definition. Numbers of the form $m \cdot n^{-1}$, where $m, n \in \mathbb{Z}$, are called rational numbers. The set of rational numbers is denoted as \mathbb{Q} .

The Irrational Numbers

Definition. The real numbers that are not rational are called *irrational*. The set of irrational numbers is $\mathbb{R} \setminus \mathbb{Q}$.

Example 2.1. $\sqrt{2}$ is irrational.

Proof. Let X and Y be the sets of positive real numbers such that $\forall x \in X (x^2 < 2)$ and $\forall y \in Y (2 < y^2)$. $X \neq \emptyset$ and $Y \neq \emptyset$, since $1 \in X$ and $2 \in Y$.

Further, $(x < y) \Leftrightarrow (x^2 < y^2)$, and by the completeness axiom there exists $s \in \mathbb{R}$ such that $\forall x \in X \forall y \in Y (x \leq s \leq y)$. The next step is to show that $s^2 = 2$.

Case 1: $s^2 < 2$. Then we can see, for example, the number $s + \frac{2-s^2}{3s}$, which is larger than s , would have a square less than 2. Indeed, we know that $1 \in X$, $1^2 \leq s^2 < 2$, and $0 < \Delta := 2 - s^2 \leq 1$. It follows that

$$(s + \frac{\Delta}{3s})^2 = s^2 + 2 \cdot \frac{\Delta}{3s} + (\frac{\Delta}{3s})^2 < s^2 + 3s \cdot \frac{\Delta}{3s} = s^2 + \Delta = 2$$

Therefore $(s + \frac{\Delta}{3s}) \in X$, and this contradicts to the fact that $\forall x \in X (x \leq s)$.

Similarly, we can prove that $s^2 \not> 2$ by consider the number $s - \frac{\Delta}{3s}$ and $0 < \Delta := s^2 - 2 < 3$, and we have $s^2 = 2$.

Finally, now we'll show that $s \notin \mathbb{Q}$. Let's suppose the contrary that $\exists m \in \mathbb{N} \exists n \in \mathbb{N} (s = \frac{m}{n})$ and m as well as n is a prime number. Because m and n are both prime, we can see that the only common factor between them is 1. But

$$\begin{aligned} \frac{m}{n} &= \sqrt{2} \\ (\frac{m}{n})^2 &= 2 \\ m^2 &= n^2 \cdot 2 \end{aligned}$$

Hence m is even. Now let $m = 2k$, and $2k^2 = n^2$. Also, n has to be divisible by 2. This contradicts with the fact that the only common factors between m and n is 1. \square

Definition. (Algebraic Numbers and Transcendental Numbers) A real number is called *algebraic* if it is the root of an algebraic equation

$$a_0x^n + \cdots + a_{n-1}x + a_n = 0$$

with rational coefficients. Otherwise, it is called *transcendental*.

2.2. THE MOST IMPORTANT CLASSES OF REAL NUMBERS AND COMPUTATIONAL ASPECTS

2.2.3 The Principle of Archimedes

¹⁰ Any nonempty subset of natural numbers that is bounded from above contains a maximal element.

Corollary. *The set of natural numbers is not bounded above.*

²⁰ Any nonempty subset of the integers that is bounded from above (resp. from below) contains a maximal element (resp. minimal element).

³⁰ The set of integers is unbounded above and unbounded below.

Theorem 2.2.2 (The principle of Archimedes). *For any fixed positive number h and any real number x there exists a unique integer k such that $(k-1)h \leq x \leq kh$*

Proof. The set $\{n \in \mathbb{Z} \mid \frac{x}{h} < n\}$ is not empty, since we can always find an integer that is greater than $\frac{x}{h}$ for any value of x and h . This set is also bounded below and contains a minimal element k . We can see that $(k-1) \leq \frac{x}{h} < k$. These inequalities are equivalent to the principle of Archimedes because $h > 0$. The uniqueness of k can be derived from the uniqueness of the minimal element of a set of numbers. \square

Corollary. *For any positive number ε there exists a natural number n such that $0 < \frac{1}{n} < \varepsilon$.*

Proof. By the principle of Archimedes there exists $n \in \mathbb{Z}$ such that $1 < \varepsilon \cdot n$. Since $0 < 1$ and $0 < \varepsilon$, we have $0 < n$. Thus $n \in \mathbb{N}$ and $0 < \frac{1}{n} < \varepsilon$. \square

Corollary. *If the number $x \in \mathbb{R}$ is such that $0 \leq x$ and $\forall n \in \mathbb{Z} (x < \frac{1}{n})$, then $x = 0$.*

Corollary. *For any numbers $a, b \in \mathbb{R}$ such that $a < b$ there is a rational number $r \in \mathbb{Q}$ such that $a < r < b$.*

Proof. According to what we had proved, we can choose $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < (b-a)$. Then by the principle of Archimedes, there exists an integer $m \in \mathbb{Z}$ and $\frac{m-1}{n} \leq a < \frac{m}{n}$. Hence the relationship $b < \frac{m}{n}$ is impossible, since then we'll have $\frac{m-1}{n} \leq a < b \leq \frac{m}{n}$. Now we subtract each side by a , and $(b-a) \leq \frac{m}{n} - a$. Because $\frac{m-1}{n} \leq a$, it is obvious that $(b-a) \leq \frac{m}{n} - \frac{m-1}{n} \Rightarrow (b-a) \leq \frac{1}{n}$, which contradicts with the fact that $\frac{1}{n} < b-a$. We can choose $r = \frac{m}{n} \in \mathbb{Q}$ and $a < \frac{m}{n} < b$. \square

Corollary. *For any number $x \in \mathbb{R}$ there exists a unique integer $k \in \mathbb{Z}$ such that $k \leq x < k+1$.*

Proof. Replace h with 1 in the principle of Archimedes. \square

The number k just mentioned is denoted $[x]$ and is called the *integer part* of x . The quantity $\{x\} := x - [x]$ is called the *fractional part* of x . Thus $x = [x] + \{x\}$, and $\{x\} \geq 0$.

2.2.4 Miscellaneous

Theorem 2.2.3 (Triangle Inequality). $|a+b| \leq |a|+|b|$ holds for all $a, b \in \mathbb{R}$.

Proof. We know that $|x| = \max\{x, -x\}$ and $\pm x \leq |x|$. Thus

$$\begin{aligned} a+b &\leq |a|+b \leq |a|+|b| \\ -a-b &\leq |a|-b \leq |a|+|b| \\ a-b &\leq |a|-b \leq |a|+|b| \end{aligned}$$

□

By the principle of induction, we can prove the following theorem.

Theorem 2.2.4. *The inequality*

$$|x_1 + \cdots + x_n| \leq |x_1| + \cdots + |x_n|$$

holds and equality holds if $\forall n \in \mathbb{N} (x_n \leq 0) \vee \forall n \in \mathbb{N} (0 \leq x_n)$.

Definition. An open interval containing the point $x \in \mathbb{R}$ will be called a *neighborhood* of this point. The interval $(x - \delta, x + \delta)$ is the δ -*neighborhood* about x .

Estimation for errors in arithmetic operations

Definition. If x is the exact value of a quantity and \tilde{x} is a known approximation to the quantity, the numbers

$$\Delta(\tilde{x}) := |x - \tilde{x}|$$

and

$$\delta(\tilde{x}) := \frac{\Delta(\tilde{x})}{|\tilde{x}|}$$

are called respectively the *absolute* and *relative* error of approximation by \tilde{x} . The relative error is not defined when $\tilde{x} = 0$.

Proposition. If

$$|x - \tilde{x}| = \Delta(\tilde{x}), \quad |y - \tilde{y}| = \Delta(\tilde{y}),$$

then

$$\begin{aligned} \Delta(\tilde{x} + \tilde{y}) &:= |(x + y) - (\tilde{x} + \tilde{y})| \leq \Delta(\tilde{x}) + \Delta(\tilde{y}), \\ \Delta(\tilde{x} \cdot \tilde{y}) &:= |(x \cdot y) - (\tilde{x} \cdot \tilde{y})| \leq |\tilde{x}| \Delta(\tilde{y}) + |\tilde{y}| \Delta(\tilde{x}) + \Delta(\tilde{x}) \cdot \Delta(\tilde{y}), \end{aligned}$$

2.2. THE MOST IMPORTANT CLASSES OF REAL NUMBERS AND COMPUTATIONAL ASPECTS

if, in addition

$$y \neq 0, \quad \tilde{y} \neq 0, \quad \delta(\tilde{y}) = \frac{\Delta(\tilde{y})}{|\tilde{y}|} < 1$$

then

$$\Delta\left(\frac{\tilde{x}}{\tilde{y}}\right) := \left|\frac{x}{y} - \frac{\tilde{x}}{\tilde{y}}\right| \leq \frac{|\tilde{x}|\Delta(\tilde{y}) + |\tilde{y}|\Delta(\tilde{x})}{\tilde{y}^2} \cdot \frac{1}{1 - \delta(\tilde{y})}$$

Proof. Let $x = \tilde{x} + \alpha$ and $y = \tilde{y} + \beta$. Thus, $|\alpha| = \Delta(\tilde{x})$ and $|\beta| = \Delta(\tilde{y})$. Then

$$\begin{aligned} \Delta(\tilde{x} + \tilde{y}) &= |(x + y) - (\tilde{x} + \tilde{y})| = |\alpha + \beta| \leq |\alpha| + |\beta| = \Delta(\tilde{x}) + \Delta(\tilde{y}) \\ \Delta(\tilde{x} \cdot \tilde{y}) &= |(x \cdot y) - (\tilde{x} \cdot \tilde{y})| = |(\tilde{x} + \alpha)(\tilde{y} + \beta) - \tilde{x} \cdot \tilde{y}| = \\ &= |\tilde{x}\beta + \tilde{y}\alpha + \alpha\beta| \leq |\tilde{x}||\beta| + |\tilde{y}||\alpha| + |\alpha||\beta| = \\ &= |\tilde{x}|\Delta(\tilde{y}) + |\tilde{y}|\Delta(\tilde{x}) + \Delta(\tilde{x}) \cdot \Delta(\tilde{y}) \\ \Delta\left(\frac{\tilde{x}}{\tilde{y}}\right) &= \left|\frac{x}{y} - \frac{\tilde{x}}{\tilde{y}}\right| = \left|\frac{x\tilde{y} - y\tilde{x}}{y\tilde{y}}\right| = \\ &= \left|\frac{(\tilde{x} + \alpha)\tilde{y} - (\tilde{y} + \beta)\tilde{x}}{\tilde{y}^2}\right| \cdot \left|\frac{1}{1 + \beta/\tilde{y}}\right| \leq \frac{|\tilde{x}||\beta| + |\tilde{y}||\alpha|}{\tilde{y}^2} \cdot \frac{1}{1 - \delta(\tilde{y})} = \\ &= \frac{|\tilde{x}|\Delta(\tilde{y}) + |\tilde{y}|\Delta(\tilde{x})}{\tilde{y}^2} \cdot \frac{1}{1 - \delta(\tilde{y})} \end{aligned}$$

□

These statements imply that

$$\begin{aligned} \delta(\tilde{x} + \tilde{y}) &\leq \frac{\Delta(\tilde{x} + \tilde{y})}{|\tilde{x} + \tilde{y}|} \\ \delta(\tilde{x} \cdot \tilde{y}) &\leq \delta(\tilde{x}) + \delta(\tilde{y}) + \delta(\tilde{y}) \cdot \delta(\tilde{y}) \\ \delta\left(\frac{\tilde{x}}{\tilde{y}}\right) &\leq \frac{\delta(\tilde{x}) + \delta(\tilde{y})}{1 - \delta(\tilde{y})} \end{aligned}$$

The Positional Computation System

Lemma. *If a number $q > 1$ is fixed, then for every positive number $x \in \mathbb{R}$ there exists a unique integer $k \in \mathbb{Z}$ such that*

$$q^{k-1} \leq x < q^k$$

Proof. We first verify that the set of numbers of the form q^k , $k \in \mathbb{N}$ is not bounded above. Suppose the contrary, we'll have a least upper bound such that there exists some $m \in \mathbb{N}$ such that $q^m < s$. Also, we can see that $\frac{s}{q} < q^m$ (If this is not the case, then we can have $q^m \leq \frac{s}{q} \Rightarrow q^{m+1} \leq s$, which makes the biggest element in this set be q^{m+1} , and this is impossible, since q^m must be the biggest element if the least upper bound is s). Here we'll have $\frac{s}{q} < q^m \leq s \Rightarrow s < q^{m+1}$, so s could not be the least upper bound of the set.

Because $1 < q$, $\forall m, n \in \mathbb{Z} \wedge (m < n) \Rightarrow q^m < q^n$. We already show that this set is not bounded from above, so $\forall c \in \mathbb{R} \Rightarrow \exists N \in \mathbb{N}(\forall n > N(c < q^n))$ (or c will be the upper bound of this set).

Now let's set $c = \frac{1}{\varepsilon}$ and $M = N$, it follows that $\forall \varepsilon > 0(\exists M \in \mathbb{N} \forall n > M(\frac{1}{q^m} < \varepsilon))$.

The set $K \subset \mathbb{Z}$ that $K\{m | (0 < x) \wedge (x < q^m)\}$ is bounded below (because when $x = \frac{1}{\varepsilon}$, for $m < M$ we will have $q^m < x$). Therefore if the minimal element is denoted as k , for this integer x it is obvious that $q^{k-1} < x < q^k$.

Next we have to prove the uniqueness of such integer k . For $m, n \in \mathbb{Z} \Rightarrow ((m < n) \Rightarrow (m \leq n-1))$, and $q > 1 \Rightarrow (q^m \leq q^{n-1})$. Thus, if m and n are both the minimal element of the set K , it can be derived that $q^{m-1} \leq x < q^m$ and $q^{n-1} \leq x < q^n$, which imply $q^{n-1} \leq x < q^m$, are incompatible if $m \neq n$. \square

Definition. The number p satisfying the Lemma($p = k$) is called the *order of x in the base q* or (when q is fixed) simply the *order of x* .

By the principle of Archimedes, $\exists! \alpha_p(\alpha_p q^p \leq x < (\alpha_p + 1)q^p)$. We can use r_n , a sequence of numbers, can be used to approximate some real number x . Take $\alpha_p, \alpha_{p-1}, \dots, \alpha_{p-n} \dots$ from the set $\{0, 1, \dots, q-1\}$ then

$$r_n = \alpha_p q^p + \dots + \alpha_{p-n} q^{p-n}$$

and such that

$$r_n \leq x < r_n + \frac{1}{q^{n-p}}$$

The set $\{0, 1, \dots, q-1\}$ is all the digits under base q while the power of q is the order of this corresponding digit.

2.3 Basic Lemmas Connected with the Completeness of the Real Numbers

2.3.1 The Nested Interval Lemma (Cauchy-Cantor Principle)

Definition. A function $f : \mathbb{N} \rightarrow X$ of a natural-number argument is called a *sequence* or a *sequence of elements* of X . $f(n)$ corresponding to the number $n \in \mathbb{N}$ is often denoted x_n and called the n th term of the sequence.

Definition. Let $X_1, X_2, \dots, X_n \dots$ be a sequence of sets. If $X_{n+1} \subset X_n$ for all $n \in \mathbb{N}$, we say the sequence is *nested*.

Lemma (Cauchy-Cantor). *For any nested sequence $I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots$ of closed intervals, there exists a point $c \in \mathbb{R}$ belonging to all of these intervals.*

If in addition it is known that for any $\varepsilon > 0$ there is an interval I_k such that $|I_k| < \varepsilon$, then c is the unique point common to all the intervals.

Proof. Let's find two sets in this sequence, denoted as $M = [a_m, b_m]$ and $N = [a_n, b_n]$ where $m, n \in \mathbb{N}$. Suppose $n < m$, and we obtain $a_m \leq b_n$. Thus the two numerical sets $A = \{a_m\}$ and $B = \{b_n\}$ satisfy the axiom of completeness. So $\exists c \in \mathbb{R} (a_m \leq c \leq b_n)$. This also means that $\forall n \in \mathbb{N} (a_n \leq c \leq b_n)$. Therefore this point c belongs to all the intervals.

Now let c_1 and c_2 be two points having this property. Without loss of generality, let $c_1 < c_2$, then $\forall n \in \mathbb{N} (a_n \leq c_1 < c_2 \leq b_n) \Rightarrow (0 < c_2 - c_1 < b_n - a_n)$, and the length of any interval in this sequence cannot be less than $c_2 - c_1$. Hence if there are intervals of arbitrarily small length in the sequence, their common point is unique. \square

2.3.2 The Finite Covering Lemma (Borel-Lebesgue Principle, or Heine-Borel Theorem)

Definition. A system $S = \{X\}$ of sets X is said to *cover* a set Y if $Y \subset \bigcup_{X \in S} X$.

A subset of S that is also a system of sets will be called a *subsystem* of S .

Lemma (Borel-Lebesgue). *Every system of open intervals covering a closed interval contains a finite subsystem that covers the closed interval.*

Proof. Let $S = \{U\}$ be a system of open intervals U that cover the closed interval $[a, b] = I_1$. If the interval I_1 could not be covered by a finite set of intervals of the system S , then, dividing I_1 into two halves, we could find that at least one of the two halves, which we denoted by I_2 , does not admit a finite covering. We now repeat this procedure with the interval I_2 , and so on.

In this way a nested sequence I_1

supset I_2

supset \cdots

supset I_n

supset \cdots of closed intervals arises, none of which admit a covering by a finite subsystem of S . Since the length of the interval $|I_n| = |I_1| \cdot 2^{-n}$, the sequence $\{I_n\}$ contains intervals of arbitrarily small length (The second part of **Cauchy-Cantor Principle**), and thus there exists a unique point c belonging to all the intervals $I_n, n \in \mathbb{N}$. Since $c \in I_1 = [a, b]$, there exists an

open interval $(\alpha, \beta) = U \in S$ containing c . Let $\varepsilon = \min\{c - \alpha, \beta - c\}$. In the sequence just constructed, we find an interval I_n such that $|I_n| < \varepsilon$. Since $c \in I_n$ and $|I_n| < \varepsilon$, we conclude that $I_n \subset (\alpha, \beta)$. But this contradicts the fact that the interval I_n cannot be covered by a finite set of intervals from the system. \square

2.3.3 The Limit Point Lemma(Bolzano-Weierstrass Principle)

Definition. A point $p \in \mathbb{R}$ is a *limit point* of the set $X \subset \mathbb{R}$ if every neighborhood of the point contains an infinite subset of X .

Examples:

If $X = \{\frac{1}{n} \in \mathbb{R} | n \in \mathbb{N}\}$, the only limit point of X is the point $0 \in \mathbb{R}$.

For an open interval (a, b) every point of the closed interval $[a, b]$ is a limit point, and there are no others.

For the set \mathbb{Q} every point of \mathbb{R} is a limit point; for, as we know, every open interval of the real numbers contains rational numbers.

Lemma (Bolzano-Weierstrass). *Every bounded infinite set of real numbers has at least one limit point.*

Proof. Let X be the given subset of \mathbb{R} . It follows from the definition of boundedness that X is contained in some closed interval $I \subset \mathbb{R}$. The next step is to show that at least one point of I is a limit point of X .

If, suppose the contrary, each point $x \in I$ would have a neighborhood $U(x)$ containing either no points of X or at most a finite number. The totality of such neighborhoods $\{U(x)\}$ constructed for the points $x \in I$ forms a covering of I by open intervals $U(x)$. By the finite covering Lemma(Borel-Lebesgue) we can extract a system $U(x_1), \dots, U(x_n)$ of open intervals that cover I . But, since $X \subset I$, this same system also covers X . However, there are only finitely many points of X in $U(x_i)$ (the definition of $U(x)$), and hence only finitely many in their union. That is, X is a finite set. This contradiction completes the proof. \square

2.4 Countable and Uncountable Sets

Definition. A set X is *countable* if it is equipollent with the set \mathbb{N} of natural numbers, that is, $\text{card}X = \text{card}\mathbb{N}$.

Proposition. An infinite subset of a countable set is countable.

Proof. Let's consider a countable set E . There is a minimal element of $E_1 := E$, which we assign to $1 \in \mathbb{N}$ and denote $e_1 \in E$. E is infinite, so $E_2 := E \setminus e_1$ is not empty. Following the principle of induction, we can

construct an injective mapping from $\{1, 2, \dots\}$ to $\{e_1, e_2, \dots\}$.

