Lecture Notes for Math 104 — Introduction to Analysis

(following the structure of Rudin's *Principles of Mathematical Analysis*)

Rui Wang
University of California, Berkeley

Introduction to Analysis
Rui Wang
Draft

These notes are based on lectures given in Math 104 – Introduction to Analysis, UC Berkeley. They have been continuously developed since 2019, and are not intended as a substitute for Rudin, but as a companion resource with additional explanations and examples. We sincerely thank the students of past courses for their valuable feedback on these notes.

Introduction to Analysis
Rui Wang
Draft

© 2019–2025 Rui Wang. All rights reserved.

Contents

Introduction to Analysis
Rui Wang
Draft

Introduction to Analysis
Rui Wang
Draft

Introduction to Analysis
Rui Wang
Draft

Introduction to Analysis
Rui Wang
Draft

CHAPTER 1

The number systems

In this course we will study sequences, series, and functions, with the goal of understanding their key properties: convergence, continuity, differentiability, and integrability.

All of these notions depend on precise definitions of numbers and their properties. Without a rigorous foundation for the number systems we use, concepts such as "limit" or "continuity" would remain vague.

If you have studied abstract algebra, you may recall the formal construction of the field of rational numbers as the field of fractions of the integral domain of integers, and the field of complex numbers as an algebraic extension of the real numbers. However, a crucial gap remains: the construction and definition of the real number system itself. Filling this gap is our first task in analysis.

It is perfectly fine if you do not have an algebra background. We will explicitly state the basic requirements for algebraic structures such as groups, rings, and fields, but in this course we will restrict attention to the concrete examples $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

1. The ordered field of rational numbers $\mathbb{Q}$

1.1. The integral domain of integers $\mathbb{Z}$. We use $\mathbb{Z}$ to denote the set of integers:

$$\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}.$$

On $\mathbb{Z}$ we can perform addition $+$ and subtraction $-$. Subtraction is introduced as the inverse operation of addition.

In the language of abstract algebra, $\mathbb{Z}$ forms an abelian group with respect to addition. Concretely, this means

(1) Addition $+$ is associative: $(a + b) + c = a + (b + c)$;
(2) Addition $+$ has an identity element, which is zero: $a + 0 = 0 + a = a$;
(3) Every element has an inverse element, its negation: $a + (-a) = 0 = (-a) + a$;
(4) Addition $+$ is commutative: $a + b = b + a$;

for all $a, b, c \in \mathbb{Z}$.

The first three properties ensure that $(\mathbb{Z}, +)$ forms a group. Adding the fourth property gives that $(\mathbb{Z}, +)$ is an abelian group (also called a commutative group).

Since additive inverses exist, subtraction can be defined by

$$a - b := a + (-b).$$

On $\mathbb{Z}$ there is another natural binary operation, multiplication $\cdot$. The structure $(\mathbb{Z}, \cdot)$ is not a group: although multiplication is associative, commutative and has an identity element 1, any integer other than ± 1 fails to have a multiplicative inverse in Z.

Moreover, notice that multiplication and addition are compatible through the distributive laws: for all $a, b, c \in \mathbb{Z}$,

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (b + c) \cdot a = b \cdot a + c \cdot a.$$

Introduction to Analysis
Rui Wang
Draft

Put all these properties together, i.e.,

(1) $(\mathbb{Z}, +)$ is an abelian group;
(2) $(\mathbb{Z}, \cdot)$ is associative;
(3) $(\mathbb{Z}, \cdot)$ is commutative;
(4) $(\mathbb{Z}, \cdot)$ has the identity element 1;
(5) multiplication and addition satisfy the distributive laws,

$(\mathbb{Z}, +, \cdot)$ is a commutative (from (3)), unital (from (4)) ring (from (1)(2)(5)).

The ring $\mathbb{Z}$ has another property that the condition $a \cdot b = 0$ implies either $a = 0$ or $b = 0$. A commutative, unital ring with such a property is called an integral domain. As a conclusion,

THEOREM 1.1. $(\mathbb{Z}, +, \cdot)$ *is an integral domain.*

1.2. The field of rational numbers $\mathbb{Q}$. In abstract algebra you may have seen the general construction of the field of fractions of an integral domain. Concretely, this is exactly how the rational numbers $\mathbb{Q}$ are obtained from the integers $\mathbb{Z}$.

A rational number can be written as

$$\frac{m}{n}, \quad \text{where } m, n \in \mathbb{Z} \text{ and } n \neq 0.$$

By cancelling common factors, every rational number is either 0 or can be uniquely written as

$$\frac{p}{q}, \quad \text{where } p, q \in \mathbb{Z},\ q > 0,\ \text{and } \gcd(p, q) = 1.$$

It is a good exercise to check the following fact that $(\mathbb{Q}, +, \cdot)$ is a commutative, unital ring and actually also an integral domain.

Moreover, every nonzero rational number has a reciprocal, which is exactly its multiplicative inverse. This property is precisely what distinguishes a field from a commutative unital ring.

We can therefore state:

THEOREM 1.2. $(\mathbb{Q}, +, \cdot)$ *is a field.*

REMARK 1.3. It is a standard exercise in abstract algebra to show that every field is an integral domain. We will not prove this here, but interested readers may consult algebra texts for details.

The field properties allow us to perform addition, subtraction, multiplication, and division (by any nonzero element). These operations are fundamental for conducting analysis.

We will give a rigorous construction of $\mathbb{R}$ later and show that $\mathbb{R}$ is also a field with respect to addition and multiplication. The same holds for $\mathbb{C}$.

1.3. Totally ordered sets. To construct $\mathbb{R}$ from $\mathbb{Q}$, we need to equip $\mathbb{Q}$ with an additional structure: a total order.

DEFINITION 1.4 (Partial order). Suppose X is a set and $\preceq$ is a relation on X. The relation $\preceq$ is called a **partial order** if it satisfies the following three properties:

(1) **Reflexive:** For all $x \in X$, $x \preceq x$.
(2) **Antisymmetric:** For all $x, y \in X$, if $x \preceq y$ and $y \preceq x$, then $x = y$.
(3) **Transitive:** For all $x, y, z \in X$, if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

A set equipped with a partial order $(X, \preceq)$ is called a **partially ordered set**, or a **poset** for short.

Introduction to Analysis
Rui Wang
Draft

REMARK 1.5. More precisely, a relation on a set X is simply a subset of $X \times X$. In this language, writing $x \preceq y$ means that the ordered pair (x, y) belongs to this subset.

EXAMPLE 1.6. The usual order $\leq$ is a partial order on $\mathbb{Q}$, and similarly, $\geq$ is also a partial order on $\mathbb{Q}$. The same statements hold if we replace $\mathbb{Q}$ by $\mathbb{Z}$.

Another simple example is the equality relation $=$ on $\mathbb{Q}$, which defines a partial order. This is called the *trivial order*, and it exists on every set.

In fact there are many interesting partial orders on $\mathbb{Q}$.

EXAMPLE 1.7. Define $x \preceq y$ if $x \neq 0$ and $\frac{y}{x} \in \mathbb{Z}^+$ or $x = y = 0$. Check: This is a partial order.

Similarly, one can consider this partial order on $\mathbb{Z}$.

REMARK 1.8. Any partial order restricts to any subset by keeping all existing comparisons among elements of that subset.

Conversely, a partial order on a subset extends to the whole set via the minimal extension: keep the old comparisons inside the subset, and let every outside element be comparable only with itself. (Other extensions are also possible.)

EXAMPLE 1.9. Let X be a set, and write 2^X for the power set of X, whose elements are subsets of X.

For example, if $X = \{1, 2, 3\}$, then

$$2^X = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

This set has $8 = 2^3$ elements.

The inclusion relation $\subseteq$ is a natural partial order on 2^X.

DEFINITION 1.10 (Total order). Suppose X is a set and $\preceq$ is a relation on X. The relation $\preceq$ is called a **total order** if it is a partial order and for any two elements $x, y \in X$, there are

$$x \preceq y \text{ or } y \preceq x.$$

A set equipped with a total order $(X, \preceq)$ is called a **totally ordered set**, or simply an **ordered set**.

We write $x \preceq y$ also as $y \succeq x$. If $x \preceq y$ but $x \neq y$, we write $x \prec y$. Similarly, if $y \succeq x$ but $x \neq y$, we write $y \succ x$.

EXAMPLE 1.11.
(1) $\leq, \geq$ are total orders on $\mathbb{Q}$, but $=$ is not.
(2) If $\preceq$ is a total order on a set X, then restrict the relation to a subset S of X, $\preceq$ is also a total order on S. For example, $\leq, \geq$ are also total orders on $\mathbb{Z}$.
(3) The inclusion of power set 2^X is not a total order but only a partial order.

1.4. Infimum and supremum. For an ordered set, we can talk about upper (lower) bounds and least upper (greatest lower) bounds for its subsets.

DEFINITION 1.12. Assume S is a subset of an order set $(X, \preceq)$.

(1) An element $x \in X$ is called an upper bound of the subset S, if for every $s \in S$, there is

$$s \preceq x.$$

duction to Analysis Rui Wang Draft

(2) An element $x \in X$ is called a lower bound of the subset S, if for every $s \in S$, there is

$$x \preceq s.$$

EXAMPLE 1.13. (1) Find an upper bound and a lower bound of $\{x \in \mathbb{Z} | x^2 \leq 2\} \subseteq \mathbb{Z}$.

(2) Find an upper bound and a lower bound of $\{x \in \mathbb{Q} | x^2 \leq 2\} \subseteq \mathbb{Q}$.

Are there a smallest upper bound or a biggest lower bound?

DEFINITION 1.14. (1) Let S be a subset of an ordered set $(X, \preceq)$. If $x_0 \in X$ is an upper bound of S and no element $x \prec x_0$ in X is an upper bound of S, then x_0 is called the **least upper bound** (lub) of S. The least upper bound is also called the **supremum** (sup) of S and is denoted by $\sup_X S$.

(2) Let S be a subset of an ordered set $(X, \preceq)$. If $x_0 \in X$ is a lower bound of S and no element $x \succ x_0$ in X is a lower bound of S, then x_0 is called the **greatest lower bound** (glb) of S. The greatest lower bound is also called the **infimum** (inf) of S and is denoted by $\inf_X S$.

LEMMA 1.15. *The least upper bound (lub) or the greatest lower bound (glb) is unique if it exists.*

PROOF. We prove uniqueness for the least upper bound; the case of the greatest lower bound is analogous and left to readers.

Suppose x_1 and x_2 are both least upper bounds of a subset $S \subseteq X$. Since x_2 is a least upper bound, any element $x \prec x_2$ fails to be an upper bound of S. But x_1 is an upper bound, hence we must have $x_1 \succeq x_2$. By the same argument, $x_2 \succeq x_1$. Therefore, by antisymmetry of the order, $x_1 = x_2$. This proves the least upper bound, if it exists, is unique.

□

Not every subset of an ordered set has a supremum (or infimum), even when it has an upper bound (or lower bound).

EXAMPLE 1.16. Denote by

$$S = \{x \in \mathbb{Q}^+ | x^2 \leq 2\}.$$

Then $\sup_{\mathbb{Q}} S$ does not exist. (Here $\mathbb{Q}^+$ is the set of all positive rational numbers.)

PROOF. First, we can easily check for example $1, 1.1 \in S$ and $2 \notin S$, and these imply the lub, if exists, denoted by x_0, must satisfy

$$x_0 \geq 1.1 > 1, \quad x_0 \leq 2.$$

We now try to obtain contradiction to the existence of x_0 by considering the following three cases.

- Case $x_0^2 = 2$.

  Since $x_0 \in \mathbb{Q}$ and $x_0 > 0$, we can write

  $$x_0 = \frac{p}{q}, \quad p, q \in \mathbb{Z}^+, \quad \gcd(p, q) = 1.$$

  Then it follows from $x_0^2 = 2$ that

  $$p^2 = 2q^2,$$

  and so $2 \mid p^2$.

  Recall the property of a prime number: If d is a prime number, and $a, b \in \mathbb{Z}$, then $d \mid (ab)$ implies $d \mid a$ or $d \mid b$. We have now $2 \mid p$ for our case.

Introduction to Analysis Rui Wang Draft

Write $p = 2k$, for $k \in \mathbb{Z}$, then there is

$$2k^2 = q^2,$$

and so $2 \mid q^2$. By the same argument, $2 \mid q$.

Since 2 is a common factor of p, q, we obtain a contradiction to the assumption that $\gcd(p, q) = 1$.

Thus this case is not possible.

- Case $x_0^2 < 2$.

  For this case, we can construct a rational number $\epsilon > 0$ such that

  $$(x_0 + \epsilon)^2 < 2.$$

  Then since

  $$x_0 + \epsilon \in S, \quad x_0 + \epsilon > x_0,$$

  it will lead to a contradiction to x_0 being an upper bound.

  To construct such an ϵ, we can do some estimates (estimates are the most common tool in analysis). Calculate

  $$(x_0 + \epsilon)^2 = x_0^2 + 2\epsilon x_0 + \epsilon^2.$$

  If we require $\epsilon < 1$ (that can be checked later), then

  $$x_0^2 + 2\epsilon x_0 + \epsilon^2 < x_0^2 + 2\epsilon x_0 + \epsilon = x_0^2 + \epsilon(2x_0 + 1),$$

  and we can try to set

  $$\epsilon = \frac{2 - x_0^2}{2x_0 + 1}.$$

  Then it is straightforward to check that the so-defined ϵ satisfies

  $$\epsilon \in \mathbb{Q}^+, \quad \epsilon < 1,$$

  and then $(x_0 + \epsilon)^2 < 2$, under the assumption $x_0^2 < 2$.

  This brings the contradiction we expect.

- Case $x_0^2 > 2$.

  For this case, we will construct $\epsilon > 0$ so that $x_0 - \epsilon$ is an upper bound of S. Then this will lead to a contradiction that x_0 is the least upper bound.

  To find such $\epsilon \in \mathbb{Q}$, we may consider

  $$(x_0 - \epsilon)^2 = x_0^2 - 2\epsilon x_0 + \epsilon^2 > x_0^2 - 2\epsilon x_0 = 2,$$

  and solve

  $$\epsilon = \frac{x_0^2 - 2}{2x_0}.$$

  By the assumption $x_0^2 > 2$, we have $\epsilon > 0$. It is also straightforward to check that $\epsilon < x_0$, and then we have

  $$x_0 - \epsilon \in \mathbb{Q}^+.$$

  Now we claim that every $x \in S$ must satisfy

  $$x \leq x_0 - \epsilon,$$

  since otherwise, if some $x > x_0 - \epsilon$, there follows

  $$2 > x^2 > (x_0 - \epsilon)^2 > 2,$$

Introduction to Analysis
Rui Wang
Draft

which is a contradiction.

Thus $x_0 - \epsilon$ is also an upper bound of S but strictly smaller than x_0. By this, we exclude this case too.

Above all, $\sup_{\mathbb{Q}} S$ doesn't exist.

□

REMARK 1.17. In the above proof, the choice of ϵ is very flexible—any value of ϵ that leads to a contradiction will do the job. This is a common phenomenon in analysis: we often obtain the desired conclusion through estimates and inequalities, but the choice of parameters in such estimates is usually not unique.

Rudin gives a more sophisticated (but trickier) construction: set

$$x := x_0 - \frac{x_0^2 - 2}{x_0 + 2} = \frac{2x_0 + 2}{x_0 + 2} \in \mathbb{Q}^+,$$

for which a short calculation shows

$$x^2 - 2 = \frac{2(x_0^2 - 2)}{(x_0 + 2)^2}.$$

If $x_0^2 < 2$ then $x_0 < x$ and $x^2 < 2$, contradicting that x_0 is an upper bound; if $x_0^2 > 2$ then $x < x_0$ and $x^2 > 2$, so x is also an upper bound, contradicting the leastness of x_0.

DEFINITION 1.18. We call a totally ordered set $(X, \preceq)$ has the **least-upper-bound (lub) property** if any nonempty subset of X with an upper bound has a least upper bound.

From the example above, we have seen that $(\mathbb{Q}, \leq)$ doesn't have the lub property.

REMARK 1.19. The emphasis on lub is not asymmetric: the greatest lower bound (glb) property is actually equivalent.

As an exercise, prove that $(X, \preceq)$ has the lub property if and only if it has the glb property.

The defect that $(\mathbb{Q}, \leq)$ doesn't satisfy the lub property is the main motivation that we want to extend $\mathbb{Q}$ to a larger number system, and so introduce real numbers.

1.5. Ordered field.

DEFINITION 1.20. Suppose $(F, +, \cdot)$ is a field. It is called an **ordered field** if there is a total order $(F, \preceq)$ such that

(1) $a + b \prec a + c$ for all $a, b, c \in F$ with $b \prec c$.
(2) For all $a, b \in F$ with $a \succ 0$ and $b \succ 0$, there is $ab \succ 0$.

Clearly, $\mathbb{Q}$ is an ordered field for the natural addition, multiplication and the order $\leq$. (Notice: it is not an ordered field with respect to the total order $\geq$.)

PROPOSITION 1.21. *For an ordered field $(F, +, \cdot, \preceq)$, there are the following statements.*

(1) *If $a \prec 0$, then $-a \succ 0$, and vice versa.*
(2) *If $a \prec 0$ and $b \succ c$, then $ab \prec ac$.*
(3) *If $a \succ 0$ and $b \succ c$, then $ab \succ ac$.*
(4) *If $a \neq 0$, then $a^2 \succ 0$. In particular, $0 \prec 1$.*
(5) *If $0 \prec a \prec b$, then $0 \prec 1/b \prec 1/a$.*

Introduction to Analysis
Rui Wang
Draft

PROOF. We prove the first one and the rest are left to readers (or refer Prop 1.18 from Rudin's book). If $-a \succ 0$ is not true, then by the definition of a total order,

$$-a \preceq 0.$$

It follows from the ordered field property (1) that

$$0 = a + (-a) \prec 0 + (-a) = -a,$$

and then $-a \succ 0$ which is a contradiction. □

2. Real numbers $\mathbb{R}$

2.1. Dedekind cuts. The construction of the real numbers via *cuts* was given by Dedekind in 1872. In the same year, Cantor introduced another approach, using Cauchy sequences. Here we outline Dedekind's construction without detailed proofs (for proofs, see Chapter 1, Appendix in Rudin's book).

Our goal is to construct a larger set $\mathbb{R}$ containing $\mathbb{Q}$, which preserves the algebraic and order properties of $\mathbb{Q}$ (as an ordered field) and, in addition, satisfies the *least-upper-bound property*.

The idea is to construct $\mathbb{R}$ from $\mathbb{Q}$ as a subset of the power set $2^{\mathbb{Q}}$, that is, we view the elements of $\mathbb{R}$ as particular subsets of $\mathbb{Q}$. The construct is done via the following steps.

(1) We identify each rational number $q \in \mathbb{Q}$ with the subset

$$\{x \in \mathbb{Q} | x < q\} \subseteq \mathbb{Q}.$$

Equivalently, we consider the map

$$F : \mathbb{Q} \to 2^{\mathbb{Q}}, \quad q \mapsto \{x \in \mathbb{Q} | x < q\}.$$

The map F is injective, so we may regard $\mathbb{Q}$ as its image $F(\mathbb{Q}) \subseteq 2^{\mathbb{Q}}$.