Now we have to prove that this mapping is also surjective. Suppose the contrary, that an element $e \in E$ does not have a natural number assigned to it. The set $K = \{n \in \mathbb{N} \mid n \leq e\}$ is finite, since it's a subset of \mathbb{N} bounded both from below and above. According to our previous construction, we assign 1 to $\min K$, denoted as e_1 , and we can acquire a sequence $e_1, e_2, \dots, e_{k=\text{card}K}$. But $e_{k=\text{card}K}$ is $\max K$, and because $e \in K \wedge (\forall n \in K (n \leq e))$, $e = \max K$. Therefore $e = e_k$, or otherwise it will contradict the uniqueness of maximal element. \square

Proposition. The Union of the sets of a finite or countable system of countable sets is also a countable set.

Proof. Let $X_1, X_2, \dots, X_n, \dots$ is a countable system of sets and each set $X_m = \{x_m^1, \dots, x_m^n, \dots\}$ is itself countable. Since $\forall m \in \mathbb{N} (\text{card}(X = \bigcup_{n \in \mathbb{N}} X_n) \geq \text{card}(X_m))$, X is an infinite set. The ordered pair (m, n) identifies the element $x_m^n \in X_m$. We can construct a mapping, like $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} := (m, n) \rightarrow \frac{(m+n-2)(m+n-1)}{2} + m$, such that it is bijective. Thus X is countable. Then because $\text{card}X \leq \text{card}\mathbb{N}$ and the fact that X is infinite, we conclude that $\text{card}X = \text{card}\mathbb{N}$. \square

If it is known that a set is either finite or countable, we say it is *at most countable* ($\text{card}X \leq \mathbb{N}$).

Corollary. $\text{card}\mathbb{Z} = \text{card}\mathbb{N}$

Corollary. $\text{card}\mathbb{N}^2 = \text{card}\mathbb{N}$ (The direct product of countable sets is countable).

Corollary. $\text{card}\mathbb{Q} = \text{card}\mathbb{N}$, that is, the set of rational numbers is countable.

Proof. Let (m, n) denote a rational number $\frac{m}{n}$. It is known that the pair (m, n) and (m', n') define the same number iff they are proportional. Thus \mathbb{Q} is equipollent to some infinite subset of the set $\mathbb{Z} \times \mathbb{Z}$. Since $\text{card}\mathbb{Z}^2 = \text{card}\mathbb{N}$, we can conclude that $\text{card}\mathbb{Q} = \text{card}\mathbb{N}$. \square

Corollary. The set of algebraic numbers is countable.

Proof. It can be observed that $\text{card}\mathbb{Q} \times \mathbb{Q} = \text{card}\mathbb{N}$. By the principle of induction, $\forall k \in \mathbb{N} (\text{card}\mathbb{Q}^k = \text{card}\mathbb{N})$. Let $r \in \mathbb{Q}^k$ be an ordered set (r_1, r_2, \dots, r_k) consists of k rational numbers.

An algebraic equation of degree k with rational coefficient can be written in the reduced form $x^k + r_1x^{k-1} + \dots + r_k = 0$. Thus there are as many different algebraic equations of degree k as there are different ordered sets (r_1, \dots, r_k) of rational numbers, that is, a countable set.

The algebraic equation with rational coefficients (of arbitrary degree) is the union of sets consisting of algebraic equation (of a fixed degree) which is countable, and this union is countable. Each such equation has only a finite number of roots. Hence the set of algebraic numbers is at most countable. But it is infinite, and therefore countable. \square

2.4.1 The Cardinality of the Continuum

Definition. The set \mathbb{R} of real numbers is also called the *number continuum* (from Latin *continuum*, meaning continuous, or solid), and its cardinality the *cardinality of the continuum*.

Theorem 2.4.1 (Cantor). $\text{card}\mathbb{N} < \text{card}\mathbb{R}$

Proof by Nested Interval Lemma. It is sufficient to show that even $[0, 1]$ is an uncountable set.

Assume it is countable, that is, can be written as a sequence $x_1, x_2, \dots, x_n, \dots$. Take x_1 on $I_0 = [0, 1]$, and find I_1 such that $x_1 \notin I_1$. Then construct the nested interval I_n such that $x_{n+1} \notin I_{n+1}$ and $|I_n| > 0$. It follows the nested interval lemma that there exist a point $c \in [0, 1]$ belonging to all I_n . But by our construction, $c \in \mathbb{R}$ and c cannot be any point of the sequence $x_1, x_2, \dots, x_n, \dots$. \square

Proof by Cantor's Diagonal Argument. Let's first consider an the set L and write out the infinite sequence of distinct binary numbers in it which has the form:

$$\begin{aligned} s_1 &= (0, 0, 0, 0, 0, 0, 0, \dots) \\ s_2 &= (1, 1, 1, 1, 1, 1, 1, \dots) \\ s_3 &= (0, 1, 0, 1, 0, 1, 0, \dots) \\ s_4 &= (1, 0, 1, 0, 1, 0, 1, \dots) \\ s_5 &= (1, 1, 0, 1, 0, 1, 1, \dots) \\ s_6 &= (0, 0, 1, 1, 0, 1, 1, \dots) \\ s_7 &= (1, 0, 0, 0, 1, 0, 0, \dots) \\ &\dots \end{aligned}$$

We then construct a number s such that its first digit is the complementary

(swapping 0s for 1s and vice versa) of the first digit of s_1 and etc.

$$\begin{aligned}
s_1 &= (\mathbf{0}, 0, 0, 0, 0, 0, 0, \dots) \\
s_2 &= (1, \mathbf{1}, 1, 1, 1, 1, 1, \dots) \\
s_3 &= (0, 1, \mathbf{0}, 1, 0, 1, 0, \dots) \\
s_4 &= (1, 0, 1, \mathbf{0}, 1, 0, 1, \dots) \\
s_5 &= (1, 1, 0, 1, \mathbf{0}, 1, 1, \dots) \\
s_6 &= (0, 0, 1, 1, 0, \mathbf{1}, 1, \dots) \\
s_7 &= (1, 0, 0, 0, 1, 0, \mathbf{0}, \dots) \\
&\dots \\
s &= (\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{1}, \dots)
\end{aligned}$$

By construction s differs from s_n at the n th digit, so s is not in this sequence, and thus L is uncountable.

We can now define a mapping $f : L \rightarrow \mathbb{R}$. $f(s_n) = r_n \in \mathbb{R}$ means that s_n and r_n have the same digit while r_n is under base 10 and s_n is under base 2. For $s_n \neq s_m \Rightarrow (r_n = f(s_n)) \neq (r_m = f(s_m))$, f is injective, and with the fact that all s_n corresponds to a r_n together give us $\text{card}f(L) = \text{card}L$. Since $f(L)$ is a subset of \mathbb{R} , we can see that \mathbb{R} is also uncountable. \square

The cardinality of \mathbb{R} is often denotes as \mathfrak{c} .

Corollary. $\mathbb{Q} \neq \mathbb{R}$, and so irrational numbers exist.

Corollary. There exist transcendental numbers, since the set of algebraic numbers is countable.

2.4.2 Miscellaneous

Example 2.2. The cardinality of $P(X)$, which is the power set of X , satisfy that if $\text{card}X = n$, $\text{card}P(X) = 2^n$.

Proof. We can use the principle of induction to complete the proof. If $n = 1$, $X = \{x\}$, then $P(X) = \{\emptyset, X\}$, then $\text{card}P(X) = 2^1$.

Now if $n \in \mathbb{N} \Rightarrow \text{card}P(X) = 2^n$, let X be a set that has x as one of its elements and has the cardinality of $n + 1$. Therefore $Y = X \setminus \{x\}$ has n elements. We can divide $P(X)$ into two parts: the ones containing x and the ones don't. If $x \in A \subset P(X)$, then $A \setminus \{x\} \subset P(Y)$ and vice versa. Thus we can set up a bijection between $P(Y)$ and the elements in $P(X)$ that contains x . Similarly, we can clearly see that a bijection between the subsets of $P(X)$ that does not contains x and $P(Y)$. Thus $\text{card}P(X) = 2^n + 2^n = 2^{n+1}$, and we complete the proof. \square

3

Limits

3.1 The Limit of a Sequence

3.1.1 Definitions and Examples

Definition. A number $A \in \mathbb{R}$ is called the *limit of the numerical sequence* $\{x_n\}$ if for every neighborhood $V(A)$ of A there exists an index N (depending on $V(A)$) such that all terms of the sequence having index larger than N belong to $V(A)$. $((\lim_{n \rightarrow \infty} x_n = A) := \forall V(A) \exists N \in \mathbb{N} \forall n > N (x_n \in V(A)))$

An equivalent way (or more common) way to say this is that a number $A \in \mathbb{R}$ is called the *limit of the sequence* $\{x_n\}$ if $\forall \varepsilon > 0$ there exists an index N such that $\forall n > N (|x_n - A| < \varepsilon)$. $((\lim_{n \rightarrow \infty} x_n = A) := \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n > N (|x_n - A| < \varepsilon))$

Definition. If $\lim_{n \rightarrow \infty} x_n = A$, we say that the sequence $\{x_n\}$ *converges* to A or *tends* to A and write $x_n \rightarrow A$ as $n \rightarrow \infty$. Otherwise, it's called *divergent*.

Examples:

$\lim_{n \rightarrow \infty} \frac{1}{n} = 0$, since $|\frac{1}{n} - 0| = \frac{1}{n} < \varepsilon$ when $n > N = [\frac{1}{\varepsilon}]$ (the integer part of $\frac{1}{\varepsilon}$).

$\lim_{n \rightarrow \infty} \frac{n+1}{n} = 1$, since $|\frac{n+1}{n} - 1| = \frac{1}{n} < \varepsilon$ if $n > [\frac{1}{\varepsilon}]$.

$\lim_{n \rightarrow \infty} \frac{\sin n}{n} = 0$, since $|\frac{\sin n}{n} - 0| \leq \frac{1}{n} < \varepsilon$ if $n > [\frac{1}{\varepsilon}]$.

$\lim_{n \rightarrow \infty} \frac{1}{q^n} = 0$ if $|q| > 1$.

Proof. As shown in the proof in 2.2.4(Miscellaneous), for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\frac{1}{|q|^N} < \varepsilon$. Since $|q| > 1$, we have $|\frac{1}{q^n} - 0| \leq \frac{1}{|q|^n} < \frac{1}{|q|^N} < \varepsilon$ for $n > N$. \square

The sequence $1, 2, \frac{1}{3}, 4, \frac{1}{5}, 6, \frac{1}{7}, \dots$ whose n th term is $x_n = n^{(-1)^n}$, $n \in \mathbb{N}$, is divergent.

Proof. If A is the limit of this sequence, then any neighborhood of A would contain all but a finite number of terms of the sequence.

A number $A \neq 0$ cannot be the limit, since when $\varepsilon = \frac{|A|}{2} > 0$, any point of the form $\frac{1}{2k+1}$ for which $\frac{1}{2k+1} < \frac{|A|}{2}$ lie outside the ε -neighborhood of A . At the same time $A = 0$ cannot be the limit of this sequence because there are infinitely many terms lie outside of even 1-neighborhood of 0. \square

3.1.2 Properties of the Limit of a Sequence

Definition. If there exists a number A and an index N such that $x_n = A$ for all $n > N$, the sequence $\{x_n\}$ will be called *ultimately constant*.

Definition. A sequence $\{x_n\}$ is *bounded* if there exists M such that $|x_n| < M$ for all $n \in \mathbb{N}$.

Theorem 3.1.1. *An ultimately constant sequence converges.*

Proof. if $x_N = A$, then $\forall n > N (x_n \in V(A))$. \square

Theorem 3.1.2. *Any neighborhood of the limit of a sequence contains all but a finite number of terms of the sequence.*

Theorem 3.1.3. *A convergent sequence cannot have two different limits.*

Proof. Suppose the contrary, that is, A_1 and A_2 are both the limit of the sequence x_n . Then we find two nonintersecting neighborhoods $V(A_1)$ and $V(A_2)$ of A_1 and A_2 . By the definition of limits we find two indices N_1 and N_2 such that $\forall n > N_1 (x_n \in V(A_1))$ and $\forall n > N_2 (x_n \in V(A_2))$. But then for $N = \max\{N_1, N_2\}$, we'll have $\forall n > N (x_n \in V(A_1) \cap V(A_2))$, and this is impossible since $V(A_1) \cap V(A_2) = \emptyset$. \square

Theorem 3.1.4. *A convergent sequence is bounded.*

Proof. Let $\lim_{n \rightarrow \infty} x_n = A$. Set $\varepsilon = 1$ in the common definition of limit, we find N such that $|x_n - A| < 1$ for all $n > N$. Then by the triangle inequality we have $|x_n| < |A| + 1$. Considering $n < N$, we take $M > \max\{|x_1|, |x_2|, \dots, |x_N|, |A| + 1\}$, and for all $n \in \mathbb{N}$ we have $M > |x_n|$. \square

Passage to the Limit and the Arithmetic Operations

Definition. If $\{x_n\}$ and $\{y_n\}$ are two numerical sequences, their *sum*, *product*, and *quotient* are the sequences

$$\{(x_n + y_n)\}, \quad \{(x_n \cdot y_n)\}, \quad \left\{\left(\frac{x_n}{y_n}\right)\right\}$$

while the quotient is defined when $\forall n \in \mathbb{N} (y_n \neq 0)$.

Let $\{x_n\}$ and $\{y_n\}$ be two numerical sequences, if $\lim_{n \rightarrow \infty} x_n = A$ and $\lim_{n \rightarrow \infty} y_n = B$, then

Theorem 3.1.5. $\lim_{n \rightarrow \infty} \{x_n + y_n\} = A + B$.

Proof. Set $|A - x_n| = \Delta(x_n)$, $|B - y_n| = \Delta(y_n)$. Now we have

$$|(A + B) - (x_n + y_n)| \leq \Delta(x_n) + \Delta(y_n)$$

Suppose $\varepsilon > 0$ is given. Since $\lim_{n \rightarrow \infty} x_n = A$, there exists N' such that $\Delta(x_n) < \varepsilon/2$ for all $n > N'$. Similarly, since $\lim_{n \rightarrow \infty} y_n = B$, there exists N'' such that $\Delta(y_n) < \varepsilon/2$ for all $n > N''$. Then for $n > \max\{N', N''\}$ we have

$$|(A + B) - (x_n + y_n)| \leq \varepsilon$$

and our proof completes. \square

Theorem 3.1.6. $\lim_{n \rightarrow \infty} \{x_n \cdot y_n\} = A \cdot B$.

Proof. Similar to our first proof, we show that $|x_n|\Delta(y_n)$, $|y_n|\Delta(x_n)$ and $\Delta(x_n) \cdot \Delta(y_n)$ are less than $\frac{\varepsilon}{3}$, since their sum is greater or equal to $|(A \cdot B) - (x_n \cdot y_n)|$. \square

Theorem 3.1.7. $\lim_{n \rightarrow \infty} \left\{ \frac{x_n}{y_n} \right\} = \frac{A}{B}$ when $\forall n \in \mathbb{N} (y_n \neq 0) \wedge (B \neq 0)$.

Proof. If we prove that $|x_n| \cdot \frac{1}{y_n^2} \Delta(y_n) = \frac{\varepsilon}{4}$, $|\frac{1}{y_n}| \Delta(x_n) = \frac{\varepsilon}{4}$, and $0 < \frac{1}{1 - \delta(y_n)} < 2$, we'll have $|\frac{A}{B} - \frac{x_n}{y_n}| < \varepsilon$. \square

Passage to the Limit and Inequalities

Theorem 3.1.8. Let $\{x_n\}$ and $\{y_n\}$ be two convergent sequences with $\lim_{n \rightarrow \infty} x_n = A$ and $\lim_{n \rightarrow \infty} y_n = B$. If $A < B$, then there exists an index $N \in \mathbb{N}$ such that $x_n < y_n$ for all $n > N$.

Proof. Choose a number C such that $A < C < B$. By definition of limit, we can find numbers N' and N'' such that $|x_n - A| < C - A$ for all $n > N'$ and $|y_n - B| < B - C$ for all $n > N''$. Then for $n > N = \max\{N', N''\}$ we shall have $x_n < (A + C - A) = C = (B - (B - C)) < y_n$. \square

Theorem 3.1.9. Suppose the sequences $\{x_n\}$, $\{y_n\}$, and $\{z_n\}$ are such that $x_n \leq y_n \leq z_n$ for all $n > N \in \mathbb{N}$. If the sequences $\{x_n\}$ and $\{z_n\}$ both converge to the same limit, then the sequence $\{y_n\}$ also converges to that limit.

Proof. Suppose $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} z_n = A$. Given $\varepsilon > 0$ choose N' and N'' such that $\forall n > N' (A - \varepsilon < x_n)$ and $\forall n > N'' (z_n < A + \varepsilon)$. Then for $n > N = \max\{N', N''\}$ we shall have $A - \varepsilon < x_n \leq y_n \leq z_n < A + \varepsilon$, which says $|y_n - A| < \varepsilon$, that is $A = \lim_{n \rightarrow \infty} y_n$. \square

Corollary. Suppose $\lim_{n \rightarrow \infty} x_n = A$ and $\lim_{n \rightarrow \infty} y_n = B$. If there exists N such that for all $n > N$ we have

- a) $x_n > y_n$, then $A \geq B$;
- b) $x_n \geq y_n$, then $A \geq B$;
- c) $x_n > B$, then $A \geq B$;
- d) $x_n \geq B$, then $A \geq B$.

Proof. The first two statement can be proved by contradiction using the theorem we mentioned above, and the last two statement are the special cases when $y_n \equiv B$. \square

3.1.3 Questions Involving the Existence of the Limit of a Sequence

The Cauchy Criterion

Definition. A sequence $\{x_n\}$ is called a *fundamental* or *Cauchy* sequence if for any $\varepsilon > 0$ there exists an index $N \in \mathbb{N}$ such that $|x_m - x_n| < \varepsilon$ whenever $n > N$ and $m > N$.

Theorem 3.1.10 (Cauchy's convergence criterion). *A numerical sequence converges if and only if it is a Cauchy sequence.*

Proof. First, we prove that a convergent sequence is a Cauchy sequence. If $\lim_{n \rightarrow \infty} x_n = A$, we set $|x_n - A| < \frac{\varepsilon}{2}$, and we obtain that $|x_m - x_n| \leq |x_m - A| + |x_n - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ for $m, n > N$ (imagine two points on a line, their distance to a fixed point is greater or equal to the distance between these two points).

Next, let $\{x_k\}$ be a Cauchy sequence. We find $|x_m - x_k| < \frac{\varepsilon}{3}$ for $m, k \geq N$. Fix $m = N$, we can see that $x_N - \frac{\varepsilon}{3} < x_k < x_N + \frac{\varepsilon}{3}$, and thus this sequence is bounded. Then we set $a_n := \inf_{k \geq n} x_k$ and $b_n := \sup_{k \geq n} x_k$. Apply the **Nested Interval Lemma** to the close intervals $[a_n, b_n]$ and denote that point A . The inequality can be derived from the following inequalities:

$$\begin{aligned} a_n &\leq x_k \leq b_k \\ |A - x_k| &\leq b_n - a_n \\ x_N - \frac{\varepsilon}{3} &\leq a_n \leq b_n \leq x_N + \frac{\varepsilon}{3} \end{aligned}$$

\square

A Criterion for the Existence of the Limit of a Monotonic Sequence

Definition. A sequence $\{x_n\}$ is *increasing* if $\forall n \in \mathbb{N}(x_n < x_{n+1})$, *nondecreasing* if $\forall n \in \mathbb{N}(x_n \leq x_{n+1})$, *nonincreasing* if $\forall n \in \mathbb{N}(x_n \geq x_{n+1})$, and *decreasing* if $\forall n \in \mathbb{N}(x_n > x_{n+1})$. Sequence of these four types are called *monotonic* sequences.

Definition. A sequence $\{x_n\}$ is *bounded above* if there exists a number M such that $\forall n \in \mathbb{N} (x_n < M)$.

Theorem 3.1.11 (Weierstrass). *In order for a nondecreasing sequence to have a limit, it is necessary and sufficient that it be bounded above.*

Proof. We already know that a convergent sequence is bounded, hence we just need to prove its sufficiency.

Let s be the least upper bound of this sequence. Since it's a nondecreasing sequence, we have $s - \varepsilon < x_N \leq x_n \leq s$ for all $n > N$, and thus $s - x_n < \varepsilon$. \square

Analogously one can prove that in order for a nonincreasing sequence to have a limit, it has to be bounded below.