By taking suprema, elements in $F(\mathbb{Q})$ can be mapped back to $\mathbb{Q}$.

Through this correspondence, $F(\mathbb{Q})$ inherits the operations of $\mathbb{Q}$, such as addition and multiplication. For example, we define addition on $F(\mathbb{Q})$ by

$$F(a) + F(b) := F(a + b).$$

Notice that $2^{\mathbb{Q}}$ contains subsets which have no supremum in $\mathbb{Q}$, so the map F is not surjective.

(2) A subset C of $\mathbb{Q}$ is called a **cut** if
   (a) $C \neq \emptyset$ and $C \neq \mathbb{Q}$;
   (b) If $c \in C$, then every $x < c$ also lies in C;
   (c) If $c \in C$, then there exists some $x > c$ with $x \in C$.

Clearly, for any $q \in \mathbb{Q}$, the subset $\{x \in \mathbb{Q} | x < q\}$ (that is an element of $F(\mathbb{Q})$) is a cut. However, not every cut comes from $F(\mathbb{Q})$. For example, the set

$$S = \{x \in \mathbb{Q} | x \leq 0 \text{ or } x^2 < 2\}$$

is a cut, but it is not in $F(\mathbb{Q})$ since $\sup_{\mathbb{Q}} S$ does not exist.

We define $\mathbb{R}$ as the set of all cuts in $\mathbb{Q}$. Then one can check:

(a) $\mathbb{R}$ is ordered by $\subseteq$.

Introduction to Analysis
Rui Wang
Draft

(b) Addition is defined by

$$C_1 + C_2 := \{x_1 + x_2 | x_1 \in C_1; x_2 \in C_2\}.$$

(c) Multiplication is more delicate, but can also be defined on $\mathbb{R}$.

(d) $(\mathbb{R}, +, \cdot)$ is a field.

(e) $(\mathbb{R}, +, \cdot; \subseteq)$ is an ordered field.

(f) $\mathbb{R}$ satisfies the least upper bound property: suppose $\emptyset \neq S \subseteq \mathbb{R}$ has an upper bound a. Then

$$b := \bigcup_{s \in S} s$$

is itself a cut, and it is the least upper bound of S. (This proof is a good exercise to check your understanding of the construction of Dedekind cuts, and we leave the details to the reader.)

We use $\mathbb{R}$ to denote the ordered field constructed above; this is the **field of real numbers**.

EXAMPLE 2.1. For any nonnegative rational number r, we denote by $\sqrt{r}$ the cut

$$\{x \in \mathbb{Q} | x \leq 0 \text{ or } x > 0, x^2 < r\}.$$

If $r = s^2$ for some $s \in \mathbb{Q}$ and $s \geq 0$, then the cut

$$\{x \in \mathbb{Q} | x \leq 0 \text{ or } x > 0, x^2 < r\} = \{x \in \mathbb{Q} | x < s\},$$

and the preimage, i.e., the image under supremum function, is simply the rational number s.

Otherwise, $\sqrt{r}$ is an irrational number, i.e., a real number but not a rational number.

It is a worthwhile task to recover the known properties of real numbers using the terms of Dedekind's cuts, for example try to prove

$$(\sqrt{x})^2 = x.$$

From now on, when we consider the supremum or infimum in $\mathbb{R}$ for a subset $S \subseteq \mathbb{R}$, we omit the subscript $\mathbb{R}$ in $\sup_{\mathbb{R}}$ or $\inf_{\mathbb{R}}$ and simply write $\sup S$ or $\inf S$.

For a subset $S \subseteq \mathbb{R}$, if $\sup S$ exists (equivalently, if S has an upper bound) and $\sup S \in S$, then we call it the *maximum* of S, denoted by $\max S$. Similarly, if $\inf S$ exists (equivalently, if S has a lower bound) and $\inf S \in S$, then we call it the *minimum* of S, denoted by $\min S$.

For example,

$$\sup(1, 2] = \max(1, 2],$$

but since $\inf(1, 2] = 1 \notin (1, 2]$, the minimum of $(1, 2]$ does not exist.

Denote by

$$\mathbb{R}^+ := \{x \in \mathbb{R} | x > 0\}, \quad \mathbb{R}^* := \{x \in \mathbb{R} | x \neq 0\}.$$

Formally, we can write for every nonempty subset $S \subseteq \mathbb{R}$ that

$$\sup_{\mathbb{R}} S = +\infty, \quad \text{if } S \text{ has no upper bound in } \mathbb{R};$$

and

$$\inf_{\mathbb{R}} S = -\infty, \quad \text{if } S \text{ has no lower bound in } \mathbb{R}.$$

Notice that these are only formal notations and there are no algebraic operations on $\pm\infty$.

Next, we are going to use the lub property of $\mathbb{R}$ to prove some important properties of $\mathbb{R}$: $\mathbb{R}$ is Archimedean; $\mathbb{Q}$ is dense in $\mathbb{R}$; $\mathbb{R}$ is closed under taking roots.

Introduction to Analysis Rui Wang Draft

2.2. $\mathbb{R}$ is archimedean.

THEOREM 2.2. *For any $x \in \mathbb{R}^+$ and $y \in \mathbb{R}$, there exists some $n \in \mathbb{Z}^+$ so that*

$$n \cdot x > y.$$

In particular, if we take $x = 1$ from this theorem, we immediately get the following statement.

PROPOSITION 2.3. *For any $y \in \mathbb{R}$, there exists some positive integer n so that $n > y$.*

We first prove Proposition 2.3 without using Theorem 2.2, and then we prove Theorem 2.2 from Proposition 2.3. This in fact shows that these two statements are equivalent, though Proposition 2.3 looks much simpler.

THE PROOF OF PROPOSITION 2.3. Suppose the statement is not true, i.e., there is some $y \in \mathbb{R}$ such that there is no $n \in \mathbb{Z}^+$ that can bound it. Then in another word, for each $n \in \mathbb{Z}^+$, there is

$$n \leq y,$$

which is saying y is an upper bound of $\mathbb{Z}^+$.

Apply the lub property of $\mathbb{R}$, the least upper bound $\sup_{\mathbb{R}} \mathbb{Z}^+$ exists and we use $y_0 \in \mathbb{R}$ to denote it.

Consider $y_0 - 1$. By the ordered field property (Proposition 1.21) of $\mathbb{R}$, there is

$$y_0 - 1 < y_0.$$

(Why? Fill details here.)

Since y_0 is the lub, there is some $N \in \mathbb{Z}^+$ such that

$$N > y_0 - 1.$$

Then it follows $N + 1 > y_0$. Since $N + 1 \in \mathbb{Z}$, this brings contradiction to the fact that y_0 as an upper bound.

Thus the original assumption is not possible and we are done. □

PROOF OF THEOREM 2.2. For any $x \in \mathbb{R}^+$ and $y \in \mathbb{R}$, consider $y \cdot x^{-1} \in \mathbb{R}$. From Proposition 2.3, there exits some $n \in \mathbb{Z}^+$ so that

$$n > y \cdot x^{-1}.$$

Then by the ordered field property (Proposition 1.21) of $\mathbb{R}$, this shows

$$nx > y.$$

□

REMARK 2.4. The archimedean property guarantees the decimal representation of real numbers: Given $a \in \mathbb{R}^+$, the archimedean property guarantees that there is some $N \in \mathbb{Z}^+$ such that $a < N$.

It follows the set

$$\{n \in \mathbb{Z} | -N \leq n \leq a\} \subseteq \{n \in \mathbb{Z} | -N \leq n \leq N\},$$

which is a nonempty finite set. Then can define

$$n_0 := \max\{n \in \mathbb{Z} | -N \leq n \leq a\} = \max\{n \in \mathbb{Z} | n \leq a\},$$

i.e., the largest integer which is not bigger than a.

Next consider $a_1 := 10(a - n_0) \in \mathbb{R}^{\geq 0}$. If $a_1 = 0$, then

$$a = n_0;$$

Introduction to Analysis Rui Wang Draft

Otherwise, define n_1 by the same way:

$$n_1 := \max\{n \in \mathbb{Z} | n \leq a_1\},$$

and then consider

$$a_2 := 10^2(a_1 - n_1).$$

If $a_2 = 0$, then

$$a_1 = n_1, \quad a = n_0 + \frac{n_1}{10},$$

which can be denoted as $n_0.n_1$.

If $a_2 > 0$, then define

$$n_2 := \max\{n \in \mathbb{Z} | n \leq a_2\}.$$

Keep going this way: By considering

$$a_{k+1} := 10^k(a_k - n_k).$$

If $a_{k+1} = 0$, then a can be written as

$$n_0.n_1n_2 \cdots n_k = n_0 + \frac{n_1}{10} + \frac{n_2}{10^2} + \cdots + \frac{n_k}{10^k}.$$

Otherwise, one can compute

$$n_{k+1} = \max\{n \in \mathbb{Z} | n \leq a_{k+1}\}.$$

Then each $a \in \mathbb{R}^+$, we can find a sequence

$$n_0.n_1n_2 \cdots$$

with $n_0, n_1, \cdots \in \mathbb{Z}^{\geq 0}$. This is called the decimal representation of a.

2.3. $\mathbb{Q}$ is dense in $\mathbb{R}$.

THEOREM 2.5. *For any $a, b \in \mathbb{R}$ with $a < b$, there exists some $q \in \mathbb{Q}$ so that $a < q < b$.*

PROOF. This is equal to say that one can find some $m \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$ so that

$$a < \frac{m}{n} < b.$$

which is further equivalent to find $m \in \mathbb{Z}$ and $n \in \mathbb{Z}^+$ so that

$$an < m < bn.$$

Notice that $b - a > 0$, by the archimedean property Theorem 2.2, there exits $n \in \mathbb{Z}^+$ so that

$$bn - an = (b - a)n > 1.$$

Let's argue that there exists some integer between two real numbers when their difference is bigger than 1.

LEMMA 2.6. *For any $\alpha, \beta \in \mathbb{R}$ with $\beta - \alpha > 1$, there exists some integer m so that $\alpha < m < \beta$.*

PROOF OF LEMMA 2.6. First, using archimedean property of $\mathbb{R}$, we can find some integer $N > 0$ so that

$$-N < -N + 1 < \alpha < \beta < N - 1 < N.$$

(Why? Fill details for this step.) Then

$$\max\{n \in \mathbb{Z} | -N < n \leq \alpha\} =: k_1$$

Introduction to Analysis Rui Wang Draft

and

$$\min\{n \in \mathbb{Z} | \beta \leq n < N\} =: k_2$$

exist since they are both nonempty finite sets.

If there is no integer between α and β, then k_1 and k_2 are consecutive integers and there must be

$$k_2 - k_1 = 1.$$

On the other hand,

$$k_2 - k_1 \geq \beta - \alpha > 1,$$

which is a contradiction.

Hence, there must be some integer between α and β. □

At last, apply the lemma to $\alpha = an$ and $\beta = bn$, we are done.

□

2.4. $\mathbb{R}$ is closed under taking roots (* This part was skipped from lectures.) We have already mentioned that the real number $\sqrt{2}$, the positive root of the equation $x^2 = 2$, corresponds to the Dedekind cut

$$\{ x \in \mathbb{Q} : x < 0 \text{ or } x^2 < 2 \}.$$

In this section, we outline a proof that $\mathbb{R}$ is closed under taking roots. More precisely, we will show that for every $n \in \mathbb{Z}^+$ and every $a \in \mathbb{R}^+$, the equation $x^n = a$ has a solution in $\mathbb{R}$. Our argument uses the least upper bound property of the real numbers, rather than relying on a specific construction such as Dedekind cuts.

We refer the reader to Rudin 1.21 Theorem for details in the proof.

THEOREM 2.7. *For every $y \in \mathbb{R}^+$ and every $n \in \mathbb{Z}^+$, there exists a unique $x \in \mathbb{R}^+$ so that $x^n = y$.*

PROOF. We first claim that such $x \in \mathbb{R}^+$, if exists, must be unique.

Otherwise, assume that both $x_1, x_2 \in \mathbb{R}^+$ are solutions of the equation

$$x^n = y, \quad y \in \mathbb{R}^+, n \in \mathbb{Z}^+.$$

Assume now $x_1 < x_2$, then from that fact that $\mathbb{R}$ is an ordered field, we have $x_1^n < x_2^n$ (Why? Fill details here.) and that is a contradiction. Similarly, $x_1 > x_2$ also leads to a contradiction, and so $x_1 = x_2$.

Now we prove the existence of solution.

Consider a subset of $\mathbb{R}$ as

$$S := \{a \in \mathbb{R}^+ | a^n < y\}.$$

Try to check that

(1) $S \neq \emptyset$;
(2) S has upper bound.

Then using the fact that $\mathbb{R}$ has the l.u.b. property, sup S exists. Define it as x, clearly, $x \in \mathbb{R}^+$. We show that x solves the equation. (The idea of the proof is similar to the proof of $\sup_{\mathbb{Q}}\{x \in \mathbb{Q} | x^2 < 2\}$ does not exist.)

First, we show that if $x^n < y$, then we can construct some $x_0 \in S$ which is greater than x, which says x is not an upper bound of S. So $x^n \geq y$.

Second, we show that if $x^n > y$, then we can find an upper bound of S which is smaller than x, which says x is not the least upper bound. So $x^n \leq y$.

Introduction to Analysis
Rui Wang
Draft

Above all, we must have $x^n = y$. □

From now on, we use $y^{\frac{1}{n}}$ to denote the unique solution for the equation

$$x^n = y, \quad y \in \mathbb{R}^+, n \in \mathbb{Z}^+,$$

and call it the n-th real root of y. The property

$$(ab)^{\frac{1}{n}} = a^{\frac{1}{n}} b^{\frac{1}{n}}$$

immediately follows from the uniqueness of n-th real root.

You may now return to the familiar world of calculus, using all the usual identities of the real numbers. These identities can all be proved rigorously within this framework, though we will omit the details here.

3. The Euclidean plane $\mathbb{R}^2$

3.1. The linear space $\mathbb{R}^2$. We consider the Cartesian product of $\mathbb{R}$ with $\mathbb{R}$:

$$\mathbb{R}^2 := \mathbb{R} \times \mathbb{R} = \{(x_1, x_2) | x_1, x_2 \in \mathbb{R}\}.$$

On $\mathbb{R}^2$, we consider the following operations:

- Addition:
$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2);$$
- Scalar multiplication: for $c \in \mathbb{R}$,
$$c \cdot (x_1, x_2) = (cx_1, cx_2).$$

These two operations make $\mathbb{R}^2$ into a 2-dimensional vector space (or called linear space) over the real field $\mathbb{R}$. For example, the set $\{(1, 0), (0, 1)\}$ forms a basis of it.

More generally, we may consider the Cartesian product of $\mathbb{R}$ with itself n times, namely

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}}.$$

With the analogous definitions of addition and scalar multiplication, $\mathbb{R}^n$ becomes a real vector space of dimension n, with the standard basis

$$\{e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots,\ e_n = (0, \ldots, 0, 1)\}.$$

3.2. The Euclidean space $\mathbb{R}^2$. On $\mathbb{R}^2$, we can consider the dot product, which is also called the inner product:

$$\langle (x_1, x_2), (y_1, y_2) \rangle = x_1 y_1 + x_2 y_2 = \Sigma_{j=1,2} x_j y_j.$$

The inner product induces the Euclidean norm defined as

$$|(x_1, x_2)| = \sqrt{\langle (x_1, x_2), (x_1, x_2) \rangle} = \sqrt{x_1^2 + x_2^2}.$$

PROPOSITION 3.1. *Denote by $\vec{x}$ the point (x_1, x_2) in $\mathbb{R}^2$. There are the following properties for the inner product and the Euclidean norm.*

(1) $|\vec{x}| \geq 0$ *and it is* 0 *if and only if* $\vec{x} = \vec{0}$.
(2) $|c \cdot \vec{x}| = |c||\vec{x}|$.
(3) $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$.
(4) $|\langle \vec{x}, \vec{y} \rangle| \leq |\vec{x}||\vec{y}|$.

Introduction to Analysis Rui Wang Draft

In next section, we will introduce the concept of metric. In fact, the Euclidean norm defines a natural metric

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

on $\mathbb{R}^2$.

All constructions here can be generalized to any $\mathbb{R}^n$ with $n \in \mathbb{Z}^+$.

4. The complex numbers $\mathbb{C}$

We will not work with complex numbers in this course, but this is a good opportunity to introduce some of their basic concepts.

As a set, and in fact as a vector space over $\mathbb{R}$, the complex numbers $\mathbb{C}$ can be identified with $\mathbb{R}^2$.

A complex number $z \in \mathbb{C}$ is a pair of real numbers $x, y \in \mathbb{R}$, usually written in the form

$$z = x + yi.$$

The real number x is called the real part of z and written as $x = \mathfrak{R}(z)$; the real number y is called the imaginary part of z and written as $y = \mathfrak{I}(z)$;

At this point, without yet defining multiplication on $\mathbb{C}$, the symbol i should simply be regarded as a convenient placeholder indicating the second coordinate of the ordered pair $(x, y) \in \mathbb{R}^2$.

It is the multiplication that brings out the truly interesting features of complex numbers. For

$$z = a + bi, \quad w = c + di, \quad a, b, c, d \in \mathbb{R},$$

we define

$$zw = (ac - bd) + (ad + bc)i.$$

In particular, $i = 0 + 1i$ then satisfies

$$i^2 = ii = -1.$$

Although this multiplication looks a bit complicated, it has a very clear geometric meaning. In fact, it contains both the dot product and the determinant (or 2-dimensional cross product) of vectors in $\mathbb{R}^2$. To see this, define the conjugate of a complex number by

$$\overline{a + bi} = a - bi, \qquad a, b \in \mathbb{R}.$$

Then, if we identify $z = a + bi$ with the vector $\vec{z} = (a, b) \in \mathbb{R}^2$, and similarly $w = c + di$ with $\vec{w} = (c, d) \in \mathbb{R}^2$, we have

$$z\overline{w} = \vec{z} \cdot \vec{w} + \det(\vec{z}, \vec{w})i,$$

where

$$\vec{z} \cdot \vec{w} = ac + bd, \qquad \det(\vec{z}, \vec{w}) = ad - bc.$$

The norm

$$|\vec{z}|^2 = z\overline{z}.$$

By direct verification, one can check that with respect to addition and multiplication, $\mathbb{C}$ forms a field. Moreover, $\mathbb{C}$ contains the field of real numbers, identified with the subset of complex numbers having zero imaginary part.

However, there is no total order on $\mathbb{C}$ can make it into an ordered field (See Rudin Ex 8).

Introduction to Analysis
Rui Wang
Draft

Introduction to Analysis
Rui Wang
Draft

CHAPTER 2

Basic topology

In this chapter we will view $\mathbb{R}$ from a topological perspective. Topology provides essential tools for analysis. Together with the ordered field structure of $\mathbb{R}$, this perspective allows us to prove many results from calculus that we have seen before but could not justify at that time.

Another advantage of introducing topology is that it lets us extend our discussion to much broader contexts with very little extra effort, leading to many interesting and useful results. For this reason, in this chapter we will define topology in an abstract way. Later in the course you will come to appreciate the value of this broader perspective, as it naturally encompasses the structure of spaces of functions.

The language in this chapter may feel more abstract than what we have seen before. However, the best way to understand each definition is always to connect it with concrete examples. Being able to move flexibly between precise definitions and intuitive examples is the key to mastering this material.

1. Countable and Uncountable Sets

So far we have encountered several subsets of $\mathbb{C}$:

$$\{1, 2, \dots, N\} \subseteq \mathbb{Z}^+ \subset \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C},$$

where N is a positive integer.

Using the notion of countability, we can classify these sets into three types:

(1) **Finite sets:** $\{1, 2, \dots, N\}$;
(2) **Countable sets:** $\mathbb{Z}^+$, $\mathbb{Z}$, $\mathbb{Q}$;
(3) **Uncountable sets:** $\mathbb{R}$, $\mathbb{C}$.

We will now introduce the precise definitions of these terms, and then examine both the examples above and some additional ones.

1.1. Definition of countability.

DEFINITION 1.1. Two sets X and Y are said to be in **one-to-one correspondence** if there exists a bijective map from X to Y. In this case we write $X \sim Y$.

Note that two sets may be in one-to-one correspondence through many different maps. For example, there are 6 distinct bijections between $X = \{1, 2, 3\}$ and $Y = \{A, B, C\}$.

DEFINITION 1.2. Let X be a set. A relation $\sim$ on X is called an **equivalence relation** if for all $x, y, z \in X$:

(1) Reflexive: $x \sim x$;
(2) Symmetric: if $x \sim y$, then $y \sim x$;
(3) Transitive: if $x \sim y$ and $y \sim z$, then $x \sim z$.

PROPOSITION 1.3. *Let $\mathcal{C}$ be a collection of sets. Define $X \sim Y$ if there exists a bijection between X and Y. Then $\sim$ is an equivalence relation on $\mathcal{C}$.*

Introduction to Analysis
Rui Wang
Draft

PROOF. Let $\mathcal{C}$ be a collection of sets, and define $X \sim Y$ if there exists a bijection from X to Y. We check the three properties:

(1) For any $X \in \mathcal{C}$, the identity map $id_X : X \to X$ is bijective. Hence $X \sim X$.
(2) For any $X, Y \in \mathcal{C}$, if $X \sim Y$, then there exists a bijective map

$$f : X \to Y.$$

Its inverse $f^{-1} : Y \to X$ is also bijective, and hence $Y \sim X$.
(3) For any $X, Y, Z \in \mathcal{C}$, if $X \sim Y$ and $Y \sim Z$, then there exist bijections

$$f : X \to Y, \quad g : Y \to Z.$$

Their composition $g \circ f : X \to Z$ is bijective, and hence $X \sim Z$.

□

Using this notion, a nonempty set X is called a **finite set** if there exists a positive integer N such that

$$X \sim \{1, 2, \dots, N\}.$$

In this case we write $|X| := N$, the number of elements of X. (Why is this definition of $|X|$ well defined?)

Obviously from definition, if $X \sim Y$ and one of them is finite, then the other one is also finite with $|X| = |Y|$. The bijective map between $\{1, 2, \cdots, N\}$ and X is in fact the counting map which counts the elements of X.

A nonempty set X is called **infinite**, if it is not a finite set. We use the following notations for finite or infinite sets: If a nonempty set X is finite, we write $|X| < \infty$; If a nonempty set X is infinite, we write $|X| = \infty$.

DEFINITION 1.4. (1) An infinite set X is called **countable**, if $\mathbb{Z}^+ \sim X$. A set is called **at most countable**, if it is either finite or countable.
(2) An infinite set X is called **uncountable**, if X is not countable.

Now let us examine some standard examples: $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$.

EXAMPLE 1.5. $\mathbb{Z}$ is countable.

PROOF. Construct a (counting) map

$$f : \mathbb{Z}^+ \to \mathbb{Z}, \quad k \mapsto (-1)^k \lfloor \frac{k}{2} \rfloor.$$

Here

$$\lfloor a \rfloor := \max\{n \in \mathbb{Z} | n \le a\}.$$

Check: The so defined f is a bijective map. □

(Ex: Write down the inverse map of f explicitly.)

EXAMPLE 1.6. $\mathbb{Q}$ is countable.

PROOF. Every rational number $x \in \mathbb{Q}$ can be written uniquely as a reduced fraction

$$x = \frac{m}{n}, \qquad m \in \mathbb{Z},\ n \in \mathbb{Z}^+,\ \gcd(m, n) = 1.$$

Introduction to Analysis
Rui Wang
Draft

This gives an injective map from $\mathbb{Q}$ to $\mathbb{Z} \times \mathbb{Z}^+$. The map is not surjective, but infectivity is enough. By the following two general propositions Proposition 1.7 and Proposition 1.8, we will obtain $\mathbb{Q}$ is countable. □

PROPOSITION 1.7. *If the sets X, Y are both countable, then the cartesian product $X \times Y$ is countable.*

PROOF. From the assumption X and Y are both countable, it follows $X \times Y \sim \mathbb{Z}^+ \times \mathbb{Z}^+$. Hence, it is enough to show $\mathbb{Z}^+ \times \mathbb{Z}^+$ is countable. (Why?)

Elements in $\mathbb{Z}^+ \times \mathbb{Z}^+$ are of the form (m, n), $m, n \in \mathbb{Z}^+$. We can assign a counting by regrouping them into disjoint union of diagonals. Define

$$S_i := \{(m, n) \in \mathbb{Z}^+ \times \mathbb{Z}^+ | m + n = i + 1\}, \quad i = 1, 2, \cdots.$$

Each diagonal S_i contains i elements, which can be listed with increasing order of x-coordinates as

$$(1, i), (2, i-1), (3, i-2), \cdots, (i, 1),$$

From definition, for distinct i, j, S_i and S_j are disjoint; and

$$\mathbb{Z}^+ \times \mathbb{Z}^+ = \sqcup_{i \in \mathbb{Z}^+} S_i.$$

(Here we use $\sqcup$ instead of $\cup$ to emphasize that these subsets are disjoint from each other.)

Now let's construct a map f from $\mathbb{Z}^+$ to $\sqcup_{i \in \mathbb{Z}^+} S_i = \mathbb{Z}^+ \times \mathbb{Z}^+$.

First, we need to find the right diagonal to map into. To do this, we count how many numbers are used up the diagonals before one reach n, and these numbers are counted by the sequence

$$n_1, n_2, n_3, \cdots,$$

where

$$n_k = |S_1| + |S_2| + \cdots + |S_k|, \quad k = 1, 2, \cdots.$$

(This sequence is in fact the triangle sequence $1, 3, 6, 10, 15, \cdots$.) Then the positive integer set is divided into disjoint sets:

$$I_1 = \{1\}$$
$$I_2 = \{2, 3\}$$
$$I_3 = \{4, 5, 6\}$$
$$I_4 = \{7, 8, 9, 10\}$$
$$\cdots,$$

where each $I_k = (n_{k-1}, n_k] \cap \mathbb{Z}$ by defining $n_0 = 0$. Then the integer n uniquely lives in one I_k, i.e., there is one and only one $k \in \mathbb{Z}^+$ such that

$$n_{k-1} < n \leq n_k.$$

We then map n to the k-th diagonal and specifically to the element

$$(n - n_{k-1}, k + 1 - (n - n_{k-1})).$$

This so-constructed map f is clearly bijective: Write $f(n) = (x_n, y_n)$, where

$$x_n = n - n_{k-1}, \quad y_n = k + 1 - (n - n_{k-1})$$

from above. Notice k is determined by n, so in fact $k = k(n)$, is a function in n.

Introduction to Analysis
Rui Wang
Draft

- Injectivity: If $f(p) = f(q)$, i.e., $x_p = x_q$, $y_p = y_q$, for some $p, q \in \mathbb{Z}^+$, then
$$k(p) = x_p + y_p - 1 = x_q + y_q - 1 = k(q).$$
This shows p, q are mapped to the same diagonal S_k, where $k = k(p) = k(q)$.
Then
$$x_p = p - n_{k-1} = x_q = q - n_{k-1}$$
implies $p = q$, and this proves the injectivity of f.
- Surjectivity: Take any $(x, y) \in \mathbb{Z}^+ \times \mathbb{Z}^+$, compute $k = x + y - 1$. Consider
$$p = x + n_{k-1}.$$
Then from definition of f
$$f(p) = (x, y).$$

(This construction of f may look complicated, but it is simply the formal translation of the picture where we count the points on each diagonal from left to right.) □

PROPOSITION 1.8. *Any subset of a countable set is at most countable.*

PROOF. We only need to show any infinite subset of $\mathbb{Z}^+$ is countable.
We first give a lemma whose proof is left to you as an exercise.

LEMMA 1.9. *For any subset $E \subseteq \mathbb{Z}^+$,* $\min E$ *exists.*

Assume $S \subseteq \mathbb{Z}^+$ is infinite. Define $S_1 = S$, $S_{n+1} = S_n \setminus \{\min S_n\}$, for $n = 1, 2, \cdots$, and a map
$$f : \mathbb{Z}^+ \to S, \quad n \mapsto \min S_n.$$
Then we check this map is bijective.

- It is injective: Assume $f(n) = f(n')$, then $\min S_n = \min S_{n'}$. If $n \neq n'$, and WLOG, assume $n > n'$, then from the construction we can see
$$S_n \subseteq S_{n'} \setminus \{\min S_{n'}\},$$
and $S_n \neq \emptyset$ since S is not finite. In particular, $\min S_n \in S_n$ but $\min S_{n'} \notin S_n$, this shows
$$\min S_n \neq \min S_{n'},$$
and we get contradiction.
- It is surjective: Take any $k \in S$, we show that there must be some $n \in \mathbb{Z}^+$ so that $f(n) = k$.
Consider the set
$$\{m \in S | m \leq k\}.$$
It is a finite set, and assume it contains n elements. Then $f(n) = k$ by the construction of f.

We are done with the proof of Proposition 1.8. □

An immediate corollary is

COROLLARY 1.10. *Any set that contains an uncountable set is uncountable.*

The results we have proved above lead easily to the following two theorems, both of which are very useful in practice. We leave the detailed proofs to the reader.

Introduction to Analysis Rui Wang Draft

THEOREM 1.11. *For any $N \in \mathbb{Z}^+$, if each $X_1, X_2, \dots, X_N$ is countable, then their Cartesian product*

$$X_1 \times X_2 \times \cdots \times X_N := \{(x_1, x_2, \dots, x_N) \mid x_i \in X_i,\ i = 1, 2, \dots, N\}$$

is countable.

PROOF. This is an immediate corollary of Proposition 1.7. □

THEOREM 1.12. *Suppose Λ is a countable set, and for each $\alpha \in \Lambda$, the set X_α is countable. Then the disjoint union*

$$\sqcup_{\alpha\in\Lambda} X_\alpha$$

is countable.

PROOF. Hint: Since each $X_\alpha \sim \mathbb{Z}^+$ and $\Lambda \sim \mathbb{Z}^+$, one can show that

$$\sqcup_{\alpha\in\Lambda} X_\alpha \sim \mathbb{Z}^+ \times \mathbb{Z}^+.$$

Then apply Proposition 1.7. □

As a corollary, we obtain:

COROLLARY 1.13. *Let X be a nonempty set and Λ be countable. If for each $\alpha \in \Lambda$, the set X_α is an at most countable subset of X, then the union*

$$\bigcup_{\alpha\in\Lambda} X_\alpha$$

is an at most countable subset of X.

1.2. $\mathbb{R}$ is uncountable. The following proof of the uncountability of $\mathbb{R}$, due to Georg Cantor (1845–1918), is one of the most elegant arguments in mathematics. It beautifully illustrates the power and clarity of a rigorous mathematical proof.

THEOREM 1.14. *Consider a sequence of sets $X_1, X_2, \cdots$, with each X_i contains at least two elements, then their cartesian product*

$$X_1 \times X_2 \times \cdots := \{(x_1, x_2, \cdots,) | x_i \in X_i\}$$

is uncountable.

PROOF.

(1) Using Corollary 1.10, it is enough to prove the proposition for the case that each X_i contains exactly 2 elements.

(2) Now we prove the case of each set containing exactly two elements. Assume

$$X_i = \{x_i^0, x_i^1\}, \quad i = 1, 2, \cdots.$$

If $X_1 \times X_2 \times \cdots$ is a countable set, then we can find an one-to-one correspondence between $\mathbb{Z}^+$ and it, which is the same to index each element in $X_1 \times X_2 \times \cdots$ by a positive integer. Let's denote the indexed elements as

$$S_1, S_2, \cdots.$$

In fact, each S_k, $k = 1, 2, \cdots$, is a sequence

$$S_k = (s_{k1}, s_{k2}, s_{k3}, \cdots), \quad \text{with entries } s_{ki} \in X_i = \{x_i^0, x_i^1\}.$$

Introduction to Analysis Rui Wang Draft

We place these sequence $S_1, S_2, \cdots$ row by row to form an array, and use its diagonal to construct a new sequence:

$$S := (s_1, s_2, \cdots), \quad s_i := \overline{s_{ii}}$$

where $\overline{s_{ii}}$ denotes the other element of X_i different from s_{ii}.

Clearly S is in the Cartesian product $X_1 \times X_2 \times \cdots$, but it is different from every S_k, $k = 1, 2, \cdots$. That is because if

$$S = S_k$$

for some k, then they must have the same k-th element. The k-th element of the sequence S is $\overline{s_{kk}}$ by construction, while the k-th element of the sequence S_k is s_{kk}. They must be different.

This proves $X_1 \times X_2 \times \cdots$ is not countable and then is uncountable.

□

REMARK 1.15. Strictly speaking, the Cartesian product

$$X_1 \times X_2 \times \cdots$$

should be defined as the set of all maps

$$f : \mathbb{Z}^+ \to \bigcup_{i=1}^{\infty} X_i$$

such that $f(i) \in X_i$ for every $i \in \mathbb{Z}^+$.

More generally, one can define the Cartesian product of a family of sets indexed by an arbitrary set Λ, whether Λ is finite, countable, or uncountable:

$$\prod_{\alpha \in \Lambda} X_\alpha := \{ f : \Lambda \to \bigcup_{\alpha \in \Lambda} X_\alpha \mid f(\alpha) \in X_\alpha \text{ for all } \alpha \in \Lambda \}.$$

In other words, an element of the product is a function that selects one element from each X_α.

The definition itself always makes sense, but to guarantee that the product is nonempty whenever each X_α is nonempty, we need the **Axiom of Choice**. This axiom cannot be proved from the usual axioms of set theory, but is *assumed* as part of the foundations of modern analysis.

Then apply the decimal representation of $\mathbb{R}$, we can map the set of real numbers between 0 and 1 bijectively to the cartesian product $X_1 \times X_2 \times \cdots$ of $X_k = \{0, 1, \cdots, 9\}$. Theorem 1.14 proves that $(0, 1)$ is uncountable then.

Apply Corollary 1.10, $\mathbb{R}$ and $\mathbb{C}$ are both uncountable.

1.3. Basic exercises. If you do not feel very comfortable with this section on the cardinality of sets, the following basic exercises may help you build intuition. For more challenging problems, please refer to the homework assignments.

EXERCISE 1.16.
(1) Rigorously prove that any nonempty subset of a finite set is finite.
(2) Rigorously prove that the set of even (odd) integers is countable.
(3) Rigorously prove that the set of positive rational numbers is countable.
(4) Rigorously prove that the set of positive real numbers is uncountable.
(5) Rigorously prove that $\mathbb{Z} \times \mathbb{R}$ is uncountable.
(6) Rigorously prove that $\mathbb{R}^2$ is uncountable.

Introduction to Analysis Rui Wang Draft

2. Metric space

This section is the core of the chapter.

When you think of a metric space, the Euclidean metric should be the first example that comes to mind. At the same time, it is important to keep in mind other cases beyond the Euclidean one, which show that the concept of a metric space embraces a much broader range of examples. We will see some of these in this chapter, and more later on.

Metric spaces provide an extremely useful framework: they allow us to introduce powerful tools of analysis in a unified way. Later in the course, as we study sequences and spaces of functions, we will see that they too can be included in the category of metric spaces. The general properties we develop here will then apply directly to those new settings—often with surprising and elegant consequences.

2.1. Metric.

DEFINITION 2.1. A **metric space** is a nonempty set X together with a function

$$d : X \times X \to \mathbb{R},$$

called a **distance function** or **metric**, such that for all $x, y, z \in X$ the following properties hold:

(1) **Positivity:** $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$.
(2) **Symmetry:** $d(x, y) = d(y, x)$.
(3) **Triangle inequality:** $d(x, z) \leq d(x, y) + d(y, z)$.

EXAMPLE 2.2. (1) On $\mathbb{R}^n$, the **Euclidean metric** is defined by

$$d_{euclid}(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}, \quad x = (x_1, \dots, x_n),\ y = (y_1, \dots, y_n).$$

In particular, for $\mathbb{R}$,

$$d_{euclid}(x, y) = |x - y|.$$

(2) On $\mathbb{R}^n$, there are also other metrics. For example, the **taxicab metric** is defined by

$$d_{taxicab}(x, y) = \sum_{i=1}^{n} |x_i - y_i|.$$

(3) Another one is **supremum metric**:

$$d_{sup}(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|.$$

REMARK 2.3. In fact, these three examples are all metrics induced by norms on the vector space $\mathbb{R}^n$. In finite dimensions, all norms are equivalent, so these metrics induce the same topology. Later in the course (and in more advanced classes), one encounters their infinite-dimensional analogues on spaces of functions, usually denoted by L^2, L^1, and L^∞. These play a central role in analysis, but here we only mention their names as a preview.

EXAMPLE 2.4. Here is a very unusual example, showing how broad the notion of a metric space can be. On any nonempty set X, the **discrete metric** is defined by

$$d_{discrete}(x, y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$$

Introduction to Analysis Rui Wang Draft

A metric space (X, d) naturally induces a metric on any nonempty subset $S \subseteq X$ by simply restricting d to S:

$$d_S(x, y) := d(x, y), \quad x, y \in S.$$

We call this the **subspace metric**.

2.2. Open sets and closed sets.

DEFINITION 2.5. For any $x \in X$, $r > 0$,

(1) the subset $B_r(x) := \{y \in X | d(y, x) < r\}$ is called the open ball centered at x with radius r;
(2) the subset $\overline{B}_r(x) := \{y \in X | d(y, x) \leq r\}$ is called the closed ball centered at x with radius r.

An open ball centered at x is also called a neighborhood of x.

EXAMPLE 2.6. (1) In $\mathbb{R}$ with the Euclidean metric, an open (resp. closed) ball is just an open (resp. closed) interval of finite length, i.e. (a, b) (resp. $[a, b]$) with $a < b$.
(2) In $\mathbb{R}^2$ with the Euclidean metric, an open (resp. closed) ball is an open (resp. closed) disk of finite radius.

DEFINITION 2.7. Suppose (X, d) is a metric space, and S is a nonempty subset of X. A point $x \in S$ is called an **interior point** of S, if there is some $r > 0$ such that

$$B_r(x) \subseteq S.$$

We use S° to denote the set of all interior points of S.