Corollary. $\lim_{n \rightarrow \infty} \frac{n}{q^n} = 0$ if $q > 1$.

Proof. We prove that this sequence is bounded below (all terms are positive). Let $x = \lim_{n \rightarrow \infty} x_n$ for $n > N$, where N is the index that satisfies $x_{n+1} < x_n$ (from N the sequence is monotonically decreasing). From the relation $x_{n+1} = \frac{n+1}{nq} x_n$, one can have:

$$x = \frac{1}{q} x$$

and hence $(1 - \frac{1}{q})x = 0 \Rightarrow (x = 0)$. \square

Corollary. $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

Proof. There exists $N \in \mathbb{N}$ such that $1 \leq n < (1 + \varepsilon)^n$ for all $n > N$. Thus $1 \leq \sqrt[n]{n} < 1 + \varepsilon$, which implies $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$. \square

Corollary. $\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1$ for any $a > 0$.

Proof. Assume $a \geq 1$ first, use the same technique from above, and then prove $\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1$ for $0 < a < 1$. \square

Example 3.1. $\forall q \in \mathbb{R} \forall n \in \mathbb{N} (\lim_{n \rightarrow \infty} \frac{q^n}{n!} = 0)$.

Proof. If $q = 0$, the assertion is obvious. Since $|\frac{q^n}{n!}| = \frac{|q|^n}{n!}$, let's assume $q > 0$. Use the same technique as we used in proving $\lim_{n \rightarrow \infty} \frac{n}{q^n} = 0$. First prove that $0 < \frac{1}{n+1} < 1$ and $x_{n+1} < x_n$ for all $n > N \in \mathbb{N}$, thus this sequence is monotonically decreasing from N . Let the limit be x , one will have:

$$x = \lim_{n \rightarrow \infty} \frac{q}{n+1} \cdot \lim_{n \rightarrow \infty} x_n = 0 \cdot x = 0$$

\square

The Number e **Theorem 3.1.12** (Jacob Bernoulli's inequality).

$$(1 + \alpha)^n \geq 1 + n\alpha \quad (3.8)$$

holds for $n \in \mathbb{N}$ and $\alpha > -1$.

Proof. Prove by the principle of induction. □

Incidentally, strict inequality holds if $\alpha \neq 0$ and $n > 1$.

Example 3.2. The limit $\lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n$ exists.

Proof. Let $y_n = (1 + \frac{1}{n})^{n+1}$ and $n \geq 2$. Use Bernoulli's inequality, we find that $\frac{y_{n-1}}{y_n} > 1$, and since all terms are positive, this sequence is bounded, monotonically decreasing and hence, has a limit. Then

$$\lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n = \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^{n+1} \cdot \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^{-1} = \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^{n+1}$$

□

Definition. $e := \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n$.**Subsequences and Partial Limits of a Sequence**

Definition. If $x_1, x_2, \dots, x_n, \dots$ is a sequence and $n_1 < n_2 < n_3 < \dots < n_k < \dots$ an increasing sequence of natural numbers, then the sequence $x_{n_1}, x_{n_2}, \dots, x_{n_k}, \dots$ is called a *subsequence* of the sequence $\{x_n\}$.

Lemma (Bolzano-Weierstrass). *Every bounded sequence of real numbers contains a convergent subsequence.*

Proof. Let E be the set of values of the bounded sequence $\{x_n\}$. If E is finite, there exists a point $x \in E$ and a sequence $n_1 < n_2 < \dots$ of indices such that $x_{n_1} = x_{n_2} = \dots = x$. The subsequence is constant and hence converges.

If E is infinite, then by the **Bolzano-Weierstrass principle** from section 2.3.3 it has a limit point x . Use the property of limit points, one can see that $|x_{n_k} - x| < \frac{1}{k}$ and $|x_{n_{k+1}} - x| < \frac{1}{k+1}$. Because $\lim_{k \rightarrow \infty} \frac{1}{k} = 0$, the sequence $x_{n_1}, x_{n_2}, \dots, x_{n_k}, \dots$ converges to x . □

Definition. We shall write $x_n \rightarrow +\infty$ and say that the sequence $\{x_n\}$ *tends to positive infinity* if for each number c there exists $N \in \mathbb{N}$ such that $\forall n > N (x_n > c)$. We can generalize this to both positive and negative infinity:

$$(x_n \rightarrow \infty) := \forall c \in \mathbb{R} \exists N \in \mathbb{N} \forall n > N (c < |x_n|).$$

Lemma. From each sequence of real numbers one can extract either a convergent subsequence or a subsequence that tends to infinity.

Proof. When the sequence $\{x_n\}$ is not bounded, the new case occurs. Then for each $k \in \mathbb{N}$ we can choose $n_k \in \mathbb{N}$ such that $|x_{n_k}| > k$ and $n_k < n_{k+1}$. This sequence is monotonically increasing and not bounded above, hence it tends to infinity. \square

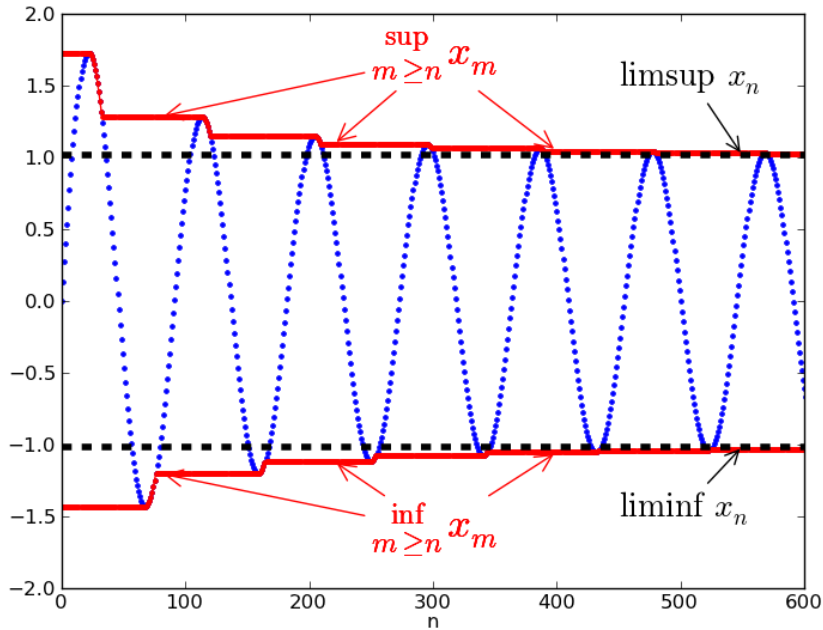
Let $\{x_k\}$ be an arbitrary sequence of real numbers that is bounded below. We can consider the sequence $i_n = \inf_{k \geq n} x_k$. The sequence $\{i_n\}$ has a finite limit $\lim_{n \rightarrow \infty} i_n = l$, or $i_n \rightarrow +\infty$, since $\forall n \in \mathbb{N} (i_n \leq i_{n+1})$.

Definition. The number $l = \lim_{n \rightarrow \infty} \inf_{k \geq n} x_k$ is called the *inferior limit* of the sequence $\{x_k\}$ and denoted $\underline{\lim}_{k \rightarrow \infty} x_k$. If $i_n \rightarrow +\infty$, it is said that the inferior limit of the sequence equals positive infinity, and we write $\underline{\lim}_{k \rightarrow \infty} x_k = +\infty$. If the original sequence $\{x_k\}$ is not bounded below, then we shall have $i_n = \inf_{k \geq n} x_k = -\infty$ and write $\underline{\lim}_{k \rightarrow \infty} x_k = -\infty$.

$$\underline{\lim}_{k \rightarrow \infty} x_k := \lim_{n \rightarrow \infty} \inf_{k \geq n} x_k$$

Similarly, the superior limit of the sequence $\{x_k\}$ can be defined as

$$\overline{\lim}_{k \rightarrow \infty} x_k := \lim_{n \rightarrow \infty} \sup_{k \geq n} x_k$$



Definition. A number (or the symbol $-\infty$ or $+\infty$) is called a *partial limit* of a sequence, if the sequence contains a subsequence converging to that number.

Proposition. The inferior and superior limits of any sequence are respectively the smallest and largest partial limits of the sequence.

Proof. Let's assume that this sequence is bounded. First consider the inferior limit $i = \underline{\lim}_{k \rightarrow \infty} x_k$. The sequence $i_n = \inf_{k \geq n} x_k$ is nondecreasing. Using the definition of the greatest lower bound, we choose by induction numbers $k_n \in \mathbb{N}$ such that $k_n < k_{n+1}$ and $i_{k_n} \leq x_{k_n} < i_{k_n} + \frac{1}{n}$ (Taking i_1 we find k_1 ; taking i_{k_1+1} we find k_2 , etc.). Since $\lim_{n \rightarrow \infty} i_n = \lim_{n \rightarrow \infty} (i_n + \frac{1}{n}) = i$, we have $\lim_{n \rightarrow \infty} x_{k_n} = i$. It is the smallest partial limit since for every $\varepsilon > 0$ there exists $n \in \mathbb{N}$ such that $i - \varepsilon < i_n$, that is $i - \varepsilon < i_n = \inf_{k \geq n} x_k \leq x_k$ for any $k \geq n$. Now we have $i - \varepsilon < x_k$ for $k > n$ means that no partial limit of the sequence can be less than $i - \varepsilon$. But $\varepsilon > 0$ is arbitrary, and hence no partial limit can be less than i . The proof for the superior limit is of course analogous.

Now if the sequence is not bounded below (resp. above), one can select a subsequence of it tending to $-\infty$ (resp. $+\infty$). But then we also have $\underline{\lim}_{k \rightarrow \infty} x_k = -\infty$ (resp. $\overline{\lim}_{k \rightarrow \infty} x_k = +\infty$). Finally, if $\underline{\lim}_{k \rightarrow \infty} x_k = -\infty$ (resp. $\underline{\lim}_{k \rightarrow \infty} x_k = +\infty$), the sequence itself tends to $-\infty$ (resp. $+\infty$). \square

Corollary. A sequence has a limit or tends to $\pm\infty$ iff its inferior and superior limits are the same.

Proof. The cases when $\underline{\lim}_{k \rightarrow \infty} x_k = \overline{\lim}_{k \rightarrow \infty} x_k = \pm\infty$ have been investigated above, and so we may assume that $\underline{\lim}_{k \rightarrow \infty} x_k = \overline{\lim}_{k \rightarrow \infty} x_k = A \in \mathbb{R}$. Since $(i_n = \inf_{k \geq n} x_k) \leq x_k \leq (\sup_{k \geq n} x_k = s_n)$, we have $\lim_{n \rightarrow \infty} x_n = A$. \square

Corollary. A sequence converges iff every subsequence of it converges.

Proof. The inferior and superior limits of a subsequence lie between those of the sequence itself. If the sequence converges, then its subsequences must converge, and their limits are the same. The converse assertion is obvious, since the subsequence can be chosen as the sequence itself. \square

Corollary. The Bolzano-Weierstrass Lemma in its restricted and wider formulations (corresponding to Lemmas at page 32 and page 33) follows from the Proposition we just proved.

Proof. If the sequence $\{x_k\}$ is bounded, then $i = \underline{\lim}_{k \rightarrow \infty} x_k$ and $s = \overline{\lim}_{k \rightarrow \infty} x_k$ are finite and partial limits of the sequence. When $i = s$ some subsequences have a unique limit, and at least two when $i < s$. If the sequence is unbounded on one side or the other, there exists a subsequence tending to the corresponding infinity. \square

3.1.4 Elementary Facts about Series

The Sum of a Series and the Cauchy Criterion for Convergence of Series

Definition. The expression $a_1 + a_2 + \cdots + a_n + \cdots$ is denoted by the symbol $\sum_{n=1}^{\infty} a_n$ and usually called a *series* or an *infinite series*.

Definition. The elements of the sequence $\{a_n\}$, when regarded as elements of the series, are called the *terms* of the series. The element a_n is called the *nth term*.

Definition. The sum $s_n = \sum_{k=1}^n a_k$ is called the *partial sum of the series* or the *nth partial sum of the series*.

Definition. If the sequence $\{s_n\}$ of partial sums of a series converges (resp. diverges), we say the series is *convergent* (resp. *divergent*).

Definition. The limit $\lim_{n \rightarrow \infty} s_n = s$ of the sequence of partial sums of the series, if it exists, is called the *sum of the series*.

One can see that $\sum_{n=1}^{\infty} a_n = s$.

Theorem 3.1.13 (The Cauchy convergence criterion for a series). *The series $a_1 + \cdots + a_n + \cdots$ converges iff for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that the inequalities $m \geq n > N$ imply $|a_n + \cdots + a_m| < \varepsilon$.*

Corollary. *If only a finite number of terms of a series are changed, the resulting new series will converge if the original series did and diverge if it diverged.*

Corollary. *A necessary condition for convergence of the series $a_1 + \cdots + a_n + \cdots$ is that the terms tend to zero as $n \rightarrow \infty$, that is, it is necessary that $\lim_{n \rightarrow \infty} a_n = 0$.*

Proof. Set $m = n$ in the Cauchy convergence criterion and use the definition of the limit of a sequence.

Alternatively, $a_n = s_n - s_{n-1}$, and, given that $\lim_{n \rightarrow \infty} s_n = s$, we have $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = s - s = 0$. \square

Example 3.3. *The series $1 + q + q^2 + \cdots + q^n + \cdots$ is often called the geometric series. It converges iff $|q| < 1$.*

Proof. Suppose $|q| \geq 1$, then we have $|q^n| \geq 1$, and in this case this series does not converge.

Now let $|q| < 1$, and we'll have $s_n = 1 + q + \cdots + q^{n-1} = \frac{1-q^n}{1-q}$ and $\lim_{n \rightarrow \infty} s_n = \frac{1}{1-q}$, since $\lim_{n \rightarrow \infty} q^n = 0$ if $|q| < 1$. \square

Example 3.4. The series $1 + \frac{1}{2} + \cdots + \frac{1}{n} + \cdots$, called the harmonic series, diverges because its partial sums $s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ diverges.

Proof. It's sufficient to prove that its partial sum $s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ diverges. For all $n \in \mathbb{N}$ we have

$$|x_{2n} - x_n| = \frac{1}{n+1} + \cdots + \frac{1}{n+n} > n \cdot \frac{1}{2n} = \frac{1}{2}$$

and our proof completes. \square

Remark Usual laws for dealing with finite sums does not apply to series in general (e.g. insert parentheses to a divergent series).

Absolute Convergence. The Comparison Theorem and Its Consequences

Definition. The series $\sum_{n=1}^{\infty}$ is *absolutely* convergent if the series $\sum_{n=1}^{\infty} |a_n|$ converges.

Since $|a_n + \cdots + a_m| \leq |a_n| + \cdots + |a_m|$, the Cauchy convergence criterion implies that an absolutely convergent series converges, but the converse is generally not true.

Example 3.5. The series $1 - 1 + \frac{1}{2} - \frac{1}{2} + \frac{1}{3} - \frac{1}{3} + \cdots$, whose partial sums are either $\frac{1}{n}$ or 0, converges to 0. However, this sequence does not absolutely converges, and its proof is similar to the proof for the divergence of harmonic series.

Theorem 3.1.14 (Criterion for convergence of series of non-negative terms). A series whose terms are non-negative converges iff the sequence of partial sums is bounded above.

Theorem 3.1.15 (Comparison Theorem). Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series with non-negative terms. If there exists an index $N \in \mathbb{N}$ such that $a_n \leq b_n$ for all $n > N$, then the convergence of the series $\sum_{n=1}^{\infty} b_n$ implies the convergence of $\sum_{n=1}^{\infty} a_n$, and the divergence of $\sum_{n=1}^{\infty} a_n$ implies the divergence of $\sum_{n=1}^{\infty} b_n$.

Proof. Omit terms for these two series for all $n < N$, since a finite number of terms has no effect on the convergence of a series. Denote the partial sum of the sequences as $A_n = \sum_{k=1}^n a_k \leq \sum_{k=1}^n b_k = B_n$. If the series $\{B_n\}$ converges, then $\{A_n\}$ is bounded above. Because $\{A_n\}$ is also non-decreasing (all terms are non-negative), it has a limit, and so do $\sum_{n=1}^{\infty} a_n$. The second assertion is similar and can be proved by contradiction. \square

Corollary (The Weierstrass M-test for absolute convergence). Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be series. Suppose there exists an index X such that $|a_n| \leq b_n$

for all $n > N$. Then a sufficient condition for absolute convergence of the series $\sum_{n=1}^{\infty} a_n$ is that the series $\sum_{n=1}^{\infty} b_n$ converge.

It's often summarized as following: If the terms of a series are majorized (in absolute value) by the terms of a convergent numerical series, then the original series converges absolutely.

Proof. By the comparison theorem the series $\sum_{n=1}^{\infty} |a_n|$ will then converge, and that is what is meant by the absolute convergence of $\sum_{n=1}^{\infty} a_n$. \square

Corollary (Cauchy's Test). Let $\sum_{n=1}^{\infty} a_n$ be a given series and $\alpha = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$. Then the following are true:

- a) if $\alpha < 1$, the series converges absolutely;
- b) if $\alpha > 1$, the series diverges;
- c) there exist both absolutely convergent and divergent series for which $\alpha = 1$.

Proof. For $\alpha < 1$, we find a $q \in \mathbb{R}$ such that $\alpha < q < 1$, and show that the sequence $\sum_{n=1}^{\infty} q^n$ converges and its terms are always greater than $|a_n|$. Thus the original sequence absolutely converges.

For $\alpha > 1$, we find that α is the greatest partial limit of the sequence $\{\sqrt[n]{|a_n|}\}$. Hence for some $k > K(|a_{n_k}| > 1)$, and the necessary condition for convergence ($a_n \rightarrow 0$) does not meet for the original sequence.

If $\alpha = 1$, for example, the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges absolutely, but their superior limit under n th root are both 1. \square

Corollary (d'Alembert's Test). Suppose the limit $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \alpha$ exists for the series $\sum_{n=1}^{\infty} a_n$. Then,

- a) if $\alpha < 1$, the series converges absolutely;
- b) if $\alpha > 1$, the series diverges;
- c) there exist both absolutely convergent and divergent series for which $\alpha = 1$.

Proof. If $\alpha < 1$, then we find a number q such that $\alpha < q < 1$. Fixing q and use the properties of limits, we find an index $N \in \mathbb{N}$ such that $\left| \frac{a_{n+1}}{a_n} \right| < q$ for $n > N$. Since

$$\left| \frac{a_{n+1}}{a_n} \right| \cdot \left| \frac{a_n}{a_{n-1}} \right| \cdots \left| \frac{a_2}{a_1} \right| = \frac{a_{n+1}}{a_1}$$

and therefore we have $|a_{n+1}| < |a_1| \cdot q^n$. But the geometric series $\sum_{n=1}^{\infty} |a_1| q^n$ converges for $|q| < 1$, hence the original series converges.

For $\alpha > 1$, we can find some terms that $\left| \frac{a_{n+1}}{a_n} \right| > 1$, thus it diverges. For case involving $\alpha = 1$, the examples from our proof for **Cauchy's Test** are sufficient. \square

Proposition (Cauchy). If $a_1 \geq a_2 \geq \dots \geq 0$, the series $\sum_{n=1}^{\infty} a_n$ converges iff the series $\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 \dots$ converges.

Proof. By the inequality

$$2^n a_{2^{n+1}} \leq a_{2^n+1} + \dots + a_{2^{n+1}} \leq 2^n a_{2^n} \quad (3.-14)$$

we have

$$\frac{1}{2}(S_{n+1} - a_1) \leq A_{2^{n+1}} - a_1 \leq S_n \quad (3.-14)$$

where $A_k = a_1 + \dots + a_k$ and $S_n = a_1 + 2a_2 + \dots + 2^n a_{2^n}$. $\{A_k\}$ and $\{S_n\}$ are non-decreasing, and hence they are both bounded above or unbounded above. Since all their terms are non-negative, our proof completes. \square

Corollary. The series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges for $p > 1$ and diverges for $p \leq 1$.