DEFINITION 2.8. Suppose (X, d) is a metric space. A subset $S \subseteq X$ is called an **open set**, if $S = \emptyset$ or if every $x \in S$, there exists some $r > 0$ such that the open ball $B_r(x) \subseteq S$. In another word, if S is not empty, then it requires that $S = S^\circ$, i.e., every point in S is an interior point.

EXAMPLE 2.9. In any metric space (X, d), an open ball is open.

PROOF. By definition, any open ball $B_r(x)$, with $x \in X$, $r > 0$, is not empty since x is in it.

To check it is open, we consider an arbitrary point $y \in B_r(x)$ and show that there is some $r' > 0$ so that

$$B_{r'}(y) \subset B_r(x).$$

Since $y \in B_r(x)$, there is

$$d(y, x) < r.$$

Define $r' := r - d(y, x) > 0$.

Take any $z \in B_{r'}(y)$, let's show $z \in B_r(x)$.

For this, recall from the triangle inequality, there is

$$d(z, x) \leq d(z, y) + d(y, x) < r' + d(y, x) = (r - d(y, x)) + d(y, x) = r.$$

This proves $z \in B_r(x)$, and

$$B_{r'}(y) \subseteq B_r(x).$$

□

The following theorem is easy to prove, yet it captures some of the most fundamental properties of open sets in a metric space.

THEOREM 2.10. *For any metric space (X, d), the open sets in it satisfy the following properties.*

Introduction to Analysis Rui Wang Draft

(1) *Both $\emptyset$ and X are open sets.*
(2) *If $U_i \subseteq X$, $i = 1, \cdots, n$, are open sets, then their intersection $\cap_{i=1}^n U_i$ is an open set.*
(3) *For arbitrarily many open sets $U_\alpha \subseteq X$, $\alpha \in \Lambda$, their union $\cup_{\alpha\in\Lambda} U_\alpha$ is an open set.*

PROOF. (1) This is immediate from definition.

(2) If the intersection is empty, from definition, it is an open set. Otherwise, take an arbitrary point x from the intersection $\cap_{i=1}^n U_i$. Then $x \in U_i$ for every $i = 1, \cdots, n$.

By the definition of open sets, for each i, there is $r_i > 0$ such that

$$B_{r_i}(x) \subseteq U_i.$$

Define

$$r = \min\{r_1, \cdots, r_n\}.$$

Each $r_i > 0$ guarantees that $r > 0$, and

$$B_r(x) \subseteq B_{r_i}(x).$$

It follows

$$B_r(x) \subseteq \cap_{i=1}^n U_i,$$

and we have proved that $\cap_{i=1}^n U_i$ is open in X.

(3) If every U_α is empty, then the union is also empty and then an open set. Otherwise, the union is not empty, and every point x in the union lives in some U_{α_0}.

For this U_{α_0}, since it is open, there is an open ball

$$B_r(x) \subseteq U_{\alpha_0},$$

where $r > 0$. It follows

$$B_r(x) \subseteq U_{\alpha_0} \subseteq \cup_{\alpha\in\Lambda} U_\alpha.$$

This shows $\cup_{\alpha\in\Lambda} U_\alpha$ is open.

□

REMARK 2.11. In fact, the notion of a topology on a set is defined precisely as a collection of subsets that satisfies the three properties stated in Theorem 2.10. In this way, every metric space naturally becomes a topological space.

It is interesting to note, however, that not every topology comes from a metric. There are topological spaces which cannot be realized as metric spaces, though one can still study many important topological properties in that general setting. Deciding which topologies are metrizable is a subtle problem, and we will not address it here.

For the purposes of this course—and indeed for most (if not every) courses in real analysis and in differential geometry—we will work entirely within the framework of metric spaces.

However, the Zariski topology, which algebraists' favorite which plays a central role in algebra and in algebraic geometry, is not induced by any metric.

The following example shows that the finiteness assumption in part (2) of Theorem 2.10 cannot be omitted.

Introduction to Analysis Rui Wang Draft

EXAMPLE 2.12. For each $n \in \mathbb{Z}^+$, the interval

$$I_n := \left(-\tfrac{1}{n}, \tfrac{1}{n}\right) \subseteq \mathbb{R}$$

is open. However, their infinite intersection is

$$\bigcap_{n\in\mathbb{Z}^+} I_n = \{0\},$$

which is not open in $\mathbb{R}$.

Given a subset $S \subseteq X$, its complement in X is defined as

$$S^c := \{x \in X | x \notin S\}.$$

DEFINITION 2.13. A subset $S \subseteq X$ is called a **closed set**, if its complement is open.

EXAMPLE 2.14. (1) The closed interval $[a, b]$, $a \leq b$ is closed in $\mathbb{R}$ with respect to the Euclidean metric.
(2) A closed ball $\overline{B}_r(x)$, $r > 0$, $x \in X$, is closed. (Ex: Prove it.)

Using the definition of closed sets and Proposition 2.10, we immediately get the following analogue for closed sets, whose proof is left to the reader as an exercise.

THEOREM 2.15. *For any metric space (X, d), the closed sets in it satisfy the following properties.*
(1) *Both $\emptyset$ and X are closed.*
(2) *If $S_i \subseteq X$, $i = 1, \cdots, n$, are closed sets, then their union $\cup_{i=1}^n S_i$ is a closed set.*
(3) *For arbitrarily many closed sets $S_\alpha \subseteq X$, $\alpha \in \Lambda$, their intersection $\cap_{\alpha\in\Lambda} S_\alpha$ is a closed set.*

Again, the finiteness assumption in part (2) of Theorem 2.15 cannot be omitted.

EXAMPLE 2.16. Consider a sequence of closed sets $[-1 + \frac{1}{n}, 1 - \frac{1}{n}]$, $n \in \mathbb{Z}^+$, of $\mathbb{R}$. Take their union

$$\cup_{n\in\mathbb{Z}^+}[-1 + \frac{1}{n}, 1 - \frac{1}{n}] = (-1, 1),$$

which is not closed, since its complement

$$(-\infty, -1] \cup [1, +\infty)$$

is not open.

2.3. Limit points in a metric space.

DEFINITION 2.17. Suppose S is a nonempty subset of X. A point $x \in X$ is called a **limit point of** S, if for any $r > 0$, the intersection of the open ball $B_r(x)$ with S contains some point which is not x.

We use S' to denote the set of limit points of S in X, and the set

$$\overline{S} := S \cup S'$$

is called the **closure of** S **in** X.

Points in S but not in S' are called **isolated points** of S.

EXAMPLE 2.18. Consider the Euclidean metric space $(\mathbb{R}^n, d)$, with

$$S = B_r(x)$$

for some $x \in \mathbb{R}^n$, $r > 0$.

Introduction to Analysis Rui Wang Draft

(1) Any $y \in B_r(x)$ is a limit point of S.
(2) Any y with $d(y,x) = r$, i.e., a point on the bounding circle, is a limit point of S, even y is not in S.
(3) Any y with $d(y,x) > r$, i.e., a point which is outside the closed ball, is not a limit point of S.
Even when we modify S to
$$S = B_r(x) \cup \{y\},$$
such y is NOT a limit point of S.
(4) The limit set of $B_r(x)$ and $\overline{B_r(x)}$ are the same, which is $\overline{B_r(x)}$. In particular, the closure of $B_r(x)$ is exactly $\overline{B_r(x)}$.

REMARK 2.19 (Closed ball vs.closure notations). In a general metric space, a closed ball needs not coincide with the closure of the open ball with the same radius.

For example, the discrete metric space $(X, d_{discrete})$. Take a point $x \in X$. The open ball $B_1(x)$ is a single point set containing only x, and so is its closure. However the closed ball with radius 1 contains everything in X.

In this course—apart from this chapter—we will work almost exclusively with the Euclidean metric, so there is little risk of confusion. However, be careful when extending results to an arbitrary metric space.

EXAMPLE 2.20. Consider the Euclidean metric space $\mathbb{R}$. The subset
$$S := \{\frac{1}{n} | n \in \mathbb{Z}^+\}$$
has only one limit point, which is 0. Notice 0 is not in S.

Every point in S is an isolated point, and the closure of S is
$$\{\frac{1}{n} | n \in \mathbb{Z}^+\} \cup \{0\}.$$

EXAMPLE 2.21. Consider the Euclidean metric space $\mathbb{R}$. The set of rational numbers $\mathbb{Q}$ has
$$\mathbb{Q}' = \mathbb{R},$$
and then $\overline{\mathbb{Q}} = \mathbb{R}$.

In the proof of this result, we need to use Theorem 2.5.

DEFINITION 2.22. In a metric space (X, d), a subset S is called **dense** in X, if $\overline{S} = X$.

Then this example shows that $\mathbb{Q}$ is dense in $\mathbb{R}$.

PROPOSITION 2.23. *In a metric space (X, d), the closure of any subset $S \subseteq X$ is a closed set.*

PROOF. If $\overline{S} = X$, then it is closed automatically since $\overline{S}^c = \emptyset$.

We now prove it for $\overline{S} \subsetneq X$. To show it is closed, we need to prove the complement is open, which is the same to show every point $x \in (\overline{S})^c$ is an interior point by showing there is some $r > 0$ so that
$$B_r(x) \subseteq (\overline{S})^c, \quad \text{equivalently, } B_r(x) \cap \overline{S} = \emptyset.$$

Suppose this is not true. Then for every $\epsilon > 0$, there is some point
$$y \in B_\epsilon(x) \cap \overline{S}.$$
By the definition of $\overline{S}$, such y is either in S or in $S' \setminus S$.

Introduction to Analysis Rui Wang Draft

If $y \in S' \setminus S$, then we take take $\epsilon' > 0$ such that

$$B_{\epsilon'}(y) \subseteq B_{\epsilon}(x).$$

By the definition of limit point, $y \in S' \setminus S$ implies

$$B_{\epsilon'}(y) \cap S$$

contains some point y'. Then such $y' \in S \cap B_{\epsilon}(x)$.

The above arguments show that for any $\epsilon > 0$, there is some y or y' in $S \cap B_{\epsilon}(x)$.

Notice also y and y' are distinct from x since we have supposed that $x \notin \overline{S}$. It follows x is a limit point of S, which then contradicts with the assumption that $x \in \overline{S}^c$.

Thus, we proved $\overline{S}^c$ is open, and then $\overline{S}$ is closed.

□

PROPOSITION 2.24. *Suppose (X, d) is a metric space. A subset S of X is closed, if and only if $S = \overline{S}$.*

PROOF. We have shown the "if" part from Proposition 2.23.

Now we show the "only if" part.

From the definition $\overline{S} = S \cup S'$, there is immediately

$$S \subseteq \overline{S}.$$

For this other direction, we only need to show $S' \subset S$.

For any $x \in S'$, if $x \notin S$, then $x \in S^c$. Since S is closed, S^c is open. Then there is $r > 0$ such that

$$B_r(x) \subset S^c,$$

and then equivalently,

$$B_r(x) \cap S = \emptyset.$$

Then x is not a limit point of S, which contradicts the assumption.

Thus $x \in S$, and we are done.

□

PROPOSITION 2.25. *Suppose (X, d) is a metric space. For any $S \subseteq X$, the closure of S is the smallest closed subset of X that contains S. In fact,*

$$\overline{S} = \cap_{C \in \mathcal{C}} C,$$

where

$$\mathcal{C} = \{C \subseteq X | C \text{ is a closed set}, S \subseteq C\}.$$

PROOF. Denote by $A = \cap_{C \in \mathcal{C}} C$, where

$$\mathcal{C} = \{C \subseteq X | C \text{ is a closed set}, S \subseteq C\}.$$

By Theorem 2.15, A is closed set. It contains S since every C in $\mathcal{C}$ contains S. Thus A is the smallest closed set that contains S.

From $S \subseteq A$ and definition of closure, there is

$$\overline{S} \subseteq \overline{A} = A.$$

On the other hand, from Proposition 2.23, $\overline{S}$ is a closed set and thus $A \subseteq \overline{S}$.

Hence $\overline{S} = A$, the smallest closed set in X that contains S. □

Introduction to Analysis Rui Wang Draft

2.4. Sequences and their limits.

DEFINITION 2.26. Let X be a nonempty set. A **sequence** in X is a map from $\mathbb{Z}^+$ to X. Equivalently, it is an element in the Cartesian product $X \times X \times \cdots$.

Usually, one use

$$(x_n)_{n\in\mathbb{Z}^+} = (a_1, a_2, \cdots)$$

to denote a sequence.

DEFINITION 2.27. Suppose $(x_n)_{n\in\mathbb{Z}^+}$ is a sequence in X. A sequence

$$(x_{n_k})_{k\in\mathbb{Z}^+}$$

with $n_1 < n_2 < \cdots$ is called a **subsequence** of $(x_{n_k})_{k\in\mathbb{Z}^+}$.

EXAMPLE 2.28. $x_n = \frac{1}{n}$, $n = 1, 2, \cdots$, is a sequence in $\mathbb{R}$.

Consider the sequence of prime numbers:

$$n_1 = 2, n_2 = 3, n_3 = 5, n_4 = 7, n_5 = 11, \cdots.$$

Then

$$x_{n_k} = \frac{1}{n_k}, \quad k = 1, 2, \cdots$$

is a subsequence which is

$$\frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{7}, \frac{1}{11}, \cdots$$

concretely.

DEFINITION 2.29. Let $(x_n)_{n\in\mathbb{Z}^+}$ be a sequence in $\mathbb{R}$. A real number $x \in \mathbb{R}$ is called the **limit** of the sequence, written as

$$\lim_{n\to\infty} x_n = x$$

or written as

$$x_n \to x, \quad \text{as } n \to \infty,$$

if for any $\epsilon > 0$, there exists some $N \in \mathbb{Z}^+$ so that for every $n > N$ there is

$$|x_n - x| < \epsilon.$$

EXAMPLE 2.30. (1) $\lim_{n\to\infty} 1 = 1$.

(2) $\lim_{n\to\infty} \frac{1}{n} = 0$.

(3) $\lim_{n\to\infty}(-1)^n$ DNE.

Try to prove them using $\epsilon - N$ definition.

DEFINITION 2.31. Suppose (X, d) is a metric space. A point $x \in X$ is a called a **limit** of a sequence $(x_n)_{n\in\mathbb{Z}^+}$ if

$$\lim_{n\to\infty} d(x_n, x) = 0.$$

A sequence which doesn't have a limit is called **divergent**.

PROPOSITION 2.32. *In a metric space, if a sequence has a limit, then the limit must be unique.*

Introduction to Analysis
Rui Wang
Draft

PROOF. We provide two proofs here.

Proof by contradiction. Suppose a sequence $(x_n)_{n\in\mathbb{Z}^+}$ has two distinct limits x and x' in X.

Take $\epsilon = \frac{d(x,x')}{2}$.

Since $x_n \to x$, we can find $N \in \mathbb{Z}^+$ such that any $n > N$, there is

$$d(x_n, x) < \epsilon.$$

Similarly since $x_n \to x'$, we can find $N' \in \mathbb{Z}^+$ such that any $n > N'$, there is

$$d(x_n, x') < \epsilon.$$

Then when $n > \max\{N, N'\}$, there is

$$d(x, x') \le d(x, x_n) + d(x_n, x') = d(x_n, x) + d(x_n, x') < \epsilon + \epsilon = 2\epsilon.$$

Then

$$\epsilon > \frac{d(x, x')}{2},$$

which contracts $\epsilon = \frac{d(x,x')}{2}$.

Hence the limit is unique if exists.

A direct proof.

Assume both x and x' are limits of the sequence $(x_n)_{n\in\mathbb{Z}^+}$.

Then for any $\epsilon > 0$, there exists some $N, N' > 0$ so that

$$d(x_n, x) < \frac{\epsilon}{2}, \quad \text{for all } n > N,$$

and

$$d(x_n, x') < \frac{\epsilon}{2}, \quad \text{for all } n > N'.$$

Then whenever $x > \max\{N, N'\}$, both hold.

We estimate by the triangle inequality that

$$0 \le d(x, x') \le d(x, x_n) + d(x_n, x') = d(x_n, x) + d(x_n, x') < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

for any $\epsilon > 0$. This implies $d(x, x') = 0$, and thus there must be $x = x'$. □

EXERCISE 2.33. Formulate the ϵ–N definition for a **divergent** sequence. (Hint: modify the usual definition of convergence by replacing each "for every" with "there exists," and by reversing the inequality in the conclusion.)

PROPOSITION 2.34. *In a metric space, a convergent sequence $(x_n)_{n\in\mathbb{Z}^+}$ is bounded, i.e., there exists some $x_0 \in X$ and $M > 0$ so that $d(x_n, x_0) < M$ for any $n = 1, 2, \cdots$.*

PROOF. Assume $x_n \to x_0$. Then for $\epsilon = 1$, take N so that $n > N$, $d(x_n, x_0) < 1$.

Then define

$$M = \max\{x_1, \cdots, x_N, 1\},$$

we are done. □

Next, we see the relation between limit points of a set and limits of sequences.

PROPOSITION 2.35. *Assume (X, d) is a metric space and S is a nonempty subset. A point $x \in X$ is a limit point of S, if and only if there exists some sequence $(x_n)_{n\in\mathbb{Z}^+}$ with $x_n \in S \setminus \{x\}$ so that* $\lim_{n\to\infty} x_n = x$.

Introduction to Analysis Rui Wang Draft

PROOF. (1) " $\Longrightarrow$ ". Suppose x is a limit point of S, then for each $n \in \mathbb{Z}^+$, the open ball $B_{\frac{1}{n}}(x)$ intersects S at some point which is not x. We pick a such point, and denote it by x_n.

By this way, we obtain a sequence $(x_n)_{n\in\mathbb{Z}^+}$ with $x_n \in S \setminus \{x\}$. Also

$$d(x_n, x) < \frac{1}{n}$$

since $x_n \in B_{\frac{1}{n}}(x)$.

Then for each $\epsilon > 0$, there exists some $N > 0$ so that $\frac{1}{N} < \epsilon$. (This is due to the archimedean property, and actually we can just set $N = [\frac{1}{\epsilon}] + 1$.)

It follows each $n > N$,

$$d(x_n, x) < \frac{1}{n} < \frac{1}{N} < \epsilon.$$

This shows that $\lim_{n\to\infty} x_n = x$.

(2) " $\Longleftarrow$ ". Suppose there exists a sequence $(x_n)_{n\in\mathbb{Z}^+}$ with $x_n \in S \setminus \{x\}$ so that $\lim_{n\to\infty} x_n = x$.

Then for each open ball $B_\epsilon(x)$, we can find some $N \in \mathbb{Z}^+$ so that

$$x_n \in B_\epsilon(x), \quad \text{for all } n > N.$$

Since $x_n \in S \setminus \{x\}$, this shows that x is a limit point of S.

□

3. Compactness

3.1. Two definitions for compactness in metric spaces.

DEFINITION 3.1. Assume (X, d) is a metric space. A nonempty subset $K \subseteq X$ is called a **sequentially compact** subset in X, if every sequence in K has a convergent subsequence converging to a point in K.

EXAMPLE 3.2. (1) $\mathbb{R}$ is not sequentially compact, e.g., $\{x_n = n\}$ has no convergent subsequence;

(2) In any metric space, a finite set is a sequentially compact. (Why?)

(3) $(0, 1)$ is not sequentially compact, e.g., $(\frac{1}{n})_{n=2,3,\ldots}$ has no convergent subsequence;

(4) $[0, 1]$ is sequentially compact. This result looks obvious but its proof in fact needs l.u.p. of $\mathbb{R}$, boundedness and closeness of $[0, 1]$. It is a special case of the Heine–Borel theorem (also the Weierstrass theorem) which we will prove in later lectures.

DEFINITION 3.3. Assume (X, d) is a metric space. A collection of open sets $\{U_\alpha | \alpha \in \Lambda\}$ is called an **open cover** of a subset S of X, if

$$S \subseteq \cup_{\alpha\in\Lambda} U_\alpha.$$

For $\Lambda' \subseteq \Lambda$, if the subcollection $\{U_\alpha | \alpha \in \Lambda'\}$ is also an open cover of S, i.e.,

$$S \subseteq \cup_{\alpha\in\Lambda'} U_\alpha,$$

then $\{U_\alpha | \alpha \in \Lambda'\}$ is called a **subcover**. If moreover, Λ' is finite, then it is called a **finite subcover**.

DEFINITION 3.4. Assume (X, d) is a metric space. A subset $K \subseteq X$ is called a **compact subset** in X, if every open cover of K has a finite subcover of K.

EXAMPLE 3.5. (1) $\mathbb{R}$ is not compact, e.g., $\{(k, k + 2) | k \in \mathbb{Z}\}$ is an open cover but doesn't have a finite subcover.

Introduction to Analysis Rui Wang Draft

(2) In any metric space, a finite set is a compact. (Why?)
(3) Consider $(0,1) = \cup_{n=1}^{\infty}(\frac{2^{n-1}-1}{2^n}, \frac{2^n-1}{2^n})$. However, any finite subcover cannot cover $(0,1)$. (Why?)
(4) Consider $[0.01, 0.99] \subset \cup_{n=1}^{\infty}(\frac{2^{n-1}-1}{2^n}, \frac{2^n-1}{2^n})$. It has finite subcover. This is NOT enough to show $[0.01, 0.99]$ is compact, but we will prove later any closed interval in compact in $\mathbb{R}$.

PROPOSITION 3.6. *Assume (X,d) is a metric space. K is (sequentially) compact subset and S is a closed subset. Then $K \cap S$ is (sequentially) compact in X.*

PROOF.
- Compact statement:

  For any open cover of $\{U_\alpha | \alpha \in \Lambda\}$ of $K \cap S$, the collection of open sets

  $$\{S^c, U_\alpha | \alpha \in \Lambda\}$$

  is an open cover of K. Since K is compact, there must be a finite subcover of K.

  Moreover, in this finite subcover, the ones that belong to $\{U_\alpha | \alpha \in \Lambda\}$ form a finite cover of $K \cap S$. This shows that $K \cap S$ is compact.
- Sequentially compact statement

  Suppose $(x_n)_{n\in\mathbb{Z}^+}$ is a sequence in $K \cap S$, then it is also a sequence in K. Since K is sequentially compact, there is a subsequence $(x_{n_k})_{k\in\mathbb{Z}^+}$ that converges to a point x in K.

  At the same time, since S is a closed set, the subsequence $(x_{n_k})_{k\in\mathbb{Z}^+}$, which is also in S, has its limit point x in S. This shows that $(x_{n_k})_{k\in\mathbb{Z}^+}$ converges in $K \cap S$ to some point $x \in K \cap S$.

  Thus $K \cap S$ is Sequentially compact.

□

COROLLARY 3.7. *Assume (X,d) is a metric space. If K is (sequentially) compact subset and S is a closed subset with $S \subseteq K$, then S is a (sequentially) compact subset.*

We conclude this subsection with the following important theorem, which shows that in a metric space the two notions of compactness coincide.

THEOREM 3.8. *In any metric space (X,d), a subset K is a compact subset if and only if it is a sequentially compact subset in X.*

PROOF.
- (Compactness implies sequential compactness.) Assume $K \subseteq X$ is compact. Take any sequence (y_n) from K, let's prove it must have a convergent subsequence in K.

  Suppose this is not the case, i.e., any point $x \in K$ is not a limit of any subsequence of (y_n). Then for every $x \in K$, there exits some $r_x > 0$ such that the open ball $B_{r_x}(x) \setminus \{x\}$ contains no point in $\{y_n\}$.

  Consider $\{B_{r_x}(x) | x \in K\}$. It forms an open cover of K. By the compactness of K, there exists a finite subcover of K. We assume this subcover is

  $$\{B_{r_{x_1}}(x_1), \cdots, B_{r_{x_N}}(x_N)\}.$$

  In particular, every point in the sequence (y_n) is covered, and then there must be a subsequence $(y_{n_k})_{k\in\mathbb{Z}^+}$ that stays inside a fixed ball, say $B_{r_{x_i}}(x_i)$, for some $i = 1, \cdots, N$.

  Then it follows this is a constant x_i sequence, thus must converges to x_i, but then brings a contradiction to our assumption.

  As a result, (y_n) has a convergent subsequence in K and this proves K is sequentially compact.

Introduction to Analysis Rui Wang Draft

- (Sequentially compactness implies compactness.) We claim a fact without proof first. (It is an optional homework problem, with reference from Rudin's Chapter 2 Ex 24, 26.) Claim: Any infinite open cover of K has a countable subcover of K.

  With this claim, we start by assuming K has a countable covering

  $$\{U_n | n \in \mathbb{Z}^+\}.$$

  We now prove it must have a finite subcover. More concretely, we can find some $N \in \mathbb{Z}^+$ such that

  $$\{U_1, \cdots, U_N\}$$

  is a cover of K.

  Suppose there is no such N, then for each $n \in \mathbb{Z}^+$, there is some $x_n \in K \setminus (\cup_{i=1}^n U_i)$.

  By this way, we obtain a sequence $(x_n)_{n\in\mathbb{Z}^+}$ in K. Since K is sequentially compact, there is subsequence (x_{n_k}) and $x \in K$ so that $x_{n_k} \to x$ as $k \to \infty$.

  Now since $x \in K$, there must be some $\ell \in \mathbb{Z}^+$ such that $x \in U_\ell$. Notice that U_ℓ is open, there is some $\epsilon > 0$ such that

  $$B_\epsilon(x) \subseteq U_\ell.$$

  Then for this ϵ, we can choose some $N' \in \mathbb{Z}^+$ such that

  $$x_{n_k} \in B_\epsilon(x), \quad \text{for all } k > N'.$$

  In particular, when $n_k > \max\{n_{N'}, \ell\}$, it brings contradiction to the fact that

  $$x_{n_k} \notin U_\ell,$$

  which is due to $U_\ell \subset \cup_{i=1}^{n_k} U_i$ and

  $$x_{n_k} \notin K \setminus (\cup_{i=1}^{n_k} U_i).$$

  We are done with the proof now.

□

4. The Heine–Borel Theorem

The Heine–Borel theorem characterizes which subsets of the metric space $\mathbb{R}^n$ (with the standard Euclidean metric) are compact.

THEOREM 4.1 (The Heine–Borel Theorem). *A subset $K \subset \mathbb{R}^n$ with the standard Euclidean metric is compact if and only if it is a nonempty bounded and closed subset.*

In fact, the "only if" direction holds in any metric space. We restate this part separately and provide a proof first.

4.1. Some general results about compact sets in a metric space.

PROPOSITION 4.2. *Let (X, d) be a metric space. Then every sequentially compact subset of X is both bounded and closed.*

PROOF. Assume K is sequentially compact subset of X.

Introduction to Analysis Rui Wang Draft

(1) We first see K must be closed.

Suppose $x \in X$ is a limit point of K, then by Proposition 2.35, there is a sequence $(y_n)_{n\in\mathbb{Z}^+}$ in K with $y_n \neq x$ for every $n \in \mathbb{Z}^+$ so that

$$y_n \to x, \quad \text{as } n \to \infty.$$

Then by the sequentially compactness of K, the limit $x \in K$, and this proves K is compact.

(2) Next let's show K is bounded, and more concretely, take an arbitrary point $x_0 \in X$, we will show there is some $M \in \mathbb{R}^+$ such that

$$d(y, x_0) \leq M, \quad \text{for all } y \in K.$$

Suppose this is not true, then for each $n \in \mathbb{Z}^+$, there is some $y_n \in K$ such that

$$d(y_n, x_0) > n.$$

By this way, we constructed a sequence $(y_n)_{n\in\mathbb{Z}^+}$ in K. Then by the sequential compactness of K, it has a convergent subsequence $(y_{n_k})_{k\in\mathbb{Z}^+}$ with

$$y_{n_k} \to y, \quad \text{as } k \to \infty$$

and $y \in K$.

Then for $\epsilon = 1$, there is some $N \in \mathbb{Z}^+$ so that whenever $k > N$,

$$d(y_{n_k}, y) < 1,$$

and we obtain the following inequalities

$$n_k < d(y_{n_k}, x_0) \leq d(y_{n_k}, y) + d(y, x_0) < 1 + d(y, x_0),$$

which bring contradiction when $k > 1 + d(y, x_0)$.

By this way, we prove that such M exists and so K is bounded.

□

REMARK 4.3. In a metric space, every compact set is closed since every compact set is a sequentially compact set. This argument does not extend to general topological spaces: in a non-metrizable topology, a compact set need not be closed. Hence no compact version proof available for the above proposition.

As a side note, we record the following classical result, which follows from the above proposition and the equivalence of compactness and sequential compactness in metric spaces.

THEOREM 4.4. *Assume (X, d) is a metric space and let $\{K_\alpha | \alpha \in \Lambda\}$ be a collection of compact subsets in X. Then the intersection*

$$\cap_{\alpha\in\Lambda} K_\alpha \neq \emptyset$$

if and only the intersection of any finite subcollection of $\{K_\alpha | \alpha \in \Lambda\}$ is nonempty.

PROOF. (1) Since $\cap_{\alpha\in\Lambda} K_\alpha$ is a subset of the intersection any subcollection, the "only if" part immediately follows.

(2) Suppose now the intersection of any finite subcollection is not empty, but $\cap_{\alpha\in\Lambda} = \emptyset$. We will reach a contradiction from it.

Take some K_{α_0} from the collection and consider the collection

$$\{K_\alpha^c | \alpha \neq \alpha_0\}.$$

Introduction to Analysis Rui Wang Draft

Since $\cap_{\alpha\in\Lambda} = \emptyset$, this is a cover of K_{α_0}.

Further the compactness of each K_α implies each K_α is closed, and then the complement is open.

Hence this is an open over of the compact set K_α, and then must have a finite subcover, say $K_1, \cdots, K_n$.

Then it follows

$$K_{\alpha_0} \cap K_1 \cap \cdots \cap K_n = \emptyset,$$

which is a contradiction to the assumption.

□

COROLLARY 4.5. *Suppose (X, d) is a metric space and $\{K_n | n \in \mathbb{Z}^+\}$ is nested (nonempty) compact subsets in X with*

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots.$$

Then the intersection $\cap_{n\in\mathbb{Z}^+} K_n$ is not empty.

PROOF. This is immediate from Theorem 4.4 and details are left the reader as exercise. □

We end this part by the following example which shows the compactness assumption is necessary.

EXAMPLE 4.6. (1) The sequence of nested open intervals $I_n = (0, \frac{1}{n})$, $n \in \mathbb{Z}^+$, with

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots.$$

Then $\cap_{n\in\mathbb{Z}^+} I_n = \emptyset$.

(2) The sequence of nested closed intervals (which will be proved to be compact sets by the Heine–Borel theorem) $\bar{I}_n = [0, \frac{1}{n}]$, $n \in \mathbb{Z}^+$, with

$$\bar{I}_1 \supseteq \bar{I}_2 \supseteq \bar{I}_3 \supseteq \cdots.$$

Then $\cap_{n\in\mathbb{Z}^+} \bar{I}_n = \{0\} \neq \emptyset$.

4.2. The Heine–Borel theorem. We now prove the other direction of the Heine–Borel theorem.

We have seen that, in $\mathbb{R}$, for any two real numbers $a \leq b$, the closed interval $[a, b]$ is bounded and closed. Similarly, it is not hard to prove that in $\mathbb{R}^2$, for any two pairs of real numbers $a_1 \leq b_1$, $a_2 \leq b_2$, the 2-cell $[a_1, b_1] \times [a_2, b_2]$ is bounded and closed. In this section, we focus on $\mathbb{R}$, but all results work for any $\mathbb{R}^2$ and in fact, for any $\mathbb{R}^n$.

LEMMA 4.7. *For any sequence of nonempty closed intervals in $\mathbb{R}$ with*

$$\bar{I}_1 \supseteq \bar{I}_2 \supseteq \bar{I}_3 \cdots \supseteq \cdots,$$

their intersection $\cap_{n=1}^{\infty} \bar{I}_n$ is a nonempty closed interval.

PROOF. Denote by $\bar{I}_n = [a_n, b_n]$, $n \in \mathbb{Z}^+$. Consider the subset

$$L := \{a_n | n \in \mathbb{Z}^+\} \subseteq \mathbb{R}.$$

It has upper bound b_1 so must have the least upper bound, which we denote by a_∞. Similarly, the subset

$$R := \{b_n | n \in \mathbb{Z}^+\} \subseteq \mathbb{R}$$

has the greatest lower bound, which we denote by b_∞.

Introduction to Analysis Rui Wang Draft

We now show that $a_\infty \le b_\infty$. Otherwise, if $b_\infty < a_\infty$, then because a_∞ is the least upper bound, there exists some $a_n > b_\infty$.

Notice that b_∞ is the greatest lower bound, there exists some $b_{n'}$ so that $b_{n'} < a_n$.

Further notice (b_n) is decreasing, we can assume $n' > n$ and get the contradiction to the increasing property of (a_n) from the inequality

$$a_{n'} \le b_{n'} < a_n.$$

Thus we proved $a_\infty \le b_\infty$. It follows the closed interval $[a_\infty, b_\infty]$ is nonempty and lives in the intersection of all such intervals by the construction.

Further, any point x that lives in the intersection should satisfy

$$a_n \le x \le b_n.$$

It follows $a_\infty \le x \le b_\infty$.

Above all, we proved

$$[a_\infty, b_\infty] = \cap_{n=1}^{\infty} [a_n, b_n].$$

□

PROPOSITION 4.8. *For any two real numbers $a < b$, the closed interval $[a, b] \subseteq \mathbb{R}$ is compact.*

PROOF. Denote by $\bar{I}_0 = [a, b]$. Assume $\bar{I}_0$ is not compact. Then there exists some open cover

$$\{U_\alpha | \alpha \in \Lambda\}$$

which doesn't have any finite subcover of $\bar{I}_0$. Notice that $\bar{I}_0$ is the union of two closed intervals

$$[a, \frac{a+b}{2}] \cup [\frac{a+b}{2}, b],$$

this says at least one of these two closed intervals has no finite subcover. Choose one with no finite subcover and denote it by $\bar{I}_2$.

Then we repeat this construction and obtain a sequence of nested nonempty closed intervals

$$\bar{I}_1 \supset \bar{I}_2 \supset \bar{I}_3 \cdots \supset \cdots.$$

By Lemma 4.7, their intersection is a nonempty closed interval, which we denote by $[a_\infty, b_\infty]$.

On the other hand, notice that

$$0 \le b_\infty - a_\infty \le b_n - a_\infty \le b_n - a_n = \frac{b-a}{2^n},$$

and $\lim_{n\to\infty} \frac{b-a}{2^n} = 0$. This shows that $a_\infty = b_\infty$ and the intersection is a point $x_0 \in [a, b]$.

Assume $x_0 \in U_{\alpha_0}$ with some $\alpha_0 \in \Lambda$. Since U_{α_0} is open, there exits some $r > 0$ so that

$$(x_0 - r, x_0 + r) \subseteq U_\alpha.$$

Then since for any point $x \in \bar{I}_n$,

$$|x - x_0| \le b_n - a_n = \frac{b-a}{2^n},$$

we can take n large enough so that $\frac{b-a}{2^n} < r$, and then it follows

$$\bar{I}_n \subseteq (x_0 - r, x_0 + r) \subseteq U_\alpha.$$

This contradicts with the construction of $\bar{I}_n$ that we assume it has no finite subcover from $\{U_\alpha | \alpha \in \Lambda\}$.

□

Now we are ready to finish the proof of the Heine–Borel theorem.

Introduction to Analysis Rui Wang Draft

PROOF OF THE "IF" PART OF THE HEINE–BOREL. Suppose K is a nonempty bounded closed subset of $\mathbb{R}$.

First, since K is bounded, there exists some closed interval $[a, b] \supset K$ with $a, b \in \mathbb{R}$, $a \leq b$. Then use Proposition 4.8 and Corollary 3.7, it follows K is compact. □

COROLLARY 4.9. *[The Weierstrass theorem] Each bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

PROOF. Assume $\{x_n\}$ is bounded sequence and bounded by a closed interval $[a, b]$ with $a, b \in \mathbb{R}$, $a \leq b$. Then from the Heine-Borel theorem, $[a, b]$ is compact, and hence sequentially compact. It follows $\{x_n\}$ must have a convergent subsequence. □

EXAMPLE 4.10.

(1) For any sequence $\{x_n\}$ in $\mathbb{R}$ with
$$x_1 \leq x_2 \leq x_3 \leq \cdots,$$
the limit exists if and only if $\{x_n\}$ has an upper bound, and when it has upper bound, the limit is $\sup\{x_n\}$.

(2) For any sequence $\{x_n\}$ in $\mathbb{R}$ with
$$x_1 \geq x_2 \geq x_3 \geq \cdots,$$
the limit exists if and only if $\{x_n\}$ has a lower bound, and when it has lower bound, the limit is $\inf\{x_n\}$.

Before concluding this chapter, let us note that the Heine–Borel theorem, together with Proposition 4.2 and the first part of the proof of Theorem 3.8, provides a complete proof of Theorem 3.8 in the case of $\mathbb{R}$ (and in fact for any $\mathbb{R}^n$). The proof for a general metric space, however, is left as a homework problem to the reader.

Introduction to Analysis
Rui Wang
Draft

Introduction to Analysis
Rui Wang
Draft

CHAPTER 3

Numerical Sequences and Series

In this chapter, we will study the properties of sequences and series in the Euclidean metric space $\mathbb{R}$. Unless otherwise stated, all discussions take place in $\mathbb{R}$ equipped with the standard Euclidean metric. The only exception is when we introduce Cauchy sequences, which are defined in any metric space and thus apply to a much broader setting.

1. Sequences in $\mathbb{R}$

1.1. Review of convergent sequences in $\mathbb{R}$. We now recall the main definitions and properties of sequences in $\mathbb{R}$. Although stated here for $\mathbb{R}$, all of them remain valid in any metric space. (These definitions already appeared in Section 2.4; we restate them here for convenience.)