Proof. If $p \geq 0$, by our proposition it converges when $2^{1-p} < 1$, that is, $p > 1$. The case when $p \leq 0$ is obvious since all terms are not smaller than 1. \square

The Number e as the Sum of a Series

We know that $e = \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n$. By Newton's binomial formula:

$$\begin{aligned} (1 + \frac{1}{n})^n &= 1 + \frac{n}{1!} \frac{1}{n} + \frac{n(n-1)}{2!} \frac{1}{n^2} + \dots + \frac{(n(n-1) \dots (n-k+1))}{k!} \frac{1}{n^k} + \dots + \frac{1}{n^n} \\ &= 1 + 1 + \frac{1}{2!} (1 - \frac{1}{n}) + \dots + \frac{1}{k!} (1 - \frac{1}{n}) (1 - \frac{2}{n}) \times \dots \\ &\quad \times (1 - \frac{k-1}{n}) + \dots + \frac{1}{n!} (1 - \frac{1}{n}) \dots (1 - \frac{n-1}{n}) \end{aligned}$$

Setting $(1 + \frac{1}{n}) = e_n$ and $1 + 1 + \dots + \frac{1}{2!} + \dots + \frac{1}{n!} = s_n$, we thus have $\forall n \in \mathbb{N} (e_n < s_n)$.

On the other hand, for any fixed k and $n \geq k$, as can be seen from the same expansion, we have

$$1 + 1 + 1 + 1 + \frac{1}{2!} (1 - \frac{1}{n}) + \dots + \frac{1}{k!} (1 - \frac{1}{n}) (1 - \frac{2}{n}) \dots (1 - \frac{k-1}{n}) < e_n$$

As $n \rightarrow \infty$ the left-hand side of the inequality tends to s_k and the right-hand side to e . We can now conclude that $s_k \leq e$ for all $k \in \mathbb{N}$. Then from the relation $e_n < s_n \leq e$ we find that $\lim_{n \rightarrow \infty} s_n = e$.

Definition. $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!} + \dots$

The difference between e and its estimation s_n can be expressed as following

$$\begin{aligned}
 0 < e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots = \\
 &= \frac{1}{(n+1)!} \left[1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots \right] < \\
 &< \frac{1}{(n+1)!} \left[1 + \frac{1}{n+2} + \frac{1}{(n+2)^2} + \cdots \right] \\
 &= \frac{1}{(n+1)!} \frac{1}{1 - \frac{1}{n+2}} = \frac{n+2}{n!(n+1)^2} < \frac{1}{n!n}
 \end{aligned}$$

This estimate of the difference $e - s_n$ can be written as the equality

$$e = s_n + \frac{\theta_n}{n!n}$$

where $0 < \theta_n < 1$.

Hence e is irrational.

e's irrationality. Suppose the contrary, that $e = \frac{p}{q}$, where $p, q \in \mathbb{N}$. Then the number $q!e$ must be an integer, while

$$q!e = q!(s_q + \frac{\theta_q}{q!q}) = q! + \frac{q!}{1!} + \frac{q!}{2!} + \cdots + \frac{q!}{n!} + \frac{\theta_q}{q}$$

and therefore $\frac{\theta_q}{q}$ would have to be an integer, which is impossible. \square

Moreover, e is transcendental.

3.2 The Limit of a Function

3.2.1 Definitions and Examples

Let $E \subset \mathbb{R}$, a be an limit point of E , and $f : E \rightarrow \mathbb{R}$ be a real-valued function defined on E .

Definition. The function $f : E \rightarrow \mathbb{R}$ *tends to* A as x *tends to* a , or that A is the *limit* of f as x *tends to* a , if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - A| < \varepsilon$ for every $x \in E$ such that $0 < |x - a| < \delta$.

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in E (0 < |x - a| < \delta \Rightarrow |f(x) - A| < \varepsilon)$$

which is denoted as $\lim_{E \ni x \rightarrow a} f(x) = A$.

Example 3.6.

$$\lim_{E \ni x \rightarrow 0} x \sin \frac{1}{x} = 0$$

Proof. Let $\varepsilon = \delta$. □

Definition. A *deleted neighborhood* of a point is a neighborhood of the point from which the point itself has been removed.

If $U(a)$ denotes a neighborhood of a , the deleted neighborhood is denoted as $\mathring{U}(a)$.

The sets

$$\begin{aligned} U_E(a) &:= E \cap U(a) \\ \mathring{U}_E(a) &:= E \cap \mathring{U}(a) \end{aligned}$$

will be called respectively a *neighborhood of a in E* and a *deleted neighborhood of a in E* .

if the temporarily-adopted cumbersome symbols $\mathring{U}_E^\delta(a)$ and $V_\mathbb{R}^\varepsilon(A)$ denote the deleted δ -neighborhood of a in E and the ε -neighborhood of A in \mathbb{R} , then the definition of the limit of a function can be rewritten as

$$\left(\lim_{E \ni x \rightarrow a} f(x) = A \right) := \forall V_\mathbb{R}^\varepsilon(A) \exists \mathring{U}_E^\delta(a) (f(\mathring{U}_E^\delta(a)) \subset V_\mathbb{R}^\varepsilon(A)).$$

This expression says that A is the limit of the function f as x tends to a in the set E if for every ε -neighborhood $V_\mathbb{R}^\varepsilon(A)$ of A there exists a deleted δ -neighborhood $\mathring{U}_E^\delta(a)$ of a in E whose image $f(\mathring{U}_E^\delta(a))$ under the mapping $f : E \rightarrow \mathbb{R}$ is entirely contained in $V_\mathbb{R}^\varepsilon(A)$.

Since every neighborhood of a point on the real line contains a symmetric neighborhood (a δ -neighborhood) of the same point, the **final version of our definition for a limit** is:

Definition.

$$\left(\lim_{E \ni x \rightarrow a} f(x) = A \right) := \forall V_\mathbb{R}(A) \exists \mathring{U}_E(a) (f(\mathring{U}_E(a)) \subset V_\mathbb{R}(A)).$$

Example 3.7. The function

$$\operatorname{sgn} x = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ -1 & \text{for } x < 0 \end{cases}$$

(read "signum x ") has no limit as $x \rightarrow 0$.

Proof. Apparently no number distinct from $-1, 0, 1$ can be the limit of the function. But no matter what $\mathring{U}(0)$ we choose, some points of it does not belong to the ε -neighborhood of A with $\varepsilon = \frac{1}{2}$, since $\mathring{U}(0)$ contains both positive and negative points while $V(A)$ can't contain both 1 and -1 at the same time. □

When the function f is defined on a whole deleted neighborhood of a point $a \in \mathbb{R}$, that is, when $\mathring{U}_E(a) = \mathring{U}_{\mathbb{R}}(a) = \mathring{U}(a)$, we adopt the expression $x \rightarrow a$ instead of $E \ni x \rightarrow a$.

Example 3.8.

$$\lim_{x \rightarrow 0} |\operatorname{sgn}(x)| = 1$$

Proof.

$$\forall V(1)(f(\mathring{U}(0)) = 1 \in V(1))$$

□

Example 3.9.

$$\lim_{\mathbb{R}_- \ni x \rightarrow 0} \operatorname{sgn}(x) = -1, \quad \lim_{\mathbb{R}_+ \ni x \rightarrow 0} \operatorname{sgn}(x) = 1$$

Example 3.10. $\lim_{x \rightarrow 0} \sin \frac{1}{x}$ has no limit.

Proof. In any deleted neighborhood of 0 $\mathring{U}(0)$ there are always points of the form $\frac{1}{-\pi/2 + 2\pi n}$ and $\frac{1}{\pi/2 + 2\pi n}$ which assume the values -1 and 1 respectively, but for $\varepsilon < 1$ these two numbers can't both lie in the ε -neighborhood. □

Example 3.11. If

$$E_- = \{x \in \mathbb{R} | x = \frac{1}{-\pi/2 + 2\pi n}, n \in \mathbb{N}\}$$

$$E_+ = \{x \in \mathbb{R} | x = \frac{1}{\pi/2 + 2\pi n}, n \in \mathbb{N}\}$$

then

$$\lim_{E_- \ni x \rightarrow 0} \sin \frac{1}{x} = -1$$

$$\lim_{E_+ \ni x \rightarrow 0} \sin \frac{1}{x} = 1$$

The next proposition, also called the statement of the equivalence of the Cauchy definition of a limit (in terms of neighborhoods) and the Heine definition (in terms of sequences), is:

Proposition. The relation $\lim_{E \ni x \rightarrow a} f(x) = A$ holds iff for every sequence $\{x_n\}$ of points $x_n \in E \setminus a$ converging to a , the sequence $\{f(x_n)\}$ converges to A .

Proof. First, $(\lim_{E \ni x \rightarrow a} f(x) = A) \Rightarrow (\lim_{n \rightarrow \infty} f(x_n) = A)$ is obvious. Now for the converse, if A is not the limit of $f(x)$ as $E \ni x \rightarrow a$, then there exists a neighborhood $V(A)$ such that for any $n \in \mathbb{N}$, there is a point x_n in the deleted $\frac{1}{n}$ -neighborhood of a in E such that $f(x_n) \notin V(A)$. But this means that the sequence $\{f(x_n)\}$ does not converge to A . □

3.2.2 Properties of the Limit of a Function

Properties of Deleted Neighborhood of a Limit Point of a Set

Example 3.12. $\dot{U}_E(a) \neq \emptyset$, that is, the deleted neighborhood of the point in E is nonempty.

Example 3.13. $\forall \dot{U}'_E(a) \forall \dot{U}''_E(a) \exists \dot{U}_E(a) (\dot{U}_E(a) \subset \dot{U}'_E(a) \cap \dot{U}''_E(a))$, that is, the intersection of any pair of deleted neighborhoods contains a deleted neighborhood.

General Properties of the Limit of a Function

Definition. A function $f : E \rightarrow \mathbb{R}$ assuming only one value is called *constant*. A function $f : E \rightarrow \mathbb{R}$ is called *ultimately constant* as $E \ni x \rightarrow a$ if it is constant in some deleted neighborhood $\dot{U}_E(a)$, where a is a limit point of E .

Definition. A function $f : E \rightarrow \mathbb{R}$ is *bounded*, *bounded above*, or *bounded below* respectively if there is a number $C \in \mathbb{R}$ such that $|f(x)| < C$, $f(x) < C$, or $C < f(x)$ for all $x \in E$.

Theorem 3.2.1. a) $(f : E \rightarrow \mathbb{R} \text{ is ultimately the constant } A \text{ as } E \ni x \rightarrow a) \Rightarrow (\lim_{E \ni x \rightarrow a} f(x) = A)$.

b) $(\exists \lim_{E \ni x \rightarrow a} f(x)) \Rightarrow (f : E \rightarrow \mathbb{R} \text{ is ultimately bounded as } E \ni x \rightarrow a)$.

c) $(\lim_{E \ni x \rightarrow a} f(x) = A_1) \wedge (\lim_{E \ni x \rightarrow a} f(x) = A_2) \Rightarrow (A_1 = A_2)$.

Proof. The proof is similar to how we prove that a ultimately constant sequence converges and its limits is unique. \square

Passage to the Limit and Arithmetic Operations

Definition. If two numerical-valued functions $f : E \rightarrow \mathbb{R}$ and $g : E \rightarrow \mathbb{R}$ have a common domain of definition E , the *sum*, *product*, and *quotient* are respectively the functions defined on the same set by the following formulas

$$\begin{aligned} (f + g)(x) &:= f(x) + g(x) \\ (f \cdot g)(x) &:= f(x) \cdot g(x) \\ \left(\frac{f}{g}\right)(x) &:= \frac{f(x)}{g(x)} \text{ if } \forall x \in E (g(x) \neq 0) \end{aligned}$$

Definition. A function $f : E \rightarrow \mathbb{R}$ is said to be *infinitesimal* as $E \ni x \rightarrow a$ if $\lim_{E \ni x \rightarrow a} f(x) = 0$.

Proposition. a) If $\alpha : E \rightarrow \mathbb{R}$ and $\beta : E \rightarrow \mathbb{R}$ are infinitesimal functions as $E \ni x \rightarrow a$, then their sum $\alpha + \beta : E \rightarrow \mathbb{R}$ is also infinitesimal as $E \ni x \rightarrow a$.

b) If $\alpha : E \rightarrow \mathbb{R}$ and $\beta : E \rightarrow \mathbb{R}$ are infinitesimal functions as $E \ni x \rightarrow a$, then their product $\alpha \cdot \beta : E \rightarrow \mathbb{R}$ is also infinitesimal as $E \ni x \rightarrow a$.

c) If $\alpha : E \rightarrow \mathbb{R}$ is infinitesimal as $E \ni x \rightarrow a$ and $\beta : E \rightarrow \mathbb{R}$ is ultimately bounded as $E \ni x \rightarrow a$, then the product $\alpha \cdot \beta : E \rightarrow \mathbb{R}$ is infinitesimal as $E \ni x \rightarrow a$.

Proof. For assertion **a)**, we set $|\alpha(x)| < \frac{\varepsilon}{2}$.

Assertion **b)** is a special case of assertion **c)**, since every function that has a limit is ultimately bounded.

Let M bounds $|\beta(x)|$, and set $|\alpha(x)| < \frac{\varepsilon}{M}$ will help to prove assertion **c)**. \square

Example 3.14.

$$\left(\lim_{E \ni x \rightarrow a} f(x) = A \right) \Leftrightarrow (f(x) = A + \alpha(x) \wedge \lim_{E \ni x \rightarrow a} \alpha(x) = 0)$$

Proof. This follows immediately from the definition of limit, by virtue which

$$\lim_{E \ni x \rightarrow a} f(x) = A \Leftrightarrow \lim_{E \ni x \rightarrow a} (f(x) - A) = 0$$

\square

Theorem 3.2.2. Let $f : E \rightarrow \mathbb{R}$ and $g : E \rightarrow \mathbb{R}$ be two functions with a common domain of definition. If $\lim_{E \ni x \rightarrow a} f(x) = A$ and $\lim_{E \ni x \rightarrow a} g(x) = B$, then

$$\begin{aligned} \lim_{E \ni x \rightarrow a} (f + g)(x) &= A + B \\ \lim_{E \ni x \rightarrow a} (f \cdot g)(x) &= A \cdot B \\ \lim_{E \ni x \rightarrow a} \left(\frac{f}{g} \right)(x) &= \frac{A}{B} \text{ if } \forall x \in E (g(x) \neq 0 \text{ and } B \neq 0) \end{aligned}$$

Proof. These properties can be derived from the properties of the limit of sequences according to the proposition above. In order to complete the proof, one only need to convert "for some $N \in \mathbb{N}$ " to $\dot{U}_E(a)$. \square

alternative proof with infinitesimal functions. Let $f(x) = \alpha(x)$ and $g(x) = \beta(x)$, where $\alpha(x)$ and $\beta(x)$ are infinitesimal functions. The first two assertions are obvious. For the quotient one, find the value of $\frac{f(x)}{g(x)} - \frac{A}{B}$ after proving that $\frac{1}{g(x)}$ is ultimately bounded as $E \ni x \rightarrow a$. \square

Passage to the Limit and Inequalities

Theorem 3.2.3. a) If the functions $f : E \rightarrow \mathbb{R}$ and $g : E \rightarrow \mathbb{R}$ are such that $\lim_{E \ni x \rightarrow a} f(x) = A$, and $\lim_{E \ni x \rightarrow a} g(x) = B$ and $A < B$, then there exists a deleted neighborhood $\dot{U}_E(a)$ of a in E at each point of which $f(x) < g(x)$.

b) If the relations $f(x) \leq g(x) \leq h(x)$ hold for the functions $f : E \rightarrow \mathbb{R}$, $g : E \rightarrow \mathbb{R}$ and $h : E \rightarrow \mathbb{R}$, and if $\lim_{E \ni x \rightarrow a} f(x) = \lim_{E \ni x \rightarrow a} h(x) = C$, then the limit of $g(x)$ exists as $E \ni x \rightarrow a$, and $\lim_{E \ni x \rightarrow a} g(x) = C$.

Proof. **a)** Choose a number C such that $A < C < B$. By definition of limit, we find deleted neighborhoods $\dot{U}'_E(a)$ and $\dot{U}''_E(a)$ such that $|f(x) - A| < C - A$ for $x \in \dot{U}'_E(a)$ and $|g(x) - B| < B - C$ for $x \in \dot{U}''_E(a)$. Then at any point of a deleted neighborhood $\dot{U}_E(a)$ contained in $\dot{U}'_E(a) \cup \dot{U}''_E(a)$, we have

$$f(x) < A + (C - A) = C = B - (B - C) < g(x)$$

b) This proof is similar. □

Corollary. Suppose $\lim_{E \ni x \rightarrow a} f(x) = A$ and $\lim_{E \ni x \rightarrow a} g(x) = B$. Let $\dot{U}_E(a)$ be a deleted neighborhood of a in E .

a) If $f(x) > g(x)$ for all $x \in \dot{U}_E(a)$, then $A \geq B$;

b) $f(x) \geq g(x)$ for all $x \in \dot{U}_E(a)$, then $A \geq B$;

c) $f(x) > B$ for all $x \in \dot{U}_E(a)$, then $A \geq B$;

d) $f(x) \geq B$ for all $x \in \dot{U}_E(a)$, then $A \geq B$.

Proof. Assertion a) and b) can be proved by contradiction and the theorem mentioned above. Set $g(x) \equiv B$ and we prove assertion c) and d). □

Two Important Examples

Example 3.15.

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

Proof. The geometric proof is sufficient with the following conditions:

$$\begin{aligned} |\sin x| &\leq |x| \\ 0 \leq |\sin x| &\leq |x| \Rightarrow \left(\lim_{x \rightarrow 0} |\sin x| = 0 \right) \end{aligned}$$

□

Now we define the exponential, logarithmic, and power functions using the theory of real numbers and limits. **A) The exponential function**

Let $a > 1$

1⁰ For $n \in \mathbb{N}$ we define inductively $a^1 := a$, $a^{n+1} := a^n \cdot a$

2⁰ $a^0 = 1$, $a^{-n} = \frac{1}{a^n}$

3⁰ We define $a = a^1 = (a^{1/n})^n$ and $a^{m/n} := (a^{1/n})^m$, Now we have defined a^r for $r \in \mathbb{Q}$.

4⁰ By induction, for $x > 0$ and $y > 0$ we have $(x < y) \Leftrightarrow (x^n < y^n)$, and in particular, $(x = y) \Leftrightarrow (x^n = y^n)$.

5⁰ $a^{(mk)/(nk)} = a^{m/n}$ for $k \in \mathbb{Z}$ and $a^{m_1/n_1} \cdot a^{m_2/n_2} = a^{m_1/n_1 + m_2/n_2}$

6⁰ $(r_1 < r_2) \Rightarrow (a^{r_1} < a^{r_2})$ for any $r_1, r_2 \in \mathbb{Q}$.

7⁰ For $r^0 \in \mathbb{Q}$ $(\lim_{\mathbb{Q} \ni r \rightarrow r_0} a^r = a^{r_0})$.

Proof. First we prove that $a^p \rightarrow 1$ as $\mathbb{Q} \ni p \rightarrow 0$ by use the inequality $a^{-1/n} < a^p < a^{1/n}$. Then we choose δ such that $1 - \varepsilon a^{-r_0} < a^p < 1 + \varepsilon a^{-r_0}$ for $|p| < \delta$. If now $|r - r_0| < \delta$, we have $a^{r_0}(1 - \varepsilon a^{-r_0}) < a^{r_0} \cdot a^{r-r_0} < a^{r_0}(1 + \varepsilon a^{-r_0})$, which says $a^{r_0} - \varepsilon < a^r < a^{r_0} + \varepsilon$. \square

8⁰ Let $x \in \mathbb{R}$, $s = \sup_{\mathbb{Q} \ni r < x} a^r$, and $i = \inf_{\mathbb{Q} \ni r > x} a^r$. We show that $s = i$.

Proof.

$$a^{r_1} \leq s \leq i \leq a^{r_2}$$

and for $0 < |r_2 - r_1| < \delta$ we have $0 \leq i - s \leq \varepsilon$. \square

We now define $a^x := s = i$.

9⁰ $a^x = \lim_{\mathbb{Q} \ni r \rightarrow x} a^r$.

Proof. We find $r' < x$ such that $s - \varepsilon < a^{r'} \leq s = a^x$ and $r'' > x$ such that $a^x = i \leq a^{r''} < i + \varepsilon$. Then for all $r \in \mathbb{Q}$ in (r', r'')

$$a^x - \varepsilon < a^r < a^x + \varepsilon$$

\square

10⁰ For $x_1, x_2 \in \mathbb{R}$ and $a > 1$, $(x_1 < x_2) \Rightarrow (a^{x_1} < a^{x_2})$.