DEFINITION 1.1. (1) A **sequence** in $\mathbb{R}$, denoted by $(x_n)_{n\in\mathbb{Z}^+}$, is a map

$$x : \mathbb{Z}^+ \to \mathbb{R}, \quad n \mapsto x_n.$$

We call the image of this map the *range* of the sequence.

(2) A **subsequence** of $(x_n)_{n\in\mathbb{Z}^+}$ is a sequence $(x_{n_k})_{k\in\mathbb{Z}^+}$ obtained by composing with a strictly increasing function

$$s : \mathbb{Z}^+ \to \mathbb{Z}^+, \qquad n_k := s(k),$$

so that $n_1 < n_2 < \cdots$.

(3) A sequence $(x_n)_{n\in\mathbb{Z}^+}$ in $\mathbb{R}$ is said to **converge** to $x \in \mathbb{R}$ if for every $\epsilon > 0$, there exists $N \in \mathbb{Z}^+$ such that whenever $n > N$, we have

$$|x_n - x| < \epsilon.$$

We write this as $\lim_{n\to\infty} x_n = x$ or simply $x_n \to x$.

(4) If a sequence $(x_n)_{n\in\mathbb{Z}^+}$ in $\mathbb{R}$ does not converge to any real number, it is called **divergent**.

PROPOSITION 1.2 (Basic properties of sequences). *Let $\{x_n\}$ be a sequence in $\mathbb{R}$. Then:*

(1) *The limit of a convergent sequence is unique.*

(2) *$(x_n)_{n\in\mathbb{Z}^+}$ converges to x if and only if every open interval centered at x contains all but finitely many terms of the sequence.*

(3) *$(x_n)_{n\in\mathbb{Z}^+}$ converges to x if and only if every (proper) subsequence of it converges to x.*

(4) *Every convergent sequence in $\mathbb{R}$ is bounded.*

(5) *The set of all subsequential limits of $(x_n)_{n\in\mathbb{Z}^+}$ is closed.*

Here are some typical examples that can be checked using definition easily.

EXAMPLE 1.3 (Typical sequences). (1) Constant sequence: $x_n = c$ converges to c.

(2) Sequence $x_n = \frac{1}{n}$ converges to 0.

(3) The sequence $x_n = n$ and $x_n = n^2$ are unbounded, so divergent.

(4) The sequence $x_n = 1 + \frac{(-1)^n}{n}$ converges to 1.

Introduction to Analysis
Rui Wang
Draft

(5) The sequence $\{x_n = (-1)^n\}$ is bounded but divergent. However it has convergent constant subsequences

$$x_{2k-1} = 1 \quad \text{and} \quad x_{2k} = 1.$$

THEOREM 1.4. *[Weierstrass' theorem 4.9] Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

In the rest of this chapter, we focus on the properties of sequences in $\mathbb{R}$ that essentially rely on the fact that $\mathbb{R}$ is an ordered field.

1.2. Numerical properties of sequences in $\mathbb{R}$.

PROPOSITION 1.5. *Let $(x_n)_{n\in\mathbb{Z}^+}$ and $(y_n)_{n\in\mathbb{Z}^+}$ be convergent sequences in $\mathbb{R}$, with $x_n \to x$ and $y_n \to y$. Then:*

(1) *$(x_n \pm y_n)_{n\in\mathbb{Z}^+}$ converges to $x \pm y$;*
(2) *$(x_n y_n)_{n\in\mathbb{Z}^+}$ converges to xy;*
(3) *If $y \neq 0$, then the sequence $(\frac{x_n}{y_n})_{n\in\mathbb{Z}^+}$ makes sense for large n and it converges to $\frac{x}{y}$.*

PROOF. (1) For every $\epsilon > 0$, since $x_n \to x$ and $y_n \to y$, there exist $N_1, N_2 \in \mathbb{Z}^+$ such that

$$|x_n - x| < \epsilon/2 \quad \text{for all } n > N_1$$

and

$$|y_n - y| < \epsilon/2 \quad \text{for all } n > N_2.$$

Set $N := \max\{N_1, N_2\}$. For $n > N$,

$$|(x_n \pm y_n) - (x \pm y)| \leq |x_n - x| + |y_n - y| < \epsilon.$$

(2) Every convergent sequence is bounded, so there is there exists $M > |y|$ such that $|x_n| \leq M$ for all n.

Now given any $\epsilon > 0$, by the convergence of the two sequences, there are $N_1, N_2 \in \mathbb{Z}^+$ so that

$$|y_n - y| < \epsilon/2M \quad \text{for all } n > N_1$$

and

$$|x_n - x| < \epsilon/2M \quad \text{for all } n > N_2.$$

We estimate

$$\begin{aligned} |x_n y_n - xy| &= |x_n(y_n - y) + y(x_n - x)| \\ &\leq |x_n||y_n - y| + |y||x_n - x| \\ &\leq M \cdot \frac{\epsilon}{2M} + |y| \cdot \frac{\epsilon}{2M} \\ &< M \cdot \frac{\epsilon}{2M} + M \cdot \frac{\epsilon}{2M} \\ &= \epsilon, \end{aligned}$$

for all $n > N := \max\{N_1, N_2\}$.

duction to Analysis Rui Wang Draft

(3) By the previous one, it is enough to show $\frac{1}{y_n} \to \frac{1}{y}$ as $n \to \infty$.
First, since $y \neq 0$ and $y_n \to y$, there exists N_0 such that

$$|y_n - y| < |y|/2 \quad \text{for all } n > N_0,$$

then there follows

$$|y_n| > |y|/2 > 0$$

for such n and so $(\frac{x_n}{y_n})$ makes sense.
Now given $\epsilon > 0$, choose $N \geq N_0$ such that $|y_n - y| < \frac{|y|^2}{2}\epsilon$ for $n > N$.
Then whenever $n > N$,

$$|\frac{1}{y_n} - \frac{1}{y}| = \frac{|y_n - y|}{|y||y_n|} < \frac{|y_n - y|}{|y| \cdot (|y|/2)} = \frac{2}{|y|^2}\,|y_n - y| < \epsilon.$$

□

PROPOSITION 1.6 (Squeeze theorem). *Let $(x_n)_{n\in\mathbb{Z}^+}$, $(y_n)_{n\in\mathbb{Z}^+}$, $(z_n)_{n\in\mathbb{Z}^+}$ be sequences in $\mathbb{R}$. Assume there exists $N_0 \in \mathbb{Z}^+$ such that*

$$x_n \leq y_n \leq z_n \quad \text{for all } n \geq N_0,$$

and that $x_n \to a$ and $z_n \to a$. Then there is $y_n \to a$ as $n \to \infty$.

PROOF. For $n \geq N_0$ we have $x_n - a \leq y_n - a \leq z_n - a$, hence

$$|y_n - a| \leq \max\{|x_n - a|, |z_n - a|\}.$$

Given $\epsilon > 0$, choose $N_1, N_2 \in \mathbb{Z}^+$ so that $|x_n - a| < \epsilon$ for $n > N_1$ and $|z_n - a| < \epsilon$ for $n > N_2$.
Set $N := \max\{N_0, N_1, N_2\}$. Then for all $n > N$,

$$|y_n - a| \leq \max\{|x_n - a|, |z_n - a|\} < \epsilon,$$

so $y_n \to a$. □

DEFINITION 1.7. (1) A sequence (x_n) in $\mathbb{R}$ **diverges to** $+\infty$, if for any $M > 0$, there exists some $N \in \mathbb{Z}^+$, so that whenever $n > N$, $x_n \geq M$. We denote it by $\lim_{n\to\infty} x_n = +\infty$ or $x_n \to +\infty$.

(2) A sequence (x_n) in $\mathbb{R}$ **diverges to** $-\infty$, if for any $M > 0$, there exists some $N \in \mathbb{Z}^+$, so that whenever $n > N$, $x_n \leq -M$. We denote it by $\lim_{n\to\infty} x_n = -\infty$ or $x_n \to -\infty$.

A sequence that diverges to $+\infty$ or $-\infty$ is still *divergent* as a sequence in $\mathbb{R}$. The notations $x_n \to +\infty$ and $x_n \to -\infty$ are shorthand for unbounded growth and do not indicate convergence in $\mathbb{R}$, since $+\infty$ and $-\infty$ are not real numbers.

REMARK 1.8. Although one often adjoins the symbols $\pm\infty$ and works with the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, this set does carry a canonical total order extending that of $\mathbb{R}$, and with the corresponding order topology it is compact (indeed homeomorphic to a closed interval). However, there is *no* way to extend the field operations of $\mathbb{R}$ to all of $\overline{\mathbb{R}}$ so as to obtain a ring (or an ordered field): the usual "extended arithmetic" is only partially defined and breaks algebraic laws (e.g. $\infty - \infty$, $0 \cdot \infty$, ∞/∞ are undefined).

PROPOSITION 1.9. *Let $(x_n)_{n\in\mathbb{Z}^+}$ and $(y_n)_{n\in\mathbb{Z}^+}$ be real sequences. If there exists $N_0 \in \mathbb{Z}^+$ such that $x_n \leq y_n$ for all $n \geq N_0$, and $x_n \to +\infty$, then $y_n \to +\infty$.*

Introduction to Analysis Rui Wang Draft

PROOF. Given $M > 0$, choose $N \geq N_0$ so that $x_n \geq M$ for all $n \geq N$ (since $x_n \to +\infty$). Then for $n \geq N$, $y_n \geq x_n \geq M$. Hence $y_n \to +\infty$. □

EXAMPLE 1.10. (1) For any $p \in \mathbb{Z}^+$, $n^p \to +\infty$ as $n \to \infty$.

(2) More generally, if p is a real polynomial of positive degree with leading coefficient $a_d \neq 0$, then $p(n) \to \operatorname{sgn}(a_d) \cdot \infty$ as $n \to \infty$.

DEFINITION 1.11 (One-sided convergence). Let $(x_n)_{n\in\mathbb{Z}^+}$ be a real sequence and $a \in \mathbb{R}$.

(1) We write $x_n \to a^+$ if for every $\epsilon > 0$ there exists $N \in \mathbb{Z}^+$ such that $0 \leq x_n - a < \epsilon$ for all $n > N$.

(2) We write $x_n \to a^-$ if for every $\epsilon > 0$ there exists $N \in \mathbb{Z}^+$ such that $0 \leq a - x_n < \epsilon$ for all $n > N$.

In either case, $x_n \to a$ in the usual sense.

PROPOSITION 1.12. *Let $(x_n)_{n\in\mathbb{Z}^+}$ be a sequence with $x_n \neq 0$ for all n.*

(1) *$x_n \to +\infty$ if and only if $\frac{1}{x_n} \to 0^+$.*

(2) *$x_n \to -\infty$ if and only if $\frac{1}{x_n} \to 0^-$.*

SKETCH. If $x_n \to +\infty$, then eventually $x_n > 0$ and $|1/x_n| < \epsilon$ for large n, hence $1/x_n \to 0^+$; similarly for $-\infty$. Conversely, if $1/x_n \to 0^+$, then eventually $x_n > 0$ and unbounded, so $x_n \to +\infty$; the 0^- case is analogous. □

1.3. Some useful examples of convergent sequences in $\mathbb{R}$.

EXAMPLE 1.13. (1) For any $\alpha > 0$, we have $n^\alpha \to +\infty$ as $n \to \infty$; equivalently, $\frac{1}{n^\alpha} \to 0$.

PROOF. We prove $\frac{1}{n^\alpha} \to 0$; the statement $n^\alpha \to +\infty$ is equivalent.

- $\alpha \in \mathbb{Z}^+$.
  For $n \geq 1$ and any integer $\alpha \geq 1$,
  $$0 \leq \frac{1}{n^\alpha} \leq \frac{1}{n}.$$
  Since $\frac{1}{n} \to 0$, by comparison $\frac{1}{n^\alpha} \to 0$.
- $\alpha \in \mathbb{Q}^+$.
  Write $\alpha = \frac{k}{\ell}$ with $k, \ell \in \mathbb{Z}^+$. From the previous one, we have $\frac{1}{n^k} \to 0$. Let $a_n := \frac{1}{n^k} \geq 0$. Fix $\epsilon > 0$ and choose N such that $a_n < \epsilon^\ell$ for all $n > N$. Then, for $n > N$,
  $$\frac{1}{n^{k/\ell}} = a_n^{1/\ell} < \epsilon,$$
  so $\frac{1}{n^\alpha} \to 0$.
- $\alpha \in \mathbb{R}^+$.
  Choose a rational $q \in \mathbb{Q}$ with $0 < q < \alpha$ (by density of $\mathbb{Q}$ in $\mathbb{R}$). For all $n \geq 1$, $n^\alpha \geq n^q$, hence
  $$0 \leq \frac{1}{n^\alpha} \leq \frac{1}{n^q}.$$
  By the previous one, $\frac{1}{n^q} \to 0$, and the squeeze theorem yields $\frac{1}{n^\alpha} \to 0$.

□

(2) For any $|\alpha| < 1$, $\alpha^n \to 0$.

Introduction to Analysis Rui Wang Draft

PROOF. Consider $|\frac{1}{\alpha^n}| = (\frac{1}{|\alpha|})^n$. Denote by $a = \frac{1}{|\alpha|}$, it is greater than 1, so can be written as $1 + c$ with $c > 0$. Calculate

$$a^n = (1 + c)^n \geq nc.$$

So $a^n \to +\infty$. This shows $\frac{1}{a^n} \to 0$ and hence $\alpha^n \to 0$. □

(3) $n^{\frac{1}{n}} \to 1$.

PROOF. Denote by $x_n = n^{\frac{1}{n}} - 1$. Notice it is positive whenever $n \geq 2$. We show $x_n \to 0$. Consider the following inequality

$$n = (1 + x_n)^n \geq \frac{n(n-1)}{2} x_n.$$

It follows

$$0 \leq x_n \leq \frac{2}{n-1},$$

which shows $x_n \to 0$ by the squeeze theorem. □

(4) For any $\alpha > 0$, $\alpha^{\frac{1}{n}} \to 1$.

PROOF. If $\alpha \geq 1$, for large enough $n > \alpha$, we have

$$1 \leq \alpha^{\frac{1}{n}} \leq n^{\frac{1}{n}}.$$

Then by the squeeze theorem, $\alpha^{\frac{1}{n}} \to 1$.

If $0 < \alpha < 1$, then $\frac{1}{\alpha} > 1$ and $\alpha^{\frac{1}{n}} = \frac{1}{(\frac{1}{\alpha})^{\frac{1}{n}}} \to \frac{1}{1} = 1$. □

(5) For any $\alpha \in \mathbb{R}$ and $\beta > 0$, $\frac{n^\alpha}{(1+\beta)^n} \to 0$.

PROOF. If $\alpha \leq 0$ the claim is immediate: when $\alpha < 0$, $n^\alpha \to 0$ while $(1 + \beta)^n \to \infty$; when $\alpha = 0$, the numerator is 1 and the denominator diverges to infinity.

Assume now $\alpha > 0$. Choose an integer $k > \alpha$. Use the binomial expansion,

$$(1 + \beta)^n = \sum_{j=0}^{n} \binom{n}{j} \beta^j \geq \binom{n}{k} \beta^k.$$

For $n \geq 2k$ we have

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \geq \frac{(n/2)^k}{k!},$$

since each factor $n - j \geq n/2$ for $0 \leq j \leq k - 1$. Hence, for all $n \geq 2k$,

$$(1 + \beta)^n \geq \frac{\beta^k}{2^k k!} n^k.$$

Therefore,

$$\frac{n^\alpha}{(1+\beta)^n} \leq \frac{2^k k!}{\beta^k} n^{\alpha-k} \quad \text{for } n \geq 2k.$$

Since $\alpha - k < 0$, the right-hand side tends to 0 as $n \to \infty$, which proves the claim by the squeeze theorem. □

Introduction to Analysis Rui Wang Draft

2. Cauchy sequences and the completeness of $\mathbb{R}$

DEFINITION 2.1. A sequence $(x_n)_{n\in\mathbb{Z}^+}$ in a metric space (X,d) is called a **Cauchy sequence**, if for any $\epsilon>0$, there exists some $N\in\mathbb{Z}^+$ so that

$$d(x_m,x_n)<\epsilon \quad \text{for all } m,n>N.$$

This is equivalent to say, for any $\epsilon>0$, there exists some $N\in\mathbb{Z}^+$ so that

$$d(x_{n+p},x_n)<\epsilon \quad \text{for all } n>N, p>0.$$

LEMMA 2.2. *Any convergent sequence is a Cauchy sequence.*

PROOF. Assume (x_n) is a convergent sequence that converges to $x\in X$. Then for any $\epsilon>0$, there exists some $N\in\mathbb{Z}^+$ so that any $n>N$,

$$d(x_n,x)<\frac{\epsilon}{2}.$$

Now for any $m,m'>N$, there is

$$d(x_m,x_{m'})\le d(x_m,x)+d(x_{m'},x)<\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon.$$

This shows (x_n) is a Cauchy sequence. □

PROPOSITION 2.3. *Any Cauchy sequence in a metric space is bounded.*

PROOF. Assume (x_n) is a Cauchy sequence in a metric space (X,d). Take $\epsilon=1$. There exists some $N\in\mathbb{Z}^+$ so that any $n>N$,

$$d(x_n,x_N)<1.$$

Take $C:=\max_{i=1,\cdots,N}\{d(x_i,x_{N+1}),1\}$. C is a finite number and for each $n\in\mathbb{Z}^+$, there is

$$d(x_n,x_N)\le C.$$

This proves the sequence is bounded. □

PROPOSITION 2.4. *In a metric space, if a Cauchy sequence (x_n) has a convergent subsequence, then it converges to the limit of the subsequence.*

PROOF. Assume the Cauchy sequence (x_n) has a subsequence $(x_{n_k})_{k\in\mathbb{Z}^+}$ with

$$x_{n_k}\to x\in X.$$

Then for any $\epsilon>0$, there exists some $N_1\in\mathbb{Z}^+$ so that any $n_k>N_1$ has

$$d(x_{n_k},x_0)<\frac{\epsilon}{2}.$$

At the same time, since $\{x_n\}$ is a Cauchy sequence, there exists some $N_2\in\mathbb{Z}^+$ so that any $m,m'>N_2$,

$$d(x_m,x_{m'})<\frac{\epsilon}{2}.$$

Take $N=\max\{N_1,N_2\}$. Then for any $n>N$, automatically $n_n>N$. We then have

$$d(x_n,x)\le d(x_n,x_{n_n})+d(x_{n_n},x)\le\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon.$$

This proves $x_n\to x$. □

COROLLARY 2.5. *A Cauchy sequence in a (sequentially) compact subset of a metric space (X,d) is convergent.*

Introduction to Analysis
Rui Wang
Draft

PROOF. Assume $K \subseteq X$ is compact. Then by Theorem 3.8, K is sequentially compact. It follows that any sequence (x_n) in K has a convergent subsequence. By Proposition 2.4, (x_n) is convergent. □

EXAMPLE 2.6. Not every Cauchy sequence in a metric space is convergent.

(1) The sequence
$$x_n = \frac{1}{n}$$
is not convergent in $\mathbb{R} \setminus \{0\}$.
(2) Consider the sequence (x_n) as a sequence of increasing rational numbers with $x_n^2 < 2$. It is a Cauchy sequence, but is not convergent in $\mathbb{Q}$.