11⁰ For any $x_1, x_2 \in \mathbb{R}$, $a^{x_1} \cdot a^{x_2} = a^{x_1+x_2}$.

$$12^0 \quad \lim_{x \rightarrow x_0} a^x = a^{x_0}.$$

Proof. The proof is similar with the proof for 7⁰. □

13⁰ The range of values of the function $x \mapsto a^x$ is the set \mathbb{R}_+ .

Proof. The proof is similar with our first step for proving the irrationality of $\sqrt{2}$. □

14⁰ We repeat the construction mentioned above for $0 < a < 1$. In 6⁰ and 10⁰ we find that $(x_1 < x_2) \Rightarrow (a^{x_1} > a^{x_2})$ where $0 < a < 1$.

Definition. The mapping $x \mapsto a^x$ is called the *exponential* function with base a . In the case $a = e$, it's denoted with $\exp(x)$, but in general it's denoted with $\exp_a(x)$.

B) The logarithmic function

The properties of the exponential function show that it's bijective. Hence it has an inverse.

Definition. The mapping inverse to $\exp_a : \mathbb{R} \rightarrow \mathbb{R}_+$ is called the *logarithm to base a* ($0 < a, a \neq 1$), and is denoted

$$\log_a : \mathbb{R}_+ \rightarrow \mathbb{R}$$

for base $a = e$, the logarithm is called the *natural logarithm* and is denoted $\ln : \mathbb{R}_+ \rightarrow \mathbb{R}$.

By definition of the logarithm as the function inverse to the exponential function, we have

$$\begin{aligned} \forall x \in \mathbb{R} (\log_a(a^x) &= x) \\ \forall y \in \mathbb{R}_+ (a^{\log_a(y)} &= y) \end{aligned}$$

$$1' \quad \log_a a = 1.$$

$$2' \quad \log_a(y_1 \cdot y_2) = \log_a y_1 + \log_a y_2.$$

$$3' \quad \log_a y \rightarrow \log_a y_0 \text{ as } \mathbb{R}_+ \ni y \rightarrow y_0 \in \mathbb{R}_+.$$

Proof. We verify that $y_0 a^{-\varepsilon} < y < y_0 a^\varepsilon$ when $a > 1$ and take their logarithms. □

4' $(\log_a y_1 < \log_a y_2) \Leftrightarrow (y_1 < y_2)$ if $a > 1$ and $(\log_a y_1 > \log_a y_2) \Leftrightarrow (y_1 < y_2)$ if $0 < a < 1$.

5' The range of values of the function $\log_a : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the set \mathbb{R} .

6' $\log_a(b^\alpha) = \alpha \log_a b$ holds for any $b > 0$ and any $\alpha \in \mathbb{R}$.

Proof. First we verify that the equality holds in the following conditions:
 $\alpha \in \mathbb{N}$, $\alpha = -1$, $\alpha \in \mathbb{Z}$, $\alpha = \frac{1}{n}$ for $n \in \mathbb{Z}$, $\alpha \in \mathbb{Q}$, $\log_a b^\alpha = \lim_{\mathbb{Q} \ni r \rightarrow \alpha} \log_a b^r =$
 $\lim_{\mathbb{Q} \ni r \rightarrow \alpha} r \log_a b = \alpha \log_a b$ where $r \in \mathbb{Q}$ and $\log_a b^r = r \log_a b$, finally, $(a^\alpha)^\beta = a^{\alpha\beta}$. \square

C) The power function

Definition. The function $x \mapsto x^\alpha$ defined on the set \mathbb{R}_+ is called a power function, and the number α is called its *exponent*.

$$x^\alpha = a^{\log_a(x^\alpha)} = a^{\alpha \log_a(x)}$$

3.2.3 The General Definition of the Limit of a Function (Limit over a Base)

Bases; Definition and Elementary Properties

Definition. A set \mathcal{B} of subsets $B \subset X$ of a set X is called a *base* in X if the following conditions hold: $B_1) \forall B \in \mathcal{B} (B \neq \emptyset)$.
 $B_2) \forall B_1 \in \mathcal{B} \forall B_2 \in \mathcal{B} \exists B \in \mathcal{B} (B \subset B_1 \cap B_2)$. These two conditions are related to the two properties of deleted neighborhood we mentioned in the previous subsection. "Base" here is an abbreviation for what is called a "filter base".

Notation for the base	Read	Elements of the base	Definition of and notation for elements
$x \rightarrow a$	x tends to a	Deleted neighborhoods of $a \in \mathbb{R}$	$\dot{U}(a) := \{x \in \mathbb{R} a - \delta_1 < x < a + \delta_2 \wedge x \neq a\}$, where $\delta_1 > 0$ and $\delta_2 > 0$.
$x \rightarrow \infty$	x tends to infinity	Neighborhoods of infinity	$U(\infty) := \{x \in \mathbb{R} \delta < x \}$, where $\delta \in \mathbb{R}$
$x \rightarrow a, x \in E$ or $x \rightarrow_{\in E} a$ or $E \ni x \rightarrow a$	x tends to a in E	Deleted neighborhoods ¹ of a in E	$\dot{U}_E(a) := E \cap \dot{U}(a)$
$x \rightarrow \infty, x \in E$ or $E \ni x \rightarrow \infty$ or $x \rightarrow_{\in E} \infty$	x tends to infinity in E	Neighborhoods ² of infinity in E	$U_E(\infty) := E \cap U(\infty)$

If $E = E_a^+ = \{x \in \mathbb{R} | x > a\}$ (resp. $E = E_a^- = \{x \in \mathbb{R} | x < a\}$), we write $x \rightarrow a + 0$ (resp. $x \rightarrow a - 0$), and we say that x *tends to a from the right* (resp. *x tends to a from the left*) or *through larger values* (resp. *through smaller values*). When $a = 0$ we write $x \rightarrow +0$ (resp. $x \rightarrow -0$).

The notation $E \ni x \rightarrow a + 0$ (resp. $E \ni x \rightarrow a - 0$) will be used. It means that x tends to a in E while remaining larger (resp. smaller) than a .

If

$$E = E_\infty^+ = \{x \in \mathbb{R} | c < x\} \text{ (resp. } E = E_\infty^- = \{x \in \mathbb{R} | x < c\})$$

we write $x \rightarrow +\infty$ (resp. $x \rightarrow -\infty$). When $E = \mathbb{N}$, we shall write $n \rightarrow \infty$ instead of $x \rightarrow \infty, x \in \mathbb{N}$.

The Limit of a Function over a Base

Definition. Let $f : X \rightarrow \mathbb{R}$ be a function defined on a set X and \mathcal{B} a base in X . A number A is called the *limit of the function f over the base \mathcal{B}* if for every neighborhood $V(A)$ of A there is an element $B \in \mathcal{B}$ whose image $f(B)$ is contained in $V(A)$.

$$(\lim_{\mathcal{B}} f(x) = A) := \forall V(A) \exists B \in \mathcal{B} (f(B) \subset V(A))$$

Definition. A function $f : X \rightarrow \mathbb{R}$ is *ultimately constant over the base \mathcal{B}* if there exists a number $A \in \mathbb{R}$ and an element $B \in \mathcal{B}$ such that $f(x) = A$ for all $x \in B$.

¹It is assumed that a is a limit point of E

²It is assumed that E is not bounded

Definition. A function $f : X \rightarrow \mathbb{R}$ is *ultimately bounded over the base \mathcal{B}* if there exists a number $c > 0$ and an element $B \in \mathcal{B}$ such that $|f(x)| < c$ for all $x \in B$.

Definition. A function $f : X \rightarrow \mathbb{R}$ is *infinitesimal over the base \mathcal{B}* if $\lim_{\mathcal{B}} f(x) = 0$.

3.2.4 Existence of the Limit of a Function

The Cauchy Criterion

Definition. The *oscillation* of a function $f : X \rightarrow \mathbb{R}$ on a set $E \subset X$ is

$$\omega(f, E) := \sup_{x_1, x_2 \in E} |f(x_1) - f(x_2)|$$

that is, the least upper bound of the absolute value of the difference of the values of the function at two arbitrary points $x_1, x_2 \in E$.

Theorem 3.2.4 (The Cauchy Criterion for the existence of a limit of a function). *Let X be a set and \mathcal{B} a base in X .*

A function $f : X \rightarrow \mathbb{R}$ has a limit over the base \mathcal{B} iff for every $\varepsilon > 0$ there exists $B \in \mathcal{B}$ such that the oscillation of f on B is less than ε .

$$\exists \lim_{\mathcal{B}} f(x) \Leftrightarrow \forall \varepsilon > 0 \exists B \in \mathcal{B} (\omega(f, B) < \varepsilon)$$

Proof. Setting $m_B = \inf_{x \in B} f(x)$ and $M_B = \sup_{x \in B} f(x)$, and remarking that $m_{B_1} \leq m_{B_1 \cap B_2} \leq M_{B_1 \cap B_2} \leq M_{B_2}$ for any elements B_1 and B_2 of the base \mathcal{B} , we find by the axiom of completeness that there exists a number $A \in \mathbb{R}$ separating the numerical sets $\{m_B\}$ and $\{M_B\}$, where $B \in \mathcal{B}$. Since $\omega(f, B) = M_B - m_B$, we can now conclude that, since $\omega(f, B) < \varepsilon$, we have $|f(x) - A| < \varepsilon$ at every point $x \in B$. \square

Remark: When $X = \mathbb{N}$ and \mathcal{B} is the base $n \rightarrow \infty, n \in \mathbb{N}$, $\omega(f, B) < \varepsilon$ where $B \in \mathcal{B}$ is equivalent with the assertion that this sequence is a Cauchy sequence.

The Limit of a Composite Function

Theorem 3.2.5 (The Limit of a Composite Function). *Let Y be a set, \mathcal{B}_Y a base in Y , and $g : Y \rightarrow \mathbb{R}$ a mapping having a limit over the base \mathcal{B}_Y . Let X be a set, \mathcal{B}_X a base in X , and $f : X \rightarrow Y$ a mapping of X into Y such that for every element $B_Y \in \mathcal{B}_Y$ there exists $B_X \in \mathcal{B}_X$ whose image $f(B_X)$ is contained in B_Y .*

Under these hypotheses, the composition $g \circ f : X \rightarrow \mathbb{R}$ of the mapping f and g is defined and has a limit over the base \mathcal{B}_X and $\lim_{\mathcal{B}_X} (g \circ f)(x) = \lim_{\mathcal{B}_Y} g(y)$

Proof. Suppose $\lim_{\mathcal{B}_Y} g(y) = A$. $g(B_Y) \subset V(A) \wedge (\exists B_X \in \mathcal{B}_X (f(B_X) \subset B_Y)) \Rightarrow (g \circ f)(B_X) = g(f(B_X)) \subset g(B_Y) \subset V(A)$. \square

Example 3.16.

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e$$

Example 3.17.

$$\lim_{x \rightarrow +\infty} \frac{x}{q^x} = 0 \text{ if } q > 1$$

Example 3.18.

$$\lim_{x \rightarrow +\infty} \frac{\log_a x}{x} = 0$$

The Limit of a Monotonic Function

Theorem 3.2.6 (Criterion for the Existence of a Limit of a Monotonic Function). Assume $i = \inf E$ and $s = \sup E$ are limit points of the set E and $f : E \rightarrow \mathbb{R}$ a monotonic function on E . A necessary and sufficient condition for this function that is non-decreasing on the set E to have a limit as $x \rightarrow s, x \in E$, is that it be bounded above. For this function to have a limit as $x \rightarrow i, x \in E$, it is necessary and sufficient that it be bounded below.

Proof. For $x \rightarrow s$, use the definition of least upper bound and the fact the set $\{x \in E | x_0 < x < s\}$ is an element of the base $x \rightarrow s, x \in E$, where for a given $\varepsilon > 0$ we have $A - \varepsilon < f(x_0) \leq A$ and $A = \sup_{x \in E \setminus \{s\}} f(x)$. \square

Comparison of the Asymptotic Behavior of Functions

Definition. We shall say that a certain property of functions or a certain relation between functions holds *ultimately over a given base* \mathcal{B} is there exists $B \in \mathcal{B}$ on which it holds.

Definition. The function f is said to be *infinitesimal* compared with the function g over the base \mathcal{B} , and we write $f =_{\mathcal{B}} o(g)$ or $f = o(g)$ over \mathcal{B} if the relation $f(x) = \alpha(x)g(x)$ holds ultimately over \mathcal{B} , where $\alpha(x)$ is a function that is infinitesimal over \mathcal{B} .

Definition. If $f =_{\mathcal{B}} o(g)$ and g is itself infinitesimal over \mathcal{B} , we say that f is an *infinitesimal of higher order than* g over \mathcal{B} .

Definition. A function that tends to infinity over a given base is said to be an *infinite function* or simply an *infinity* over the given base.

Definition. If f and g are infinite functions over \mathcal{B} and $f =_{\mathcal{B}} o(g)$, we say that g is a *higher order infinity* than f over \mathcal{B} .

Example 3.19.

$$\lim_{x \rightarrow +\infty} \frac{x^\alpha}{a^x} = 0$$

that is, $x^\alpha = o(a^x)$ as $x \rightarrow +\infty$ for $a > 1$ and any $\alpha \in \mathbb{R}$.

Example 3.20.

$$\lim_{x \rightarrow +\infty} \frac{\log_a x}{x^\alpha} = 0$$

that is, for $\alpha > 0$ we have $\log_a x = o(x^\alpha)$ as $x \rightarrow +\infty$.

Definition. The notation $f =_{\mathcal{B}} O(g)$ or $f = O(g)$ over the base \mathcal{B} means that the relation $f(x) = \beta(x)g(x)$ holds ultimately over \mathcal{B} where $\beta(x)$ is ultimately bounded over \mathcal{B} .

Definition. The functions f and g are of the same order over \mathcal{B} , and we write $f \asymp g$ over \mathcal{B} , if $f =_{\mathcal{B}} O(g)$ and $g =_{\mathcal{B}} O(f)$ simultaneously. This condition is equivalent to the condition that there exist $c_1 > 0$ and $c_2 > 0$ and an element $B \in \mathcal{B}$ such that the relations

$$c_1|g(x)| \leq |f(x)| \leq c_2|g(x)|$$

Example 3.21. The functions $(2 + \sin x)x$ and x are of the same order as $x \rightarrow \infty$, but $(1 + \sin x)$ and x are not of the same order as $x \rightarrow \infty$.

Definition. If the relation $f(x) = \gamma(x)g(x)$ holds ultimately over \mathcal{B} where $\lim_{\mathcal{B}} \gamma(x) = 1$, we say that the function f behaves asymptotically like g over \mathcal{B} , or that f is equivalent to g over \mathcal{B} , and we write $f \sim_{\mathcal{B}} g$ or $f \sim g$ over \mathcal{B} . This relation is indeed an equivalent relation.

Example 3.22.

$$\ln(1+x) \sim x$$

as $x \rightarrow 0$.

Example 3.23.

$$e^x = 1 + x + o(x)$$

as $x \rightarrow 0$.

Example 3.24.

$$(1+x)^\alpha = 1 + \alpha x + o(x)$$

as $x \rightarrow 0$.

Proposition. If $f \sim_{\mathcal{B}} \tilde{f}$, then $\lim_{\mathcal{B}} f(x)g(x) = \lim_{\mathcal{B}} \tilde{f}(x)g(x)$, provided one of these limits exists.

This rule should not be extended to sums and differences of functions.

Proposition. For a given base

- a) $o(f) + o(f) = o(f)$;
- b) $o(f)$ is also $O(f)$;
- c) $o(f) + O(f) = O(f)$;
- d) $O(f) + O(f) = O(f)$;
- e) if $g(x) \neq 0$, then $\frac{o(f(x))}{g(x)} = o(\frac{f(x)}{g(x)})$ and $\frac{O(f(x))}{g(x)} = O(\frac{f(x)}{g(x)})$;

Here the equality sign is used in the sense of "is". The symbols $o(\cdot)$ and $O(\cdot)$ do not really denote a function, but rather indicate its asymptotic behavior, a behavior that many functions may have simultaneously.

$$\begin{aligned}
 e^x &= 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots \text{ for } x \in \mathbb{R} \\
 \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots + \frac{(-1)^k}{2k!}x^{2k} + \cdots \text{ for } x \in \mathbb{R} \\
 \sin x &= \frac{1}{1!}x - \frac{1}{3!}x^3 + \cdots + \frac{(-1)^k}{(2k+1)!}x^{2k+1} + \cdots \text{ for } x \in \mathbb{R} \\
 \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots + \frac{(-1)^{n-1}}{n}x^n + \cdots \text{ for } |x| < 1 \\
 (1+x)^\alpha &= 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots \\
 &\quad + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + \cdots \text{ for } |x| < 1
 \end{aligned}$$

And as $x \rightarrow 0$

$$\begin{aligned}
 e^x &= 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + O(x^{n+1}) \\
 \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + \frac{(-1)^k}{2k!}x^{2k} + O(x^{2k+2}) \\
 \sin x &= \frac{1}{1!}x - \frac{1}{3!}x^3 + \cdots + \frac{(-1)^k}{(2k+1)!}x^{2k+1} + O(x^{2k+3}) \\
 \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + \frac{(-1)^{n-1}}{n}x^n + O(x^{n+1}) \\
 (1+x)^\alpha &= 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots \\
 &\quad + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + O(x^{n+1})
 \end{aligned}$$

These formulas can be derived from Taylor's formula, and they are usually the most efficient method of finding the limits of the elementary functions.

When doing so, it is useful to keep in mind that $O(x^{m+1}) = x^{m+1} \cdot O(1) = x^m \cdot xO(1) = x^m o(1) = o(x^m)$ as $x \rightarrow 0$.

4

Continuous Functions

4.1 Basic Definitions and Examples

4.1.1 Continuity of a Function at a Point

Definition. A real-valued function f is *continuous at the point a* if for any neighborhood $V(f(a))$ there is a neighborhood $U(a)$ of a whose image under the mapping f is contained in $V(f(a))$.

$$\begin{aligned} f \text{ is continuous at } a &:= (\forall V(f(a))) \exists U(a) (f(U(a)) \subset V(f(a))) \\ &\forall \varepsilon > 0 \exists U(a) \forall x \in U(a) (|f(x) - f(a)| < \varepsilon) \\ &\forall \varepsilon > 0 \exists \delta > 0 \forall x \in \mathbb{R} (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon) \end{aligned}$$

This is equivalent to the condition that $\lim_{x \rightarrow a} f(x)$ exists and $f(a)$ is defined.

$$(f : E \rightarrow \mathbb{R} \text{ is continuous at } a \in E) \Leftrightarrow (\exists \lim_{\mathcal{B}_a} f(x))$$

where \mathcal{B}_a is the base consists of neighborhoods (not deleted neighborhoods) of $a \in E$.

Definition (General Case). A function $f : E \rightarrow \mathbb{R}$ is *continuous at the point $a \in E$* if for every neighborhood $V(f(a))$ of the value $f(a)$ that the function assumes at a there exists a neighborhood $U_E(a)$ of a in E whose image $f(U_E(a))$ is contained in $V(f(a))$.

$$\begin{aligned} (f : E \rightarrow \mathbb{R} \text{ is continuous at } a \in E) &:= \\ &(\forall V(f(a))) \exists U_E(a) (f(U_E(a)) \subset V(f(a))) \\ &\forall \varepsilon > 0 \exists U_E(a) \forall x \in U_E(a) (|f(x) - f(a)| < \varepsilon) \\ &\forall \varepsilon > 0 \exists \delta > 0 \forall x \in \mathbb{R} (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon) \end{aligned}$$

Definition. The quantity $\omega(f, a) = \lim_{\delta \rightarrow +0} \omega(f, U_E^\delta(a))$ is called the *oscillation* of $f : E \rightarrow \mathbb{R}$ at a .

The oscillation of a function on a subset of a set does not exceed its oscillation on the set itself, so that $\omega(f, U_E^\delta(a))$ is a non-decreasing function of δ .

$$(f : E \rightarrow \mathbb{R} \text{ is continuous at } a \in E) \Leftrightarrow (\omega(f, a) = 0)$$

Definition. A function $f : E \rightarrow \mathbb{R}$ is continuous on the set E if it is continuous at each point of E .

The set of all continuous real-valued functions defined on a set E will be denoted $C(E, \mathbb{R})$, or, more, $C(E)$.