THEOREM 2.7. *In $\mathbb{R}$ every Cauchy sequence is convergent. In fact, in $\mathbb{R}$, a sequence is convergent if and only if it is a Cauchy sequence.*

PROOF. We have shown from Lemma 2.2 that every convergent sequence is a Cauchy sequence (in any metric space).

Now assume (x_n) is a Cauchy sequence in $\mathbb{R}$. From Proposition 2.3, it is bounded and so it is a Cauchy sequence in a closed interval $[a, b]$ with some $a, b \in \mathbb{R}$, $a \leq b$.

By the Heine-Borel theorem, $[a, b]$ is compact. Apply Corollary 2.5, (x_n) is convergent. □

DEFINITION 2.8. A metric space (X, d) is called **complete** if every Cauchy sequence is convergent.

By the preceding theorem, $\mathbb{R}$ (with the Euclidean metric) is complete. In contrast, $\mathbb{R} \setminus \{0\}$ and $\mathbb{Q}$ are not complete (e.g. $x_n = 1/n$ is Cauchy in $\mathbb{R} \setminus \{0\}$ but has no limit there; the rational truncations of $\sqrt{2}$ form a Cauchy sequence in $\mathbb{Q}$ with no rational limit).

REMARK 2.9 (Completion*). Every metric space (X, d) has a *completion*: there is a complete metric space $(\widetilde{X}, \widetilde{d})$ and an injective map $\iota : X \to \widetilde{X}$ that *preserves all distances*:
$$\widetilde{d}(\iota(x), \iota(y)) = d(x, y) \quad \text{for all } x, y \in X.$$
In other words, when we view X inside $\widetilde{X}$ via ι, the metric on X is unchanged. The space $\widetilde{X}$ is the "smallest" complete space containing X (unique up to the obvious identification).

Examples: the completion of $\mathbb{Q}$ is $\mathbb{R}$. If we equip the set of continuous functions $C^0([a, b])$ (we will see this soon) with the distance function defined by the L^2 norm
$$\|f\|_2 = \Big(\int_a^b |f(x)|^2 \, dx \Big)^{1/2},$$
then its completion is the **Hilbert space** $L^2([a, b])$.

2.1. Upper and lower limits. In general, a real sequence $(x_n)_{n \in \mathbb{Z}^+}$ needs not converge. To extend the notion of limit, we will define the upper and lower limits (also written lim sup and lim inf). They always exist in the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, and $\lim_{n\to\infty} x_n$ exists in $\overline{\mathbb{R}}$ if and only if $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.

To prepare for upper and lower limits, fix a real sequence $(x_n)_{n \in \mathbb{Z}^+}$ and we define its set of subsequential limits
$$L := \{x \in \mathbb{R} | \text{There is a subsequence } (x_{n_k})_{k \in \mathbb{Z}^+} \text{with} x_{n_k} \to x\}.$$
The set L may be empty (e.g. $x_n = n$), and it is always a closed subset of $\mathbb{R}$ (see Proposition 1.2(5)).

Introduction to Analysis
Rui Wang
Draft

It is convenient to pass to the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and set

$$\hat{L} := \{x \in \overline{\mathbb{R}} | \text{There is a subsequence } (x_{n_k})_{k\in\mathbb{Z}^+} \text{with} x_{n_k} \to x\}.$$

Then $\hat{L}$ is *never* empty: if (x_n) is bounded, Bolzano–Weierstrass gives a real subsequential limit; if it is unbounded above (resp. below), one can choose a subsequence tending to $+\infty$ (resp. $-\infty$).

DEFINITION 2.10. For any real sequence (x_n), its **upper limit** is defined as

$$\limsup_{n\to\infty} x_n := \sup \hat{L},$$

and its **lower limit** is defined as

$$\liminf_{n\to\infty} x_n := \inf \hat{L}.$$

Both values are taken in $\overline{\mathbb{R}}$.

Clearly from the definition that the upper limit and the lower limit always exist and

$$\liminf_{n\to\infty} x_n \le \limsup_{n\to\infty} x_n.$$

PROPOSITION 2.11. *The (finite) limit* $\lim_{n\to\infty} x_n$ *exists in* $\mathbb{R}$ *if and only if*

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n \in \mathbb{R},$$

in which case both equal the limit.

PROOF.

- If $\lim_{n\to\infty} x_n = x \in \mathbb{R}$, then (x_n) is bounded and every subsequence limit is x, and so
$$\hat{L} = \{x\}.$$
It follows
$$\liminf_{n\to\infty} x_n = \sup \hat{L} = x, \quad \limsup_{n\to\infty} x_n = \inf \hat{L} = x.$$
- When $\liminf_{n\to\infty} x_n, \limsup_{n\to\infty} x_n \in \mathbb{R}$, (x_n) is bounded and so $\hat{L}$ doesn't contain $\pm\infty$. For every $y \in \hat{L}$,
$$\liminf_{n\to\infty} x_n \le y \le \limsup_{n\to\infty} x_n.$$
When $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x$, it follows $y = x$. This shows
$$\hat{L} = \{x\},$$
i.e., every subsequence of (x_n) converges to x. Hence $x_n \to x$.

□

This result can also be extended to $x = \pm\infty$ and the proof for these cases is left to the reader.

The following lemma is an exercise to the reader.

LEMMA 2.12. *For any sequence* $(x_n)_{n\in\mathbb{Z}^+}$*, the corresponding set* $\hat{L}$ *contains both* $\limsup_{n\to\infty} x_n$ *and* $\liminf_{n\to\infty} x_n$*. As a corollary,*

(1) *For any* $x < \limsup_{n\to\infty}$*, there exists subsequence so that*
$$x_{n_k} > x, \quad \text{for all } k.$$
(2) *For any* $x > \liminf_{n\to\infty}$*, there exists subsequence so that*
$$x_{n_k} < x, \quad \text{for all } k.$$

Introduction to Analysis Rui Wang Draft

In general, $\limsup_{n\to\infty} x_n$ is NOT an eventual upper bound of (x_n), and $\liminf_{n\to\infty} x_n$ is NOT an eventual lower bound. They capture limiting "cluster levels," not pointwise bounds.

EXAMPLE 2.13. Let $x_n = \frac{(-1)^n}{n}$. Then $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = 0$, yet $x_n > 0$ for infinitely many n and $x_n < 0$ for infinitely many n. So 0 is neither an eventual upper nor an eventual lower bound.

However, there are the following useful lemmas.

LEMMA 2.14. (1) *For any $x > \limsup_{n\to\infty}$, there exists some $N \in \mathbb{Z}^+$ so that*

$$x_n < x, \quad \text{for all } n > N.$$

(2) *For any $x < \liminf_{n\to\infty}$, there exists some $N \in \mathbb{Z}^+$ so that*

$$x_n > x, \quad \text{for all } n > N.$$

A proof by contradiction is perhaps the clearest approach, and we leave the details to the reader.

The following two propositions are also very useful and we leave their proofs as homework problems.

PROPOSITION 2.15. *For any two sequences $(x_n)_{n\in\mathbb{Z}^+}$, $(y_n)_{n\in\mathbb{Z}^+}$ with $x_n \le y_n$ for every $n > N$ for some fixed $N \in \mathbb{Z}^+$, There are*

$$\begin{aligned} \limsup_{n\to\infty} x_n &\le \limsup_{n\to\infty} y_n \\ \liminf_{n\to\infty} x_n &\le \liminf_{n\to\infty} y_n. \end{aligned}$$

EXERCISE 2.16. Under the same hypothesis (and even when assume $x_n < y_n$), show by example that one cannot conclude

$$\limsup_{n\to\infty} x_n \le \liminf_{n\to\infty} y_n.$$

The upper limit and lower limit can also be defined using the following way. Given a sequence $(x_n)_{n\in\mathbb{Z}^+}$, consider the nested sets

$$S_1 \supseteq S_2 \supseteq S_3 \supseteq \cdots$$

defined by $S_k := \{x_n | n \ge k\}$. Then the least upper bounds of them form a decreasing sequence

$$\sup S_1 \ge \sup S_2 \ge \sup S_3 \ge \cdots,$$

and the greatest lower bounds of them form an increasing sequence

$$\inf S_1 \le \inf S_2 \le \inf S_3 \le \cdots.$$

It makes sense to define $\lim_{n\to\infty}(\sup\{x_k | k \ge n\})$ and $\lim_{n\to\infty}(\inf\{x_k | k \ge n\})$ then.

PROPOSITION 2.17.

$$\begin{aligned} \limsup_{n\to\infty} x_n &= \lim_{n\to\infty}(\sup\{x_k | k \ge n\}) \\ \liminf_{n\to\infty} x_n &= \lim_{n\to\infty}(\inf\{x_k | k \ge n\}). \end{aligned}$$

(In the proof you need to discuss the RHS is a real number and $\pm\infty$ separately.)

Introduction to Analysis Rui Wang Draft

3. Series in $\mathbb{R}$ ($\mathbb{C}$)

3.1. Definitions and examples. Most results in this section also hold for $\mathbb{C}$, but you may restrict yourselves to $\mathbb{R}$ if you are not fully comfortable with complex numbers.

Given a sequence $(x_n)_{n\in\mathbb{Z}^+}$ in $\mathbb{R}$, we define its **sequence of partial sums** $(s_n)_{n\in\mathbb{Z}^+}$ by

$$s_n := \sum_{k=1}^{n} x_k = x_1 + x_2 + \cdots + x_n.$$

We use the notation

$$\sum_{k=1}^{\infty} x_k,$$

called the **(infinite) series**, to denote this sequence of partial sums $(s_n)_{n\in\mathbb{Z}^+}$.

If the sequence $(s_n)_{n\in\mathbb{Z}^+}$ converges to some $s \in \mathbb{R}$, we say that the series $\sum_{k=1}^{\infty} x_k$ **converges** to s, and write

$$\sum_{k=1}^{\infty} x_k = s.$$

If the sequence $(s_n)_{n\in\mathbb{Z}^+}$ diverges, we say that the series $\sum_{k=1}^{\infty} x_k$ **diverges**. In particular, if $s_n \to +\infty$ (resp. $-\infty$), we say that the series diverges to $+\infty$ (resp. $-\infty$), and write

$$\sum_{k=1}^{\infty} x_k = +\infty \quad (\text{resp. } -\infty).$$

When we are concerned only with the convergence of the series, rather than the concrete value of its limit, it is common to omit the indices and write it simply as $\sum x_n$.

A series $\sum_{n=1}^{\infty} x_n$ carries the same information as the sequence $(x_n)_{n\in\mathbb{Z}^+}$, since

$$x_n = s_n - s_{n-1}, \quad \text{where we set } s_0 = 0.$$

Before we proceed to the general theory of series, we begin with a few illustrative examples.

EXAMPLE 3.1. (1) The sum of natural numbers $\sum_{n=1}^{\infty} n$. Since the partial sums

$$s_n = \sum_{k=1}^{n} k = 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \to +\infty \quad \text{as } n \to \infty,$$

this series diverges to $+\infty$.

(2) What is the value of the series

$$(-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + \cdots ?$$

If one groups the terms as

$$((-1) + 1) + ((-1) + 1) + ((-1) + 1) + \cdots,$$

it appears that each pair sums to 0, so the result would be 0. However, if we group them differently:

$$(-1) + (1 + (-1)) + (1 + (-1)) + (1 + (-1)) + \cdots,$$

then we seem to obtain

$$(-1) + 0 + 0 + 0 + \cdots = -1.$$

So what happened here?

Introduction to Analysis Rui Wang Draft

The issue is that this series does NOT converge at all — its partial sums oscillate between -1 and 0:

$$s_1 = -1, \quad s_2 = 0, \quad s_3 = -1, \quad s_4 = 0, \text{ and so on.}$$

Hence, $\sum_{n=1}^{\infty}(-1)^n$ diverges.

In fact, rearranging or regrouping terms is only legitimate when a series is *absolutely convergent* (a term we will explain later).

(3) If one eats half of the remaining cake each day, then after infinitely many days, the entire cake will have been eaten — corresponding to the convergent series

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

More generally, for $\alpha \in \mathbb{R}$, the partial sums of the series $\sum_{n=1}^{\infty} \alpha^n$ can be calculated as

$$s_n = \sum_{k=1}^{n} \alpha^k = \frac{\alpha(1-\alpha^n)}{1-\alpha}, \quad \text{when } \alpha \neq 1$$

and

$$s_n = n \quad \text{when } \alpha = 1.$$

For these, one obtains

$$\sum_{n=1}^{\infty} \alpha^n = \frac{\alpha}{1-\alpha}, \quad \text{if } |\alpha| < 1;$$

and

$$\sum_{n=1}^{\infty} \alpha^n \text{ diverges, if } |\alpha| \geq 1.$$

(4) For $p > 0$, the series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

is called the ***p*-series**. We will show that it converges when $p > 1$ and diverges to $+\infty$ when $0 < p \leq 1$.

REMARK 3.2. The p-series is closely related to the **Riemann zeta function**, defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

for complex numbers s with real part $\Re(s) > 1$.

For even integers $p = 2k$ with $k \in \mathbb{Z}^+$, the values of the Riemann zeta function can be expressed explicitly in terms of the Bernoulli numbers B_{2k}:

$$\zeta(2k) = \sum_{n=1}^{\infty} \frac{1}{n^{2k}} = (-1)^{k+1} \frac{B_{2k}}{2(2k)!}(2\pi)^{2k}.$$

Here B_{2k} denotes the $2k$-th Bernoulli number. For example,

$$\zeta(2) = \frac{\pi^2}{6}, \qquad \zeta(4) = \frac{\pi^4}{90}.$$

For odd integers p, much less is known. In 1978, Apéry proved the remarkable result that $\zeta(3)$ is irrational. It remains an open problem whether $\zeta(5)$, $\zeta(7)$, and other odd zeta values are rational or irrational.

duction to Analysis
Rui Wang
Draft

The Riemann zeta function also plays a central role in the famous **Riemann Hypothesis**, one of the most important open problems in mathematics.

3.2. Basic tests for convergence and divergence. We now introduce some basic tools for determining whether a series is convergent.

PROPOSITION 3.3. *Let $\sum x_n$ be a series with nonnegative terms, i.e. $x_n \geq 0$ for every $n \in \mathbb{Z}^+$. Then the series $\sum x_n$ converges if and only if its sequence of partial sums (s_n) is bounded above.*

PROOF. Observe that in this case

$$s_n = s_{n-1} + x_n \geq s_{n-1},$$

since $x_n \geq 0$ for all n. Hence, (s_n) is an increasing sequence in $\mathbb{R}$.

For an increasing sequence of real numbers, the limit exists if and only if it is bounded above. □

THEOREM 3.4. *[Cauchy's Criterion] A series $\sum x_n$ is convergent if and only if for any $\epsilon > 0$, there exists $N \in \mathbb{Z}^+$ so that*

$$|\Sigma_{k=n}^{n+p} x_k| < \epsilon, \quad \text{for all } n > N, \quad p \geq 0.$$

This is equivalent to say $\lim_{n\to\infty} \Sigma_{k=n}^{n+p} x_k = 0$ for any $p \geq 0$.

In particular, by taking $p = 0$, we obtain the following corollary.

COROLLARY 3.5. *For any convergent series Σx_n, there is*

$$\lim_{n\to\infty} x_n = 0.$$

PROOF OF PROPOSITION 3.4. Because $\mathbb{R}$ is complete, the series $\sum x_n$ is convergent if and only if it is a Cauchy sequence. Notice that

$$|\Sigma_{k=n}^{n+p} x_k| = |S_{n+p} - S_{n-1}|.$$

The theorem follows from the definition of Cauchy sequence. □

EXAMPLE 3.6. (1) The convergence of $\sum x_n$ implies that $x_n \to 0$, but the converse is not true in general. For example, we will show later $\sum \frac{1}{n}$ diverges, though $\frac{1}{n} \to 0$.

(2) The series $\sum \frac{n}{n+1}$ diverges, since the sequence $(\frac{n}{n+1})$ doesn't converge to zero.

The following statement gives a way to check if a series with alternating signs converges.

PROPOSITION 3.7. *If the series $\sum_{n=1}^{\infty} (-1)^n x_n$ with*

(1) *$x_n \geq 0$;*
(2) *x_n decreases, i.e, $x_1 \geq x_2 \geq \cdots$;*
(3) *$x_n \to 0$ as $n \to \infty$,*

then it is convergent.

PROOF. By Cauchy's criterion, we want to show the following sum converges to zero as $n \to \infty$

$$|\Sigma_{k=n}^{n+p} (-1)^k x_k| = |x_n - x_{n+1} + \cdots + \pm x_{n+p}|.$$

Notice that the condition (1) and (2) guarantee that

$$0 \leq x_n - x_{n+1} + \cdots + \pm x_{n+p} \leq x_n.$$

Introduction to Analysis Rui Wang Draft

Further by the condition (3) $x_n \to 0$, there follows

$$\lim_{n\to\infty} |\Sigma_{k=n}^{n+p}(-1)^k x_k| = 0$$

for all $p \geq 0$. Then Cauchy's criterion shows that $\Sigma(-1)^n x_n$ is convergent. □

EXAMPLE 3.8. By the above proposition, the series $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ is convergent .

DEFINITION 3.9. (1) A series Σx_n is called **absolutely convergent**, if $\Sigma|x_n|$ is convergent.
(2) A series Σx_n is called **conditionally convergent**, if Σx_n is convergent and $\sum |x_n|$ is divergent.

The following lemma explains the above definition.

LEMMA 3.10. *If $\Sigma|x_n|$ is absolutely convergent, then Σx_n is convergent.*

PROOF. This follows from the triangle inequality

$$|\Sigma_{k=n}^{n+p} x_k| \leq \Sigma_{k=n}^{n+p}|x_k|,$$

the squeeze theorem and Cauchy's criterion. □

EXAMPLE 3.11. The series $\sum \frac{(-1)^n}{n^p}$ with $p > 1$, converges absolutely; the series $\sum \frac{(-1)^n}{n^p}$ with $0 < p \leq 1$ converges conditionally.

3.3. Comparison tests.

THEOREM 3.12 (Comparison test). *Consider two series $\sum x_n$ and $\sum y_n$.*
(1) *If $|x_n| \leq y_n$ for all n and Σy_n converges, then Σx_n converges absolutely.* .
(2) *If $x_n \leq y_n$ for all n and $\Sigma x_n = +\infty$, then $\Sigma y_n = +\infty$.*
(3) *If $x_n \leq y_n$ for all n and $\Sigma y_n = -\infty$, then $\Sigma x_n = -\infty$.*

PROOF. (1) It follows from the inequalities

$$|\Sigma_{k=n}^{n+p} x_k| \leq \Sigma_{k=n}^{n+p}|x_k| \leq \Sigma_{k=n}^{n+p} y_k = |\Sigma_{k=n}^{n+p} y_k|,$$

the squeeze theorem and Cauchy's criterion.
(2) (3) both follow from the inequality of partial sums

$$\Sigma_{k=1}^{n} x_n \leq \Sigma_{k=1}^{n} y_n.$$

$\Sigma_{k=1}^{n} x_n \to +\infty$ implies that $\Sigma_{k=1}^{n} y_n \to +\infty$; and $\Sigma_{k=1}^{n} y_n \to -\infty$ implies that $\Sigma_{k=1}^{n} x_n \to -\infty$.