Example 4.1. The function $f(x) = \sin x$ is continuous on \mathbb{R} .

Proof. For any point $x_0 \in \mathbb{R}$ we have

$$\begin{aligned} |\sin x - \sin x_0| &= \left| 2 \cos \frac{x + x_0}{2} \sin \frac{x - x_0}{2} \right| \\ &\leq 2 \left| \sin \frac{x - x_0}{2} \right| \leq 2 \left| \frac{x - x_0}{2} \right| = |x - x_0| < \varepsilon \end{aligned}$$

provided $|x - x_0| < \delta = \varepsilon$. Here we used the inequality $|\sin x| \leq |x|$. \square

Example 4.2. Any sequence $f : \mathbb{N} \rightarrow \mathbb{R}$ is a function that is continuous on the set \mathbb{N} of natural numbers, since each point of \mathbb{N} is isolated.

4.1.2 Points of Discontinuity

Definition. If the function $f : E \rightarrow \mathbb{R}$ is not continuous at a point of E , this point is called a point of *discontinuity* or simply a *discontinuity* of f .

Definition. If a point of discontinuity $a \in E$ of the function $f : E \rightarrow \mathbb{R}$ is such that there exists a continuous function $\tilde{f} : E \rightarrow \mathbb{R}$ such that $f|_{E \setminus a} = \tilde{f}|_{E \setminus a}$, then a is called a *removable discontinuity* of the function f .

Definition. The point $a \in E$ is called a discontinuity of *first kind* for the function $f : E \rightarrow \mathbb{R}$ if the following limits exist:

$$\lim_{E \ni x \rightarrow a-0} f(x) := f(a-0), \quad \lim_{E \ni x \rightarrow a+0} f(x) := f(a+0)$$

but at least one of them is not equal to $f(a)$. If at least one of the two limits does not exist, then a is called a discontinuity of *second kind*.

Example 4.3. The Dirichlet function

$$\mathcal{D}(x) = \begin{cases} 1 & , \text{ if } x \in \mathbb{Q} \\ 0 & , \text{ if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is discontinuous at every point, and obviously all of its discontinuities are of second kind, since in every interval there are both rational and irrational numbers.

Example 4.4. *The Riemann function*

$$\mathcal{R}(x) = \begin{cases} \frac{1}{n} & , \text{ if } x = \frac{m}{n} \in \mathbb{Q} , \text{ where } \frac{m}{n} \text{ is in lowest terms, } n \in \mathbb{N} \\ 0 & , \text{ if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is continuous at any irrational number. This function is discontinuous except for $x = 0$, and all of these discontinuities are of first kind.

Proof. For any point $a \in \mathbb{R}$, any bounded neighborhood $U(a)$ of it, and any number $N \in \mathbb{N}$, $U(a)$ contains only a finite number of rational numbers $\frac{m}{n}$, $m \in \mathbb{Z}$, $n \in \mathbb{N}$, with $n < N$.

By shrinking the neighborhood, one can then assume that the denominators of all rational numbers in the neighborhood (except possibly for the point a itself if $a \in \mathbb{Q}$) are larger than N . Thus at any point $x \in \dot{U}(a)$ we have $|\mathcal{R}(x)| < \frac{1}{N}$. Now we have shown that for any point $a \in \mathbb{R} \setminus \mathbb{Q}$

$$\lim_{x \rightarrow a} \mathcal{R}(x) = 0$$

□

4.2 Properties of Continuous Functions

4.2.1 Local Properties

Theorem 4.2.1. *Let $f : E \rightarrow \mathbb{R}$ be a function that is continuous at the point $a \in E$. Then the following statements hold:*

- ¹⁰ The function $f : E \rightarrow \mathbb{R}$ is bounded in some neighborhood $U_E(a)$.
- ²⁰ If $f(a) \neq 0$, then in some neighborhood $U_E(a)$ all the values of the function have the same sign as $f(a)$.
- ³⁰ If the function $g : U_E(1) \rightarrow \mathbb{R}$ is defined in some neighborhood of a and, like f , is continuous at a , then the following functions are defined in some neighborhood of a and continuous at a :

- a) $(f + g)(x) := f(x) + g(x)$
- b) $(f \cdot g)(x) := f(x) \cdot g(x)$
- c) $\frac{f}{g}(x) := \frac{f(x)}{g(x)}$ (provided $g(a) \neq 0$)

⁴⁰ If the function $g : Y \rightarrow \mathbb{R}$ is continuous at a point $b \in Y$ and f is such that $f : E \rightarrow Y$, $f(a) = b$, and f is continuous at a , then the composite function $(g \circ f)(x)$ is defined on E and continuous at a .

Example 4.5. *All algebraic polynomials, rational functions, and the composition of a finite number of continuous functions are continuous at each point of its domain of definition.*

4.2.2 Global Properties of Continuous Functions

Theorem 4.2.2 (The Bolzano-Cauchy intermediate-value theorem). *If a function that is continuous on a closed interval assumes values with different signs at the endpoints of the interval, then there is a point in the interval where it assumes the value 0.*

$$(f \in C[a, b] \wedge f(a) \cdot f(b) < 0) \Rightarrow \exists c \in [a, b](f(c) = 0)$$

Proof. Divide the interval $[a, b]$ in half where the point of division is the point that the function does not assume the value 0 and apply the nested interval lemma, then find the limit of the sequence of these two endpoints for the nested intervals. \square

Corollary. *If the function φ is continuous on an open interval and assumes values $\varphi(a) = A$ and $\varphi(b) = B$, then for any number C between A and B , there is a point c between a and b at which $\varphi(c) = C$.*

Proof. The function $f(x) = \varphi(x) - C$ is defined and continuous on the closed interval. The intermediate-value theorem implies that $f(x) = \varphi(c) - C = 0$, since $f(a) \cdot f(b) = (A - C)(B - C) < 0$. \square

Theorem 4.2.3 (The Weierstrass maximum-value theorem). *A function that is continuous on a closed interval is bounded on that interval. Moreover there is a point in the interval where the function assumes its maximum value and a point where it assumes its minimal value.*

Proof. Let $f : E \rightarrow \mathbb{R}$ be a continuous function on the closed interval $E = [a, b]$. By the local properties of a continuous function, for any point $x \in E$ there exists a neighborhood $U(x)$ such that the function is bounded on $U_E(x)$. The set of such neighborhoods forms a covering of $[a, b]$. By the finite covering lemma, one can extract a finite system of open intervals that cover $[a, b]$. Since the function is bounded, one can find $m_k \leq f(x) \leq M_k$ where $x \in U_E(x_k)$.

$$\min\{m_1, \dots, m_n\} \leq f(x) \leq \max\{M_1, \dots, M_N\}$$

Now let $M = \sup_{x \in E} f(x)$. Set $f(x) < M$ and contradiction arises by considering the auxiliary function $\frac{1}{M - f(x)}$, which must be continuous by the local properties of the continuous functions. Similarly, one can prove that there exists a point such that $f(x_m) = m$. \square

Definition. If from every covering of a set by open intervals one can extract a finite subcovering, the set is called *compact*. According to the Heine-Borel theorem, in an Euclidean space, compact means that the set is closed (contains all of its limit points) and bounded (all of its points lie in some fixed distance).

Definition. A function $f : E \rightarrow \mathbb{R}$ is *uniformly continuous* on a set $E \subset \mathbb{R}$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x_1) - f(x_2)| < \varepsilon$ for all points $x_1, x_2 \in E$ such that $|x_1 - x_2| < \delta$. The following expression states the negation of the property of uniform continuity for a function:

$$(f : E \rightarrow \mathbb{R} \text{ is not uniformly continuous}) := (\exists \varepsilon > 0 \forall \delta > 0 \exists x_1 \in E \exists x_2 \in E (|x_1 - x_2| < \delta \wedge |f(x_1) - f(x_2)| \geq \varepsilon))$$

Example 4.6. The function $f(x) = x^2$, which is continuous on \mathbb{R} , is not uniformly continuous on \mathbb{R} .

Proof. For all the points $x'_n = \sqrt{n+1}$ and $x''_n = \sqrt{n}$, where $n \in \mathbb{N}$, $f(x'_n) - f(x''_n) = 1$, but

$$\lim_{n \rightarrow \infty} (\sqrt{n+1} - \sqrt{n}) = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n+1} + \sqrt{n}} = 0$$

so that for any $\delta > 0$ there are points x'_n and x''_n such that $|x'_n - x''_n| < \delta$, yet $f(x'_n) - f(x''_n) = 1$. \square

Example 4.7. The function $f(x) = \sin(x^2)$, which is continuous and bounded on \mathbb{R} , is not uniformly continuous on \mathbb{R} .

Theorem 4.2.4 (The Cantor-Heine Theorem on Uniform Continuity). A function that is continuous on a closed interval is uniformly continuous on that interval.

Proof. Let $f : E \rightarrow \mathbb{R}$ be a given function, $E = [a, b]$, and $f \in C(E)$. Since f is continuous at every point, we construct a system consists of δ -neighborhoods for all points $x \in E$ such that the oscillation of that neighborhood is less than ε . Let $U(x) = U^{\delta(x)}(x)$ and $V(x) = U^{\delta(x)/2}(x)$. Then the open intervals $V(x), x \in E$ cover the closed interval $[a, b]$. By the finite covering lemma we select a finite covering $V(x_1), \dots, V(x_n)$. Let $\delta = \{\frac{1}{2}\delta(x_1), \dots, \frac{1}{2}\delta(x_n)\}$. Since this system of open intervals covers E , there exists an interval $V(x_i)$ that contains x' , and $|x' - x_i| < \frac{1}{2}\delta(x_i)$. But for another point $x'' \in E$ we have

$$|x'' - x_i| \leq |x' - x''| + |x' - x_i| < \delta + \frac{1}{2}\delta(x_i) \leq \delta(x_i)$$

Consequently $x', x'' \in E$, and we have $|f(x') - f(x'')| \leq \omega(f, U_E^{\delta(x_i)}(x_i)) < \varepsilon$. \square

Proposition. A continuous mapping $f : E \rightarrow \mathbb{R}$ of a closed interval $E = [a, b]$ into \mathbb{R} is injective iff the function f is strictly monotonic on $[a, b]$.

Proof. The fact that a strictly monotonic mapping is injective is obvious. For the converse, suppose the contrary, we find three points $x_1 < x_2 < x_3$ in the interval such that $f(x_2)$ does not lie between $f(x_1)$ and $f(x_3)$. Let $f(x_2) < f(x_1) < f(x_3)$. Since f is continuous, by Bolzano-Cauchy intermediate-value theorem, there exists a point x'_1 in $[x_2, x_3]$ such that $f(x'_1) = f(x_1)$. We then have $x_1 < x'_1$, but this contradicts to the injectiveness of this function. \square

Proposition. Each strictly monotonic function $f : X \rightarrow \mathbb{R}$ defined on a numerical set $X \subset \mathbb{R}$ has an inverse $f^{-1} : Y \rightarrow \mathbb{R}$ defined on the set $Y = f(X)$ of values of f , and has the same kind of monotonicity on Y that f has on X .

Proof. The mapping $f : X \rightarrow Y$ is apparently both injective and surjective. Then $\forall y_1 \in Y \forall y_2 \in Y (f^{-1}(y_1) < f^{-1}(y_2) \Leftrightarrow y_1 < y_2)$ if f is increasing on X . The case when f is decreasing can be handled similarly. \square

Proposition. The discontinuities of a function $f : E \rightarrow \mathbb{R}$ that is monotonic on the set $E \subset \mathbb{R}$ can be only discontinuities of first kind.

Proof. Let a be the point of discontinuity. It follows that a cannot be an isolated point of E (some of its neighborhoods contain no points of E but a itself), since $U_E(a) = a$ and $f(U_E(a)) = f(a) \subset V(f(a))$ for any neighborhood of $f(a)$. Then a must be a limit point of one of two following sets: E_a^- and E_a^+ . Suppose f is non-decreasing, then the restriction $f|_{E_a^-}$ is a non-decreasing function that is bounded from above. Then the limit $\lim_{E \ni x \rightarrow a} f|_{E_a^-}(x) = \lim_{E \ni x \rightarrow a-0} f(x) = f(a-0)$ exists. The rest of this proof is analogous. \square

Corollary. If a is a point of discontinuity of a monotonic function $f : E \rightarrow \mathbb{R}$, then at least one of the limits

$$\lim_{E \ni x \rightarrow a-0} f(x) = f(a-0) \quad \lim_{E \ni x \rightarrow a+0} f(x) = f(a+0)$$

exists, and strict inequality holds in at least one of the inequalities $f(a-0) \leq f(a) \leq f(a+0)$ when f is non-decreasing and $f(a-0) \geq f(a) \geq f(a+0)$ when f is non-increasing. The function assumes no values in the open interval defined by the strict inequality. Open intervals of this kind determined by different points of discontinuity have no point in common.

Proof. The first assertion is obvious. For the second assertion, since $f(x) \leq \lim_{E \ni x \rightarrow a-0} f(x) = f(a-0)$ for $x < a \in E$, the open interval $(f(a-0), f(a))$ defined by the strict inequality $f(a-0) < f(a)$ contains no value of this function. It's similar for $f(a) < f(a+0)$. Let a_1 and a_2 be two different

points of discontinuity of f and assume $a_1 < a_2$. Since the function is non-decreasing

$$f(a_1 - 0) \leq f(a_1) \leq f(a_1 + 0) \leq f(a_2 - 0) \leq f(a_2) \leq f(a_2 + 0)$$

It follows from this that the intervals containing no values of f and corresponding to different points of discontinuity are disjoint. \square

Corollary. *The set of points of discontinuity of a monotonic function is at most countable.*

Proof. The intervals are pairwise disjoint. Since they are defined by strict inequality, one can choose a rational number in each of these intervals, so that the collection of intervals is equipollent with a subset of \mathbb{Q} . \square

Proposition (A Criterion for Continuity of a Monotonic Function). A monotonic function $f : E \rightarrow \mathbb{R}$ defined on a closed interval $E = [a, b]$ is continuous iff its set of values $f(E)$ is the closed interval with endpoints $f(a)$ and $f(b)$.

Proof. If f is a continuous monotonic function, its monotonicity implies that all the values f assumes on $[a, b]$ lie between $f(a)$ and $f(b)$. Now we consider the converse. Suppose there is a point of discontinuity c . Then one of the open intervals $(f(c - 0), f(c))$ and $(f(c), f(c + 0))$ contains no value of f , but this interval is contained between $f(a)$ and $f(b)$ because of its monotonicity. \square

Theorem 4.2.5 (The Inverse Function Theorem). *A function $f : X \rightarrow \mathbb{R}$ that is strictly monotonic on a set $X \subset \mathbb{R}$ has an inverse $f^{-1} : Y \rightarrow \mathbb{R}$ defined on the set $Y = f(X)$ of values of f . The function $f^{-1} : Y \rightarrow \mathbb{R}$ is monotonic and has the same type of monotonicity on Y that f has on X .*

Proof. This can be easily proved with our last proposition. \square

5

Differential Calculus

5.1 Differentiable Functions

5.1.1 Functions Differentiable at a Point

Definition. A function $f : E \rightarrow \mathbb{R}$ defined on a set $E \subset \mathbb{R}$ is *differentiable* at a point $a \in E$ that is a limit point of E if there exists a linear function $A \cdot (x - a)$ of the increment $x - a$ of the argument such that $f(x) - f(a)$ can be represented as

$$f(x) - f(a) = A \cdot (x - a) + o(x - a) \text{ as } x \rightarrow a, x \in E$$

In other words, a function is differentiable at a point a if the change in its values in a neighborhood of the point in question is linear up to a correction that is infinitesimal compared with the magnitude of the displacement $x - a$ from the point a .

Definition. The linear function $A \cdot (x - a)$ is called the *differential* of the function f at a . The differential of a function at a point is uniquely determined; for it follows from our former definition that

$$\lim_{E \ni x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{E \ni x \rightarrow a} \left(A + \frac{o(x - a)}{x - a} \right) = A$$

The uniqueness of A follows from the uniqueness of the limit.

Definition. The number

$$f'(a) = \lim_{E \ni x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

is called the *derivative* of the function f at a .

Differentiability of a function at a point is equivalent to the existence of its derivative at the same point.

Definition. A function $f : E \rightarrow \mathbb{R}$ defined on a set $E \subset \mathbb{R}$ is differentiable at a point $x \in E$ that is a limit point of E if

$$f(x+h) - f(x) = A(x)h + \alpha(x, h)$$

where $h \mapsto A(x)h$ is a linear function in f and $\alpha(x, h) = o(h)$ as $h \rightarrow 0, x+h \in E$.

The quantities

$$\Delta x(h) := (x+h) - x = h$$

and

$$\Delta f(x, h) := f(x+h) - f(x)$$

are called respectively the *increment of the argument* and the *increment of the function*.

Definition. The function $h \mapsto A(x)h$, which is a linear in h , is called the differential of the function $f : E \rightarrow \mathbb{R}$ at the point $x \in E$ and is denoted $df(x)$ or $Df(x)$. Thus, $df(x)(h) = A(x)h$. From the definitions above we have

$$\Delta f(x, h) - df(x)(h) = \alpha(x, h)$$

and $\alpha(x, h) = o(h)$ as $h \rightarrow 0, x+h \in E$. For this reason, we say that the differential is the (*principal*) *linear part of the increment of the function*.

$$A(x) = f'(x) = \lim_{h \rightarrow 0, x+h, x \in E} \frac{f(x+h) - f(x)}{h}$$

The derivative is frequently denoted by the symbol $\frac{df(x)}{dx}$.

5.1.2 The Tangent Line; Geometric Meaning of the Derivative and Differential

Proposition. A function $f : E \rightarrow \mathbb{R}$ that is continuous at a point $x_0 \in E$ that is a limit point of $E \subset \mathbb{R}$ admits a linear approximation iff it is differentiable at the point.

Definition. If a function $f : E \rightarrow \mathbb{R}$ is defined on a set $E \subset \mathbb{R}$ and differentiable at a point $x_0 \in E$, the line defined by $y - f(x_0) = f'(x_0)(x - x_0)$ that passing through $(x_0, f(x_0))$ and having slope $f'(x_0)$ is called the *tangent* to the graph of this function at the point $(x_0, f(x_0))$.

Definition. If the mappings $f : E \rightarrow \mathbb{R}$ and $g : E \rightarrow \mathbb{R}$ are continuous at a point $x_0 \in E$ that is a limit point of E and $f(x) - g(x) = o((x - x_0)^n)$ as $x \rightarrow x_0, x \in E$, we say that f and g have *n th order contact* at x_0 (more precisely, *contact of order at least n*). For $n = 1$ we say that the mappings f and g are tangent to each other at x_0 .

5.1.3 Examples

Example 5.1. If $f(t) = r \cos \omega t$, then $f'(t) = -r\omega \sin \omega t$.

Proof.

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{r \cos \omega(t+h) - r \cos \omega t}{h} &= r \lim_{h \rightarrow 0} \frac{-2 \sin(\frac{\omega h}{2}) \sin \omega(t + \frac{h}{2})}{h} \\ &= -r\omega \lim_{h \rightarrow 0} \sin \omega(t + \frac{h}{2}) \cdot \lim_{h \rightarrow 0} \frac{\sin(\frac{\omega h}{2})}{(\frac{\omega h}{2})} \\ &= -r\omega \sin \omega t \end{aligned}$$

□

Example 5.2. If $f(t) = r \sin \omega t$, then $f'(t) = r\omega \cos \omega t$.

Proof. The proof is analogous. □

By the definition of differentiability of a function $f : E \rightarrow \mathbb{R}$ at a point $x_0 \in E$, we have

$$f(x) - f(x_0) = A(x_0)(x - x_0) + o(x - x_0) \text{ as } x \rightarrow x_0, x \in E$$

Since the right-hand side of this equality tends to zero as $x \rightarrow x_0, x \in E$, it follows that $\lim_{E \ni x \rightarrow x_0} f(x) = f(x_0)$, so that a function that is differentiable at a point is necessarily continuous at that point. However, the converse is not always true. It's easy to see that although $f(x) = |x|$ is continuous at 0, it's not differentiable at 0.

Example 5.3. $e^{x+h} - e^x = e^x h + o(h)$ as $h \rightarrow 0$. Thus $\frac{de^x}{dx} = e^x$.

Proof.

$$e^{x+h} - e^x = e^x(e^h - 1) = e^x(h + o(h)) = e^x h + o(h)$$

Here we have used the formula $e^h - 1 = h + o(h)$ obtained in Section 3.2.4. □

Example 5.4. If $a > 0$, then $a^{x+h} - a^x = a^x(\ln a)h + o(h)$ as $h \rightarrow 0$. Thus $da^x = a^x(\ln a)dx$.

Proof.

$$\begin{aligned} a^{x+h} - a^x &= a^x(a^h - 1) = a^x(e^{h \ln a} - 1) \\ &= a^x(h \ln a + o(h \ln a)) = a^x(\ln a)h + o(h) \text{ as } h \rightarrow 0 \end{aligned}$$

□

Example 5.5. If $x \neq 0$, then $\ln|x+h| - \ln|x| = \frac{1}{x}h + o(h)$ as $h \rightarrow 0$. Thus $d \ln|x| = \frac{1}{x}dx$.

Proof.

$$\ln|x+h| - \ln|x| = \ln\left|1 + \frac{h}{x}\right|$$

For $|h| < |x|$ we have $|1 + \frac{h}{x}| = 1 + \frac{h}{x}$, and so for sufficiently small value of h we can write

$$\ln|x+h| - \ln|x| = \ln\left(1 + \frac{h}{x}\right) = \frac{h}{x} + o\left(\frac{h}{x}\right) = \frac{1}{x}h + o(h)$$

as $h \rightarrow 0$. Here we have used the relation $\ln(1+t) = t + o(t)$ as $t \rightarrow 0$, shown in Section 3.2.4. \square

Example 5.6. If $x \neq 0$ and $0 < a \neq 1$, then $\log_a|x+h| - \log_a|x| = \frac{1}{x \ln a}h + o(h)$ as $h \rightarrow 0$. Thus, $d \log_a|x| = \frac{1}{x \ln a}dx$.

Proof.

$$\begin{aligned} \log_a|x+h| - \log_a|x| &= \log_a\left|1 + \frac{h}{x}\right| = \log_a\left(1 + \frac{h}{x}\right) \\ &= \frac{1}{\ln a} \ln\left(1 + \frac{h}{x}\right) = \frac{1}{\ln a} \left(\frac{h}{x} + o\left(\frac{h}{x}\right)\right) = \frac{1}{x \ln a}h + o(h) \end{aligned}$$

\square

5.2 The Basic Rules of Differentiation

5.2.1 Differentiation and the Arithmetic Operations

Theorem 5.2.1. If two functions f and g are differentiable at x , then their sum, product, and quotient if the denominator is not zero, are differentiable.

$$\begin{aligned} (f+g)'(x) &= (f' + g')(x) \\ (f \cdot g)'(x) &= f'(x) \cdot g(x) + f(x) \cdot g'(x) \\ \left(\frac{f}{g}\right)'(x) &= \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g^2(x)} \end{aligned}$$

Proof. The theorem is easy to prove with the definition of a differentiable function and the properties of the symbol $o(\cdot)$. \square

Corollary. The derivative of a linear combination of differentiable functions equals the same linear combination of the derivatives of these functions.

Corollary. If the function f_1, \dots, f_n are differentiable at x , then

$$\begin{aligned} (f_1 \cdots f_n)'(x) &= f_1'(x)f_2(x) \cdots f_n(x) \\ &\quad + f_1(x)f_2'(x)f_3(x) \cdots f_n(x) + \cdots + f_1(x) \cdots f_{n-1}(x)f_n'(x) \end{aligned}$$

Proof. Easy to prove by induction. \square

Corollary. *Following from the relation between the derivative and the differential, the theorem can also be written in terms of differentials.*

5.2.2 Differentiation of a Composite Function(Chain Rule)

Theorem 5.2.2. *If the function $f : X \rightarrow Y \subset \mathbb{R}$ is differentiable at a point $x \in X$ and the function $g : Y \rightarrow \mathbb{R}$ is differentiable at the point $y = f(x) \in Y$, then the composite function $g \circ f : X \rightarrow \mathbb{R}$ is differentiable at x , and the differential $d(g \circ f)(x) : T\mathbb{R}(x) \rightarrow T\mathbb{R}(g(f(x)))$ of their composition equals the composition $df(y) \circ df(x)$ of their differentials*

$$df(x) : T\mathbb{R}(x) \rightarrow T\mathbb{R}(y = f(x)) \text{ and } dg(y = f(x)) : T\mathbb{R}(y) \rightarrow T\mathbb{R}(g(y))$$

Corollary. *The derivative $(g \circ f)'(x)$ of the composition of differentiable real-valued functions equals the product $g'(f(x)) \cdot f'(x)$ of the derivatives of these functions compared at the corresponding points.*

Corollary. *If the composition $(f_n \circ \cdots \circ f_1)(x)$ of differentiable functions $y_1 = f_1(x), \dots, y_n = f_n(y_{n-1})$ exists, then*

$$(f_n \circ \cdots \circ f_1)'(x) = f'_n(y_{n-1})f'_{n-1}(y_{n-2}) \cdots f'_1(x)$$

Example 5.7. *The derivative of the logarithm of the absolute value of a differentiable function is often called its logarithmic derivative.*

Since $F(x) = \ln |f(x)| = (\ln \circ || \circ f)(x)$, we have $F'(x) = (\ln |f|)'(x) = \frac{f'(x)}{f(x)}$

Example 5.8. *For a function $u(x)^{v(x)}$, where $u(x)$ and $v(x)$ are differentiable functions and $u(x) > 0$. We write $u(x)^{v(x)} = e^{v(x) \ln u(x)}$, then*

$$\frac{de^{v(x) \ln u(x)}}{dx} = u(x)^{v(x)} \cdot v'(x) \ln u(x) + v(x)u(x)^{v(x)-1} \cdot u'(x)$$

5.2.3 Differentiation of an Inverse Function

Theorem 5.2.3 (The derivative of an inverse function). *Let the functions $f : X \rightarrow Y$ and $f^{-1} : Y \rightarrow X$ be mutually inverse and continuous at points $x_0 \in X$ and $f(x_0) = y_0 \in Y$ respectively. If f is differentiable at x_0 and $f'(x_0) \neq 0$, then f^{-1} is also differentiable at the point y_0 , and*

$$(f^{-1})'(y_0) = (f'(x_0))^{-1}$$

Example 5.9. *The hyperbolic and inverse hyperbolic functions and their derivatives are:*

$$\begin{aligned}
\sinh x &= \frac{1}{2}(e^x - e^{-x}) \\
\sinh' x &= \cosh x \\
\operatorname{arsinh} y &= \ln(y + \sqrt{1 + y^2}) \\
\operatorname{arsinh}' y &= \frac{1}{\sqrt{1 + y^2}} \\
\cosh x &= \frac{1}{2}(e^x + e^{-x}) \\
\cosh' x &= \sinh x \\
\operatorname{arcosh}_- y &= \ln(y - \sqrt{y^2 - 1}) \\
\operatorname{arcosh}_+ y &= \ln(y + \sqrt{y^2 - 1}) \\
\operatorname{arcosh}'_- y &= -\frac{1}{\sqrt{y^2 - 1}}, \quad y > 1 \\
\operatorname{arcosh}'_+ y &= \frac{1}{\sqrt{y^2 - 1}}, \quad y > 1 \\
\tanh x &= \frac{\sinh x}{\cosh x} \\
\tanh' x &= \frac{1}{\cosh^2 x} \\
\operatorname{artanh} y &= \frac{1}{2} \ln \frac{1 + y}{1 - y}, \quad |y| < 1 \\
\operatorname{artanh}' y &= \frac{1}{1 - y^2}, \quad |y| < 1 \\
\coth x &= \frac{\cosh x}{\sinh x} \\
\coth' x &= -\frac{1}{\sinh^2 x} \\
\operatorname{arcoth} y &= \frac{1}{2} \ln \frac{y + 1}{y - 1}, \quad |y| > 1 \\
\operatorname{arcoth}' y &= -\frac{1}{y^2 - 1}, \quad |y| > 1
\end{aligned}$$

5.2.4 Higher-Order Derivatives

Definition. The *derivative of order n* is defined by the formula

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

The set of functions $f : E \rightarrow \mathbb{R}$ having continuous derivatives up to order n inclusive will be denoted $C^{(n)}(E, \mathbb{R})$ or $C^{(n)}(E)$. $C^{(0)}(E) = C(E)$ by $f^{(0)}(x) = f(x)$.

Example 5.10 (Leibniz's formula). Let $u(x)$ and $v(x)$ be functions having derivatives up to order n inclusive on a common set E . Then following formula of Leibniz holds for the n th derivative of their product:

$$(uv)^{(n)} = \sum_{m=0}^n \binom{n}{m} u^{(n-m)} v^{(m)}$$

5.3 The Basic Theorems of Differential Calculus

5.3.1 Fermat's Lemma and Rolle's Theorem

Definition. A point $x_0 \in E \subset \mathbb{R}$ is called a *local maximum* (resp. *local minimum*) and the value of a function $f : E \rightarrow \mathbb{R}$ at the point a *local maximum value* (resp. *local minimum value*) if there exists a neighborhood $U_E(x_0)$ of x_0 in E such that at any point $x \in U_E(x_0)$ we have $f(x) \leq f(x_0)$ (resp. $f(x) \geq f(x_0)$).

Definition. If the strict inequality $f(x) < f(x_0)$ (resp. $f(x) > f(x_0)$) holds at every point $x \in \dot{U}_E(x_0)$, the point x_0 is called *strict local maximum* (resp. *strict local minimum*) and the value of the function $f : E \rightarrow \mathbb{R}$ a *strict local maximum value* (resp. *strict local minimum value*).

Definition. The local maxima and minima are called *local extrema* and the values of the function at these points *local extreme values of the function*.

Definition. An extremum $x_0 \in E$ of the function $f : E \rightarrow \mathbb{R}$ is called an *interior extremum* if x_0 is a limit point of both sets $E_- = \{x \in E | x < x_0\}$ and $E_+ = \{x \in E | x > x_0\}$.

Lemma (Fermat). If a function $f : E \rightarrow \mathbb{R}$ is differentiable at an interior extremum, $x_0 \in E$, then its derivative at x_0 is 0: $f'(x_0) = 0$.

Proposition (Rolle's Theorem). If a function $f : [a, b] \rightarrow \mathbb{R}$ is continuous on the closed interval $[a, b]$ and differentiable on (a, b) and $f(a) = f(b)$, then there exists a point $\xi \in (a, b)$ such that $f'(\xi) = 0$.

5.3.2 The Theorems of Lagrange and Cauchy on Finite Increments

Theorem 5.3.1 (Lagrange's Finite-increment Theorem). If a function $f : [a, b] \rightarrow \mathbb{R}$ is continuous on a closed interval $[a, b]$ and differentiable on the open interval (a, b) , there exists a point $\xi \in [a, b]$ such that

$$f(b) - f(a) = f'(\xi)(b - a)$$

Proof. Consider the auxiliary function

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

and apply Rolle's theorem to it. □

Corollaries of Lagrange's Theorem

Corollary (Criterion for monotonicity of a function). *If the derivative of a function is nonnegative (resp. positive) at every point of an open interval, then the function is nondecreasing (resp. increasing) on that interval.*

An analogous assertion can be made about the nonincreasing (resp. decreasing) nature of a function with a nonpositive (resp. negative) derivative.

Corollary (Criterion for a function to be constant). *A function that is continuous on a closed interval is constant on it iff its derivative equals zero at every point of the interval $[a, b]$ (or only on (a, b)).*

If the derivative of two functions are equal on some interval, then the difference between them is constant.

Proposition (Cauchy's Finite-increment Theorem). Let $x = x(t)$ and $y = y(t)$ be functions that are continuous on a closed interval $[\alpha, \beta]$ and differentiable on the open interval (α, β) .

Then there exists a point $\tau \in [\alpha, \beta]$ such that

$$x'(\tau)(y(\beta) - y(\alpha)) = y'(\tau)(x(\beta) - x(\alpha))$$

If in addition $x'(t) \neq 0$ for each t in (α, β) , then $x(\alpha) \neq x(\beta)$ and we have the equality

$$\frac{y(\beta) - y(\alpha)}{x(\beta) - x(\alpha)} = \frac{y'(\tau)}{x'(\tau)}$$

5.3.3 Taylor's Formula

Definition. The algebraic polynomial given by

$$P_n(x_0, x) = P_n(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

is the *Taylor polynomial of order n of $f(x)$ at x_0* .

The *n th remainder term in Taylor's formula*, or the *remainder*, is the value

$$f(x) - P_n(x_0, x) = r_n(x_0, x)$$

Then

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r_n(x_0, x)$$

When $x_0 = 0$, this is often called *MacLaurin's formula*.

Theorem 5.3.2. *If the function f is continuous on the closed interval with end-points x_0 and x along with its first n derivatives, and it has a derivative of order $n + 1$ at the interior points of this interval, then for any function φ that is continuous on this closed interval and has a nonzero derivative at its interior points, there exists a point ξ between x_0 and x such that*

$$r_n(x_0, x) = \frac{\varphi(x) - \varphi(x_0)}{\varphi'(\xi)n!} f^{(n+1)}(\xi)(x - \xi)^n$$

Proof. Consider the auxiliary function

$$F(t) = f(x) - P_n(t, x)$$

Find $F'(t)$. Apply Cauchy's theorem to the pair of functions $F(t)$, $\varphi(t)$ and find a point ξ . Substitute the expression for $F'(\xi)$ here and the theorem is proved with the fact that $F(x) - F(x_0) = 0 - F(x_0) = -r_n(x_0, x)$. \square

Corollary (Cauchy's Formula for the remainder term). *Setting $\varphi(t) = x - t$, we obtain*

$$r_n(x_0, x) = \frac{1}{n!} f^{(n+1)}(\xi)(x - \xi)^n(x - x_0)$$

Corollary (The Lagrange form of the remainder). *Setting $\varphi(t) = (x - t)^{n+1}$, we have*

$$r_n(x_0, x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x - x_0)^{n+1}$$

Theorem 5.3.3.

$$(1 + x)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n$$

Definition. If the function has derivatives of all orders $n \in \mathbb{N}$ at a point x_0 , the series

$$f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \cdots$$

is called the *Taylor's series* of f at the point x_0 .

it should not be thought that the Taylor series of an infinitely differentiable function converges in some neighborhood of x_0 .

It should also not be thought that if the Taylor series converges, it necessarily converges to the function that generated it – it converges to the

function that generated it only when the generating function belongs to the class of *analytic functions*. Cauchy gave an example of a nonanalytic function:

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

The Taylor series in this case has all its terms equal to 0 and hence its sum is 0, while $f(x) \neq 0$ if $x \neq 0$.

Proposition. if there exists a polynomial $P_n(x_0, x) = c_0 + c_1(x - x_0) + \cdots + c_n(x - x_0)^n$ satisfying the following condition:

$$f(x) = c_0 + c_1(x - x_0) + \cdots + c_n(x - x_0)^n + o((x - x_0)^n) \text{ as } x \rightarrow x_0, x \in E$$

then the polynomial is unique.

Proposition (The Local Taylor Formula). Let E be a closed interval having $x_0 \in \mathbb{R}$ as an endpoint. If the function $f : E \rightarrow \mathbb{R}$ has derivatives $f'(x_0), \dots, f^{(n)}(x_0)$ up to order n inclusive at the point x_0 , then the following representation holds:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n) \text{ as } x \rightarrow x_0, x \in E$$

Lemma. If a function $\varphi : E \rightarrow \mathbb{R}$, defined on a closed interval E with endpoint x_0 , is such that it has derivatives up to order n inclusive at x_0 and $\varphi(x_0) = \varphi'(x_0) = \cdots = \varphi^{(n)}(x_0) = 0$, then $\varphi(x) = o((x - x_0)^n)$ as $x \rightarrow x_0, x \in E$.

Proof. content... □

The relation is called the *local Taylor formula* since the form of the remainder term given in it (the so-called *Peano form*)

$$r_n(x_0, x) = o((x - x_0)^n)$$

makes it possible to draw inferences only about the asymptotic nature of the connection between the Taylor polynomial and the function as $x \rightarrow x_0, x \in E$.

Example 5.11 (Taylor's formula with the Lagrange form of the remainder). If f has a derivative of order $n + 1$ on the open interval with endpoints x_0 and x , then

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

where ξ is a point between x_0 and x . This is a generalization of Lagrange's mean-value theorem (finite-increment theorem), to which it reduces when $n = 0$.

Example 5.12 (Taylor's formula with the Peano form of the remainder). If f has derivative of orders up to $n \geq 1$ inclusive at the point x_0 , then

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n)$$

This is a generalization of the definition of differentiability of a function at a point, to which it reduces when $n = 1$.

Now we write the following table of asymptotic formulas as $x \rightarrow 0$:

$$\begin{aligned} e^x &= 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + O(x^{n+1}) \\ \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + \frac{(-1)^n}{2n!}x^{2n} + O(x^{2n+2}) \\ \sin x &= \frac{1}{1!}x - \frac{1}{3!}x^3 + \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + O(x^{2n+3}) \\ \cosh x &= 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots + \frac{1}{2n!}x^{2n} + O(x^{2n+2}) \\ \sinh x &= \frac{1}{1!}x + \frac{1}{3!}x^3 + \cdots + \frac{1}{(2n+1)!}x^{2n+1} + O(x^{2n+3}) \\ \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + \frac{(-1)^{n-1}}{n}x^n + O(x^{n+1}) \\ (1+x)^\alpha &= 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots \\ &\quad + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + O(x^{n+1}) \end{aligned}$$

5.4 The Study of Functions Using the Methods of Differential Calculus

5.4.1 Conditions for a Function to be Monotonic

Proposition. The following relations hold between the monotonicity properties of a function $f : E \rightarrow \mathbb{R}$ that is differentiable on an open interval $(a, b) = E$ and the sign (positivity) of its derivative f' on that interval:

$f'(x) > 0 \Rightarrow f$ is increasing	$\Rightarrow f'(x) \geq 0$
$f'(x) \geq 0 \Rightarrow f$ is nondecreasing	$\Rightarrow f'(x) \geq 0$
$f'(x) \equiv 0 \Rightarrow f \equiv \text{const.}$	$\Rightarrow f'(x) \equiv 0$
$f'(x) \leq 0 \Rightarrow f$ is nonincreasing	$\Rightarrow f'(x) \leq 0$
$f'(x) < 0 \Rightarrow f$ is decreasing	$\Rightarrow f'(x) \leq 0$

Proof. The left-hand column of implications is already known from Lagrange's theorem. The right-hand column of implications can be obtained immediately from the definition of the derivative. \square

5.4.2 Conditions for an Interior Extremum of a Function

5.4.3 Conditions for a Function to be Convex

5.4.4 L'Hôpital's Rule

5.4.5 Constructing the Graph of a Function

5.5 Complex Numbers and Elementary Functions

5.5.1 Complex Numbers

5.5.2 Convergence in \mathbb{C} and Series with Complex Terms

5.5.3 Euler's Formula and the Connections Among the Elementary Functions

5.5.4 Power Series Representation of a Function. Analyticity

5.5.5 Algebraic Closedness of the Field \mathbb{C} of Complex Numbers

5.6 Primitives

5.6.1 The Primitive and the Indefinite Integral

5.6.2 The Basic General Methods of Finding a Primitive

5.6.3 Primitives of Rational Functions

5.6.4 Primitives of the Form $\int R(\cos x, \sin x)dx$

5.6.5 Primitives of the Form $\int R(x, y(x))dx$

6

Integration

6.1 Definition of the Integral and Description of the Set of Integrable Functions

6.1.1 Definition of the Riemann Integral

6.1.2 The Set of Integrable Functions

6.2 Linearity, Additivity and Monotonicity of the Integral

6.2.1 The Integral as a Linear Function on the Space $\mathcal{R}[a,b]$

6.2.2 The Integral as an Additive Function of the Interval of Integration

6.2.3 Estimation of the Integral, Monotonicity of the Integral, and the Mean-Value Theorem

6.3 The Integral and the Derivative

6.3.1 The Integral and the Primitive

6.3.2 The Newton-Leibniz Formula

6.3.3 Integration by Parts in the Definite Integral and Taylor's Formula

6.3.4 Change of Variable in an Integral

6.3.5 Examples

6.4 Some Applications of Integration

6.4.1 Additive Interval Functions and the Integral

6.4.2 Arc Length

6.4.3 The Area of a Curvilinear Trapezoid

6.4.4 Volume of a Solid Revolution

6.4.5 Work and Energy

6.5 Improper Integrals

7

Functions of Several Variables: Their Limits and Continuity

7.1 The Space \mathbb{R}^m and the Most Important Classes of Its Subsets

7.1.1 The Set \mathbb{R}^m and the Distance in It

Definition. The function $d : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$

$$d(x_1, x_2) = \sqrt{\sum_{i=1}^m (x_1^i - x_2^i)^2}$$

is called the *metric* or *distance* of the set \mathbb{R}^m . It has the following properties:

- a) $d(x_1, x_2) \geq 0$;
- b) $(d(x_1, x_2) = 0) \Leftrightarrow (x_1 = x_2)$;
- c) $d(x_1, x_2) = d(x_2, x_1)$;
- d) $d(x_1, x_3) \leq d(x_1, x_2) + d(x_2, x_3)$.

The last inequality (called, because of geometric analogies, the *triangle inequality*) is a special case of Minkowski's inequality (see Section 5.4.2).

Any set with a fixed metric on it is called a *metric space*.

For $i \in \{1, \dots, m\}$ in \mathbb{R}^m

$$|x_1^i - x_2^i| \leq d(x_1, x_2) \leq \sqrt{m} \max_{1 \leq i \leq m} |x_1^i - x_2^i|$$

that is, the distance between the points $x_1, x_2 \in \mathbb{R}^m$ is small iff the corresponding coordinates of these points are close together.

7.1.2 Open and Closed Sets in \mathbb{R}^m

Definition. For $\delta > 0$ the set

$$B(a, \delta) = \{x \in \mathbb{R}^m \mid d(a, x) < \delta\}$$

is called the *ball with center $a \in \mathbb{R}^m$ of radius δ* or the δ -neighborhood of the point $a \in \mathbb{R}^m$.

Definition. A set $G \subset \mathbb{R}^m$ is open in \mathbb{R}^m if for every point $x \in G$ there is a ball $B(x, \delta)$ such that $B(x, \delta) \subset G$.

Example 7.1. \mathbb{R}^m and \emptyset are open sets in \mathbb{R}^m .

Example 7.2. A ball $B(a, r)$ is open in \mathbb{R}^m .

Proof. If $x \in B(a, r)$, that is, $d(a, x) < r$, then for $0 < \delta < r - d(a, x)$, we have $B(x, \delta) \subset B(a, r)$, since

$$\begin{aligned} (\xi \in B(x, \delta)) &\Rightarrow (d(x, \xi) < \delta) \\ &\Rightarrow (d(a, \xi) \leq d(a, x) + d(x, \xi) < d(a, x) + r - d(a, x) = r) \end{aligned}$$

□

Example 7.3. A set $G = \{x \in \mathbb{R}^m \mid d(a, x) > r\}$ is open.

Definition. The set $F \subset \mathbb{R}^m$ is *closed* in \mathbb{R}^m if its complement $G = \mathbb{R}^m \setminus F$ is open in \mathbb{R}^m . For example, the set $\bar{B}(a, r) = \{x \in \mathbb{R}^m \mid d(a, x) \leq r\}$, $r \geq 0$ is closed, and is called the *closed ball with center a of radius r* .

Proposition. a) The union $\bigcup_{\alpha \in A} G_\alpha$ of the sets of any system $\{G_\alpha, \alpha \in A\}$ of open sets in \mathbb{R}^m is open.

b) The intersection $\bigcap_{i=1}^n G_i$ of a finite number of open sets in \mathbb{R}^m is an open set.

a') The intersection $\bigcap_{\alpha \in A} F_\alpha$ of the sets of any system $\{F_\alpha, \alpha \in A\}$ of closed sets F_α in \mathbb{R}^m is closed.

b') The union $\bigcup_{i=1}^n F_i$ of a finite number of closed sets in \mathbb{R}^m is a closed set.

Proof. content... □

Example 7.4. The set $S(a, r) = \{x \in \mathbb{R}^m \mid d(a, x) = r\}$, $r \geq 0$, is called the *sphere of radius r with center $a \in \mathbb{R}^m$* . By the proposition just proved, the sphere $S(a, r)$ is closed.

Definition. An open set in \mathbb{R}^m containing a given point is called a *neighborhood* of that point in \mathbb{R}^m .

7.1. THE SPACE \mathbb{R}^M AND THE MOST IMPORTANT CLASSES OF ITS SUBSETS 79

Definition. In relation to a set $E \subset \mathbb{R}^m$ a point $x \in \mathbb{R}^m$ is
an interior point if some neighborhood of it is contained in E ;
an exterior point if it is an interior point of the complement of E in \mathbb{R}^m ;
a boundary point if it is neither an interior point nor an exterior point.

Example 7.5. The sphere $S(a, r)$, $r > 0$ is the set of boundary points of both the open ball $B(a, r)$ and the closed ball $\bar{B}(a, r)$.

Example 7.6. A point $a \in \mathbb{R}^m$ is a boundary point of the set $\mathbb{R}^m \setminus a$, which has no exterior points.

Example 7.7. All points of the sphere $S(a, r)$ are boundary points of it regarded as a subset of \mathbb{R}^m , the sphere has no interior points.

Definition. A point $a \in \mathbb{R}^m$ is a limit point of the set $E \subset \mathbb{R}^m$ if for any neighborhood $O(a)$ of a the intersection $E \cap O(a)$ is an infinite set.

Definition. The union of a set E and all its limit points in \mathbb{R}^m is the *closure* of E in \mathbb{R}^m and usually denoted \bar{E} .

Example 7.8. The set $\bar{B}(a, r) = B(a, r) \cup S(a, r)$ is the set of limit points of the open ball $B(a, r)$; that is why $\bar{B}(a, r)$ is called a closed ball.

Example 7.9. $\bar{S}(a, r) = S(a, r)$

Proposition. (F is closed in \mathbb{R}^m) \Leftrightarrow ($F = \bar{F}$ in \mathbb{R}^m). That is, F is closed in \mathbb{R}^m iff it contains all its limit points.

Proof. content... □

7.1.3 Compact Sets in \mathbb{R}^m

Definition. A set $K \subset \mathbb{R}^m$ is *compact* if from every covering of K by sets that are open in \mathbb{R}^m one can extract a finite covering.

Example 7.10. A closed interval $[a, b] \subset \mathbb{R}^1$ is compact by the finite covering lemma (Heine-Borel Theorem).

Example 7.11. A generalization to \mathbb{R}^m of the concept of a closed interval is the set

$$I = \{x \in \mathbb{R}^m | a^i \leq x^i \leq b^i, i = 1, \dots, m\}$$

which is called an m -dimensional interval, or an m -dimensional block or an m -dimensional parallelepiped. I is compact in \mathbb{R}^m .

Proof. content... □

Proposition. If K is a compact space in \mathbb{R}^m , then

- a) K is closed in \mathbb{R}^m ;
- b) any closed subset of \mathbb{R}^m contained in K is itself compact.

Proof. content... □

Definition. The *diameter* of a set $E \subset \mathbb{R}^m$ is the quantity

$$d(E) := \sup_{x_1, x_2 \in E} d(x_1, x_2)$$

Definition. A set $E \subset \mathbb{R}^m$ is *bounded* if its diameter is finite.

Proposition. If K is compact set in \mathbb{R}^m , then K is a bounded subset of \mathbb{R}^m .

Proof. content... □

Proposition. The set $K \subset \mathbb{R}^m$ is compact iff K is closed and bounded in \mathbb{R}^m .

Proof. content... □

7.2 Limits and Continuity of Functions of Several Variables

7.2.1 The Limit of a Function

Definition. A point $A \in \mathbb{R}^n$ is the *limit* of the mapping $f : X \rightarrow \mathbb{R}^n$ over a base \mathcal{B} in X if for every neighborhood $V(A)$ of the point there exists an element $B \in \mathcal{B}$ of the base whose image $f(B)$ is contained in $V(A)$.

$$(\lim_{\mathcal{B}} f(x) = A) := (\forall V(A) \exists B \in \mathcal{B} (f(B) \subset V(A)))$$

Or, with explicit use of the metric in \mathbb{R}^n

$$(\lim_{\mathcal{B}} f(x) = A \in \mathbb{R}^n) := (\forall \varepsilon > 0 \exists B \in \mathcal{B} \forall x \in B (d(f(x), A) < \varepsilon))$$

$$(\lim_{\mathcal{B}} f(x) = A \in \mathbb{R}^n) := (\lim_{\mathcal{B}} (d(f(x), A)) = 0)$$

If the limit $A = (A^1, \dots, A^n)$ and $y = (y^1, \dots, y^n)$ where $f^i(x) = y^i$, we have

$$|y^i - A^i| \leq d(y, A) \leq \sqrt{n} \max_{1 \leq i \leq n} |y^i - A^i|$$

from which one can see that

$$\lim_{\mathcal{B}} f(x) = A \Leftrightarrow \lim_{\mathcal{B}} f^i(x) = A^i$$

that is, convergence in \mathbb{R}^n is coordinatewise.

Definition. A mapping $f : X \rightarrow \mathbb{R}^n$ is *bounded* if the set $f(X) \subset \mathbb{R}^n$ is bounded in \mathbb{R}^n .

Definition. Let \mathcal{B} be a base in X . A mapping $f : X \rightarrow \mathbb{R}^n$ is *ultimately bounded* over the base \mathcal{B} if there exists an element $B \in \mathcal{B}$ on which f is bounded.

As one can easily verified, a function can have at most one limit over a given base. If its limit exists over a base, the function is ultimately bounded over that base.

Definition. A sequence $\{y_k\}$, $k \in \mathbb{N}$, of points $y_k \in \mathbb{R}^n$ is *fundamental* (a *Cauchy sequence*) if for every $\varepsilon > 0$ there exists a number $N \in \mathbb{N}$ such that $d(y_{k_1}, y_{k_2}) < \varepsilon$ for all $y_{k_1}, y_{k_2} > N$. One can see that $\{y_k\}$ is a Cauchy sequence iff $\{y_k^i\}$, $i = 1, \dots, n$ is a Cauchy sequence.

In other words, the Cauchy criterion is also valid in \mathbb{R}^n .

Definition. The *oscillation* of a function $f : X \rightarrow \mathbb{R}^n$ on a set $E \subset X$ is the quantity

$$\omega(f, E) := d(f(E))$$

where $d(f(E))$ is the diameter of $f(E)$.

Theorem 7.2.1. Let X be a set and \mathcal{B} a base in X . A function $f : X \rightarrow \mathbb{R}^n$ has a limit over the base \mathcal{B} iff for every $\varepsilon > 0$ there exists an element $B \in \mathcal{B}$ of the base on which the oscillation of the function is less than ε .

$$\exists \lim_{\mathcal{B}} f(x) \Leftrightarrow \forall \varepsilon > 0 \exists B \in \mathcal{B} (\omega(f, B) < \varepsilon)$$

Theorem 7.2.2. Let Y be a set, \mathcal{B}_Y a base in Y , and $g : Y \rightarrow \mathbb{R}^n$ a mapping having a limit over the base \mathcal{B}_Y .

Let X be a set, \mathcal{B}_X a base in X , and $f : X \rightarrow Y$ a mapping of X into Y such that for each $B_Y \in \mathcal{B}_Y$ there exists $B_X \in \mathcal{B}_X$ such that the image $f(B_X)$ is contained in B_Y .

Under these conditions the composition $g \circ f : X \rightarrow \mathbb{R}^n$ of the mapping f and g is defined and has a limit over the base \mathcal{B}_X , and

$$\lim_{\mathcal{B}_X} (g \circ f)(x) = \lim_{\mathcal{B}_Y} g(y)$$

All notations about base and neighborhoods remain the same.

Example 7.12. Let $x \mapsto \pi^i(x)$ be the mapping $\pi^i : \mathbb{R}^m \rightarrow \mathbb{R}$ and $\pi^i(x) = x^i$. This function does not tend to any finite value nor to infinity as $x \rightarrow \infty$ if $m > 1$.

On the other hand

$$f(x) = \sum_{i=1}^m (\pi^i(x))^2 \rightarrow \infty \text{ as } x \rightarrow \infty$$

Example 7.13. Let the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be defined at the point $(x, y) \in \mathbb{R}^2$ as follows:

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{if } x^2 + y^2 \neq 0, \\ 0, & \text{if } x^2 + y^2 = 0. \end{cases}$$

Then $f(0, y) = f(x, 0) = 0$, while $f(x, x) = \frac{1}{2}$ for $x \neq 0$. Hence this function has no limit as $(x, y) \rightarrow (0, 0)$. We can see that the limit of a function of several variables cannot be found by computing successively the limits with respect to each of its coordinates.

Often change the order of coordinates in finding the corresponding limit will yield different results, such as $f(x, y) = \begin{cases} \frac{x^2-y^2}{x^2+y^2}, & \text{if } x^2 + y^2 \neq 0, \\ 0, & \text{if } x^2 + y^2 = 0. \end{cases}$ when $x, y \rightarrow 0$ (iteratively).

7.2.2 Continuity of a Function of Several Variables and Properties of Continuous Functions

Let E be a subset of \mathbb{R}^m and $f : E \rightarrow \mathbb{R}^n$ a function defined on E with values in \mathbb{R}^n .

Definition. The function $f : E \rightarrow \mathbb{R}^n$ is *continuous* at $a \in E$ if for every neighborhood $V(f(a))$ there exists a neighborhood $U_E(a)$ and $f(U_E(a)) \subset V(f(a))$.

The mapping $f : E \rightarrow \mathbb{R}^n$ defined by the relation

$$\begin{aligned} (x^1, \dots, x^m) = x &\xrightarrow{f} y = (y^1, \dots, y^n) = \\ &= (f^1(x^1, \dots, x^m), \dots, f^n(x^1, \dots, x^m)) \end{aligned}$$

is continuous at a point iff each of the functions $y^i = f^i(x^1, \dots, x^m)$ is continuous at that point.

We defined a *path* in \mathbb{R}^n to be a mapping $f : I \rightarrow \mathbb{R}^n$ of an interval $I \subset \mathbb{R}$ defined by continuous functions $f^1(x), \dots, f^n(x)$ in the form

$$x \mapsto y = (y^1, \dots, y^n) = (f^1(x), \dots, f^n(x))$$

Thus we can now say that a *path* in \mathbb{R}^n is a continuous mapping of an interval $I \subset \mathbb{R}$ of the real line into \mathbb{R}^n .

Let E be a subset of \mathbb{R}^m , $a \in E$, and $B_E(a, r) = E \cap B(a, r)$.

Definition. The *oscillation* of the function $f : E \rightarrow \mathbb{R}^r$ at the point $a \in E$ is the quantity

$$\omega(f, a) := \lim_{r \rightarrow +0} \omega(f, B_E(a, r))$$

Local Properties of Continuous Functions

a) A mapping $f : E \rightarrow \mathbb{R}^n$ of a set $E \subset \mathbb{R}^m$ is continuous at a point $a \in E$ iff $\omega(f, a) = 0$.

b) A mapping $f : E \rightarrow \mathbb{R}^n$ that is continuous at $a \in E$ is bounded in some neighborhood $U_E(a)$.

c) If the mapping $g : Y \rightarrow \mathbb{R}^k$ of the set $Y \subset \mathbb{R}^n$ is continuous at a point $y_0 \in Y$ and the mapping $f : X \rightarrow Y$ of the set $X \subset \mathbb{R}^m$ is continuous at a point $x_0 \in X$ and $f(x_0) = y_0$, then the mapping $(g \circ f)(x) : X \rightarrow \mathbb{R}^k$ is defined, and it is continuous at $x_0 \in X$.

Real-valued functions possess, in addition, the following properties.

d) If the function $f : E \rightarrow \mathbb{R}$ is continuous at the point $a \in E$ and $f(a) > 0$ (or $f(a) < 0$), there exists a neighborhood $U_E(a)$ such that $f(x) > 0$ (resp. $f(x) < 0$) for all $x \in U_E(a)$.

e) If the functions $f : E \rightarrow \mathbb{R}$ and $g : E \rightarrow \mathbb{R}$ are continuous at $a \in E$, then any linear combination of them $(\alpha f + \beta g) : E \rightarrow \mathbb{R}$ where $\alpha, \beta \in \mathbb{R}$, their product, and, if $g(x) \neq 0$ on E , their quotient, are defined on E and continuous at a .

The function is *continuous* on the set E if it is continuous at each point of the set. The set of functions $f : E \rightarrow \mathbb{R}^n$ that are continuous on E will be denoted $C(E, \mathbb{R}^n)$, or simply $C(E)$, if the range of values of the functions is unambiguously determined from the context.

Example 7.14. The functions, $\pi^i : \mathbb{R}^m \rightarrow \mathbb{R}$, defined before are obviously continuous at each point $a \in \mathbb{R}^m$.

Example 7.15. The function $f(x, y)$ of Example 7.13 is continuous at any point of the space \mathbb{R}^2 except $(0, 0)$, but the function is continuous in either of its two variables for each fixed value of the other variable.

Definition. A mapping $f : E \rightarrow \mathbb{R}^n$ of a set $E \subset \mathbb{R}^m$ into \mathbb{R}^n is *uniformly continuous* on E if for every $\varepsilon > 0$ there is a number $\delta > 0$ such that $d(f(x_1), f(x_2)) < \varepsilon$ for any points $x_1, x_2 \in E$ such that $d(x_1, x_2) < \delta$.

Definition. A set $E \subset \mathbb{R}^m$ is *pathwise connected* if for any pair of its points x_0, x_1 , there exists a path $\Gamma : I \rightarrow E$ with support in E and endpoints at these points.

In other words, it is possible to go from any point $x_0 \in E$ to any other point $x_1 \in E$ without leaving E .

Definition. A *domain* in \mathbb{R}^m is an open connected set.

Example 7.16. An open ball $B(a, r), r > 0$, in \mathbb{R}^m is a domain.

Proof. See Pg.424. □

Example 7.17. A circle is a connected set, but it is not a domain in \mathbb{R}^2 , since it is not open in \mathbb{R}^2 .

Global Properties of Continuous Functions

- a)** If a mapping $f : K \rightarrow \mathbb{R}^n$ is continuous on a compact set $K \subset \mathbb{R}^m$, then it is uniformly continuous on K .
- b)** If a mapping $f : K \rightarrow \mathbb{R}^n$ is continuous on a compact set $K \subset \mathbb{R}^m$, then it is bounded on K .
- c)** If a function $f : K \rightarrow \mathbb{R}$ is continuous on a compact set $K \subset \mathbb{R}^m$, then it assumes its maximal and minimal values at some points of K .
- d)** If a function $f : E \rightarrow \mathbb{R}$ is continuous on a connected set E and assumes the value $f(a) = A$ and $f(b) = B$ at points $a, b \in E$, then for any C between A and B , there is a point $c \in E$ at which $f(c) = C$.

All the properties, except **d)**, can be proved similarly as we did for single variable functions. We now prove property **d)**.

Proof. content...

□

Example 7.18. The sphere $S(0, r)$ defined in \mathbb{R}^m by the equation

$$(x^1)^2 + \cdots + (x^m)^2 = r^2$$

is a compact set.

Proof. content...

□

Example 7.19. The open set $\mathbb{R}^m \setminus S(0, r)$ for $r > 0$ is not a domain, since it is not connected.

Proof. content...

□

8

The Differential Calculus of Functions of Several Variables

8.1 The Linear Structure on \mathbb{R}^m

8.1.1 \mathbb{R}^m as a Vector Space

8.1.2 Linear Transformations $L : \mathbb{R}^m \rightarrow \mathbb{R}^n$

8.1.3 The Norm in \mathbb{R}^m

8.1.4 The Euclidean Structure on \mathbb{R}^m

8.2 The Differential of a Function of Several Variables

8.2.1 Differentiability and the Differential of a Function at a Point

8.2.2 The Differential and Partial Derivatives of a Real-Valued Function

8.2.3 Coordinate Representation of the Differential of a Mapping. The Jacob Matrix

8.2.4 Continuity, Partial Derivatives, and Differentiability of a Function at a Point

8.3 The Basic Laws of Differentiation

8.3.1 Linearity of the Operation of Differentiation

8.3.2 Differentiation of a Composition of Mappings (Chain Rule)

8.3.3 Differentiation of an Inverse Mapping

8.4 The Basic Facts of Differential Calculus of Real-Valued Functions of Several Variables

8.4.1 The Mean-Value Theorem

8.4.2 A Sufficient Condition for Differentiability of a Func