□

EXAMPLE 3.13. (1) $\Sigma \frac{1}{n^2+5}$ is convergent by the comparison test:

$$0 \leq \frac{1}{n^2+5} \leq \frac{1}{n^2}$$

and the fact that $\Sigma \frac{1}{n^2}$ is convergent (will prove).
(2) $\Sigma \frac{n}{2^n}$ is convergent by the comparison test:

$$0 \leq \frac{n}{2^n} \leq \frac{1.5^n}{2^n} = 0.75^n$$

and the convergence of the geometric series $\Sigma 0.75^n$.

As an application of the comparison test, we discuss the convergence of p-series.

Introduction to Analysis Rui Wang Draft

THEOREM 3.14. *The series* $\sum \frac{1}{n^p}$ *converges if* $p > 1$*; and diverges to* $+\infty$ *if* $p \leq 1$.

We will prove the theorem by the following lemma, whose proof can be seen from the following picture: (insert picture.)

LEMMA 3.15. *A series* $\sum x_n$ *with*

$$x_1 \geq x_2 \geq \cdots \geq 0,$$

is convergent, if and only if the series

$$\sum_{k=0}^{\infty} 2^k x_{2^k} = x_1 + 2x_2 + 4x_4 + 8x_8 + \cdots$$

is convergent.

PROOF. (1) For any $N \in \mathbb{Z}^+$, there exists unique K so that $2^{K-1} < N \leq 2^K$. Then it follows

$$\sum_{n=1}^{N} x_n \leq \sum_{k=1}^{K} 2^k x_{2^k}.$$

This shows when $\sum_{k=0}^{\infty} 2^k x_{2^k}$ is convergent, the series $\sum x_n$ is convergent.

(2) For any $K \in \mathbb{Z}^+$, take an $N > 2^K$. Then

$$\frac{1}{2}\Sigma_{k=1}^{K} 2^k x_{2^k} \leq \Sigma_{n=1}^{N} x_n.$$

This shows when Σx_n is convergent, the series $\Sigma_{k=0}^{\infty} 2^k x_{2^k}$ is convergent.

□

PROOF OF THE THEOREM. If $p \leq 0$, $\Sigma \frac{1}{n^p}$ is divergent since $\frac{1}{n^p}$ is not convergent to zero.

Now assume $p > 0$. Then the series $\Sigma \frac{1}{n^p}$ satisfies the assumption of the lemma. So we can look at the series $\Sigma_{k=0}^{\infty} 2^k x_{2^k}$ instead. We have

$$\Sigma_{k=0}^{\infty} 2^k x_{2^k} \quad = \quad \Sigma_{k=0}^{\infty} \frac{2^k}{2^{kp}} = \Sigma_{k=0}^{\infty} (\frac{1}{2^{p-1}})^k,$$

which is convergent if and only if $p > 1$. □

3.4. Root and ratio tests.

THEOREM 3.16 (Root test). *For a seires* Σx_n*, define*

$$\alpha := \limsup \sqrt[n]{|x_n|} \in \mathbb{R} \cup \{+\infty\}.$$

(1) *If* $\alpha < 1$*, then* Σx_n *is absolutely convergent;*
(2) *If* $\alpha > 1$*, then* Σx_n *is divergent.*

PROOF. (1) When $\alpha < 1$, we can pick and fix some β so that $\alpha < \beta < 1$. Then since

$$\alpha = \limsup \sqrt[n]{|x_n|} < \beta,$$

there exists $N \in \mathbb{Z}^+$ so that

$$\sqrt[n]{|x_n|} < \beta, \quad \text{for all } n > N.$$

It follows $|x_n| < \beta^n$. Notice that $\Sigma \beta^n$ is convergent, by the comparison test, Σx_n converges absolutely.

Introduction to Analysis Rui Wang Draft

(2) When $\alpha = \limsup \sqrt[n]{|x_n|} > 1$, there exists subsequence $(\sqrt[n_k]{|x_{n_k}|})_{k\in\mathbb{Z}^+}$ so that $\sqrt[n_k]{|x_{n_k}|} \geq 1$. It follows $|x_{n_k}| \geq 1$, and then divergent. As a result, $\sum x_n$ is divergent.

□

EXAMPLE 3.17. When $\alpha = 1$, root test is failed to detect for convergence or divergence of a series. For example, consider the series $\sum \frac{1}{n^p}$. Calculate

$$\limsup \sqrt[n]{|x_n|} = \limsup \sqrt[n]{\frac{1}{n^p}} = (\frac{1}{\lim \sqrt[n]{n}})^p = 1^p = 1,$$

but the convergence of $\sum \frac{1}{n^p}$ depends on values of p.

EXAMPLE 3.18. (1) $\sum x^n$, where $x \in \mathbb{R}$ (or in $\mathbb{C}$). Compute

$$\alpha = \limsup \sqrt[n]{|x_n|} = |x|.$$

By root test,

(a) when $|x| < 1$, $\sum x^n$ converges absolutely;
(b) when $|x| > 1$, $\sum x^n$ is divergent.
(c) when $|x| = 1$, $\sum x^n$ is divergent too.

The same result holds for complex numbers.

(2) $\sum \frac{x^n}{n}$. Compute

$$\alpha = \limsup \sqrt[n]{|x_n|} = |x|.$$

By root test,

(a) when $|x| < 1$, $\sum \frac{x^n}{n}$ converges absolutely;
(b) when $|x| > 1$, $\sum \frac{x^n}{n}$ is divergent.
(c) when $x = 1$, $\sum \frac{x^n}{n}$ is divergent.
(d) when $x = -1$, $\sum \frac{x^n}{n}$ converges conditionally.

This result can be extended to complex numbers: for any x satisfying $|x| = 1$ but $x \neq 1$, the series $\sum \frac{x^n}{n}$ converges conditionally. The proof of this convergence is somewhat delicate; it can be established by an argument similar to the one used in Rudin, Chapter 3, Exercise 8. We leave the details to the interested reader.

(3) $\Sigma\frac{x^n}{n^2}$. Compute

$$\alpha = \limsup \sqrt[n]{|\frac{x^n}{n^2}|} = |x|.$$

By root test,

(a) when $|x| < 1$, Σx_n converges absolutely;
(b) when $|x| > 1$, Σx_n is divergent.
(c) when $|x| = 1$, the absolute series is $\sum \frac{1}{n^2}$, which is convergent. So $\sum \frac{x^n}{n^2}$ converges absolutely.

The same result holds for complex numbers.

THEOREM 3.19 (Ratio test). *For a seires $\sum x_n$ with each $x_n \neq 0$, define*

$$\alpha := \limsup \frac{|x_{n+1}|}{|x_n|}.$$

Then

(1) *if $\alpha < 1$, Σx_n converges absolutely;*

Introduction to Analysis Rui Wang Draft

(2) *if there is* $N \in \mathbb{Z}^+$ *such that all* $n > N$, $\frac{|x_{n+1}|}{|x_n|} \geq 1$, *then* $\sum x_n$ *is divergent. In particular, if*

$$\liminf \frac{|x_{n+1}|}{|x_n|} > 1,$$

then $\sum x_n$ *is divergent.*

PROOF. (1) When $\alpha < 1$, we can find and fix some β so that $\alpha < \beta < 1$. Then since

$$\alpha = \limsup \frac{|x_{n+1}|}{|x_n|} < \beta,$$

there exists some $N \in \mathbb{Z}^+$ so that

$$\frac{|x_{n+1}|}{|x_n|} < \beta, \quad \text{for all } n \geq N.$$

By induction, it follows

$$|x_{N+k}| < \beta^k |x_N|, \quad \text{for } k = 0, 1, 2, \cdots.$$

Notice the series $\sum_{k=0}^{\infty}(\beta^k |x_N|)$ is convergent since $0 < \beta < 1$. By the comparison test, $\sum x_n$ converges absolutely.

(2) If $\frac{|x_{n+1}|}{|x_n|} \geq 1$ for $n \geq N$ for some N, then

$$|x_n| \geq |x_N|, \quad \text{all } n \geq N.$$

Thus (x_n) doesn't converge to 0 and then $\sum x_n$ must be divergent.

In particular, if

$$\liminf \frac{|x_{n+1}|}{|x_n|} > 1,$$

it follows $\frac{|x_{n+1}|}{|x_n|} \geq 1$ for all $n > N$, for some $N \in \mathbb{Z}^+$.

□

REMARK 3.20. One can't use $\limsup \frac{|x_{n+1}|}{|x_n|} > 1$ to conclude divergence: Consider the following example, the series $\sum x_n$ with

$$x_{2k-1} = \frac{2}{k^2}, \quad x_{2k} = \frac{1}{k^2}.$$

It is not hard to check that $\limsup \frac{|x_{n+1}|}{|x_n|} = 2 > 1$, but

$$\sum x_n = \frac{2}{1} + \frac{1}{1} + \frac{2}{2^2} + \frac{1}{2^2} + \frac{2}{3^2} + \frac{1}{3^2} + \frac{2}{4^2} + \frac{1}{4^2} + \cdots$$

is convergent (by comparison with $\sum \frac{1}{n^2}$).

Here is an example of ratio test.

EXAMPLE 3.21. $\sum \frac{x^n}{n!}$. Compute

$$\alpha = \limsup \frac{|x_{n+1}|}{|x_n|} = \limsup \frac{|x|}{n+1} = 0$$

for any x.

By the ratio test, for every $x \in \mathbb{R}$ (actually for every $x \in \mathbb{C}$), $\sum \frac{x^n}{n!}$ converges absolutely.

Introduction to Analysis Rui Wang Draft

3.5. The power series and the number e.

DEFINITION 3.22. A series of the form $\sum_{n=0}^{\infty} c_n x^n$, with $c_n \in \mathbb{R}$ (or $c_n \in \mathbb{C}$), is called a **power series**.

Using the root test, we have the following result.

THEOREM 3.23. *Given a power series $\sum_{n=0}^{\infty} c_n x^n$, define*

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}} \in [0, +\infty].$$

(When $\limsup_{n\to\infty} \sqrt[n]{|c_n|} = 0$, define $R = +\infty$; when $\limsup_{n\to\infty} \sqrt[n]{|c_n|} = +\infty$, define $R = 0$.)

The series $\sum_{n=0}^{\infty} c_n x^n$ is absolutely convergent when $|x| < R$; and divergent when $|x| > R$.

*R is called the **convergence radius** of the power series $\sum_{n=0}^{\infty} c_n x^n$.*

We have seen that the power series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ has convergence radius $R = +\infty$. This says $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges absolutely for every $x \in \mathbb{R}$ (in fact for every $x \in \mathbb{C}$). This defines a function, which is called the **exponential function**

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

In particular, when $x = 1$, the real number

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots,$$

is called the **Euler's number** e.

We can quickly estimate that $2.5 < e < 3$ since

$$2.5 < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots \leq 1 + 1 + \sum_{n=1}^{\infty} \frac{1}{2^n} = 3.$$

In fact, we have the following better estimate as

$$\begin{aligned} e - \sum_{n=1}^{N} \frac{1}{n!} &= \sum_{n=N+1}^{\infty} \frac{1}{n!} \\ &= \frac{1}{(N+1)!} + \frac{1}{(N+2)!} + \cdots \\ &= \frac{1}{(N+1)!}\left(1 + \frac{1}{N+2} + \frac{1}{(N+2)(N+3)} + \cdots\right) \\ &\leq \frac{1}{(N+1)!}\left(1 + \frac{1}{N+2} + \frac{1}{(N+2)^2} + \cdots\right) \\ &\leq \frac{1}{N \cdot N!}. \end{aligned}$$

When $N = 10$, $\frac{1}{N \cdot N!} < 10^{-7}$, which says that error lies between

$$1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{10!}$$

and the precise value of e is smaller than 0.0000001. ($e \approx 2.71828182845904523536028747135 27\cdots$.)

We now give another interpretation of e, illustrating its natural appearance in the context of continuously compounded interest.

PROPOSITION 3.24. *The sequence $((1 + \frac{1}{n})^n)_{n\in\mathbb{Z}^+}$ converges to e.*

Introduction to Analysis
Rui Wang
Draft

PROOF. Denote by $S_n = \sum_{k=0}^{n} \frac{1}{k!}$.

(1) We estimate

$$\begin{aligned}(1+\frac{1}{n})^n &= 1+1+\frac{n(n-1)}{2!n^2}+\cdots\frac{n(n-1)\cdots(n-(k-1))}{k!n^k}+\cdots+\frac{n!}{n!n^n}\\ &\le 1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}=S_n.\end{aligned}$$

Take $n\to\infty$, we obtain

$$\limsup_{n\to\infty}(1+\frac{1}{n})^n \le \lim_{n\to\infty} S_n = e.$$

(2) On the other hand, for any $m \le n$, we have

$$\begin{aligned}(1+\frac{1}{n})^n &= 1+1+\frac{n(n-1)}{2!n^2}+\cdots\frac{n(n-1)\cdots(n-(k-1))}{k!n^k}+\cdots\frac{n!}{n!n^n}\\ &\ge 1+1+\frac{n(n-1)}{2!n^2}+\cdots\frac{n(n-1)\cdots(n-(m-1))}{m!n^m}.\end{aligned}$$

Fix m and let $n\to\infty$, we get

$$\begin{aligned}\liminf_{n\to\infty}(1+\frac{1}{n})^n &\ge \lim_{n\to\infty}(1+1+\frac{n(n-1)}{2!n^2}+\cdots\frac{n(n-1)\cdots(n-(m-1))}{m!n^m})\\ &= 1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{m!}=S_m.\end{aligned}$$

Then take $m\to\infty$, we have

$$\liminf_{n\to\infty}(1+\frac{1}{n})^n \ge \lim_{m\to\infty} S_m = e.$$

Combine (1) and (2), we have

$$e \le \liminf_{n\to\infty}(1+\frac{1}{n})^n \le \limsup_{n\to\infty}(1+\frac{1}{n})^n \le e.$$

Hence $\lim_{n\to\infty}(1+\frac{1}{n})^n$ exists and it is e. □

3.6. The algebraic structure of the set of all (absolutely) convergent series. The following proposition is easy to check using definition.

PROPOSITION 3.25. *Assume series $\Sigma a_n = A$, $\Sigma b_n = B$. Then*

(1) $\Sigma(a_n + b_n) = A + B$;
(2) *For any $c \in \mathbb{R}$, $\Sigma(c \cdot a_n) = cA$.*

As a result, the set of all convergent series forms an (infinite-dimensional) vector space over $\mathbb{R}$.

REMARK 3.26. Over this space, one can define a norm

$$\|\sum a_n\| := \sup_N |s_N|,$$

where s_n are partial sums of $\sum a_n$.

This norm defines a distance function on the set of all convergent series, and further using the completeness of $\mathbb{R}$, one can prove that this metric space is complete.

Next, let's consider multiplication (product) of two series.

Introduction to Analysis Rui Wang Draft

DEFINITION 3.27. Given two (formal) series $\sum a_n$ and $\sum b_n$, the series $\sum c_n$ with

$$c_k = \sum_{m,n\geq 0, m+n=k} a_m b_n$$

is defined as the **Cauchy product** of these two series.

EXAMPLE 3.28. Consider $\sum a_n = a_1 + a_2$, $\sum b_n = b_1 + b_2 + b_3$, then

$$(\sum a_n)(\sum b_n) = (a_1 + a_2)(b_1 + b_2 + b_3) = a_1b_1 + (a_1b_2 + a_2b_1) + (a_1b_3 + a_2b_2) + a_2b_3 = \sum c_n,$$

with

$$c_2 = a_1b_1, \quad c_3 = a_1b_2 + a_2b_1, \quad c_4 = a_1b_3 + a_2b_2, \quad c_5 = a_2b_3,$$

and all other terms are zero.

A question we should ask: If both $\sum a_n$ and $\sum b_n$ are convergent, is their product convergent? The following example shows that the answer is NO.

EXAMPLE 3.29. Consider the series

$$\sum a_n = \sum b_n = \sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}.$$

It is convergent by Proposition 3.7.

We compute their product

$$\sum c_k = (\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}})^2,$$

where

$$c_k = \sum_{m,n\geq 1, m+n=k} \frac{(-1)^m}{\sqrt{m}} \frac{(-1)^n}{\sqrt{n}} = \sum_{m,n\geq 1, m+n=k} \frac{(-1)^k}{\sqrt{m}\sqrt{n}}$$

We focus on even k's:

$$\frac{(-1)^k}{\sqrt{m}\sqrt{n}} = \frac{1}{\sqrt{m}\sqrt{n}} \geq \frac{2}{m+n} = \frac{2}{k}, \quad \text{for} \quad m+n=k,$$

and thus

$$c_k = \sum_{m,n\geq 1, m+n=k} \frac{(-1)^k}{\sqrt{m}\sqrt{n}} \geq \sum_{m,n\geq 1, m+n=k} \frac{2}{k} = \frac{2(k-1)}{k}.$$

So for this even subsequence $c_k \to 2$ as $k \to \infty$, and this implies $\sum c_k$ is divergent.

This example shows that in the space of convergent series, multiplication is not well-defined — the product of two convergent series need not be convergent.

On the other hand, if we require that at least one of the two series converges absolutely, then their product is always convergent.

THEOREM 3.30 (Mertens' Theorem). *Suppose that*

(1) $\sum a_n = A$ *and* $\sum b_n = B$ *are convergent series;*
(2) *the series* $\sum a_n$ *converges absolutely.*

Then the Cauchy product $\sum c_n$ *defined by*

$$c_k = \sum_{m,n\geq 0, m+n=k} a_m b_n$$

also converges, and its sum satisfies

$$\sum c_n = AB.$$

Introduction to Analysis Rui Wang Draft

PROOF. We compute the partial sums of $\sum c_n$:

$$\sum_{n=0}^{m} c_n = \sum_{n=0}^{m}\sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{m} a_k \sum_{n=0}^{m-k} b_n = \sum_{k=0}^{m} a_k S^b_{m-k},$$

where S^b_j denotes the j-th partial sum of $\sum b_n$.

We then write

$$\begin{aligned}\sum_{n=0}^{m} c_n - AB &= \sum_{k=0}^{m} a_k S^b_{m-k} - AB \\ &= \sum_{k=0}^{m} a_k (S^b_{m-k} - B) + \left(\sum_{k=0}^{m} a_k - A\right) B.\end{aligned}$$

Taking absolute values gives

$$\left|\sum_{n=0}^{m} c_n - AB\right| \le \sum_{k=0}^{m} |a_k|\,|S^b_{m-k} - B| + \left|\sum_{k=0}^{m} a_k - A\right| |B|.$$

The second term tends to 0 since $\sum a_k = A$.

To show the first term tends to 0, let $\beta_n := S^b_n - B$, so that $\beta_n \to 0$ as $n \to \infty$. Given $\epsilon > 0$, choose N such that $|\beta_n| < \epsilon$ for all $n > N$. Then

$$\begin{aligned}\sum_{k=0}^{m} |a_k|\,|\beta_{m-k}| &= \sum_{k=0}^{m-N-1} |a_k|\,|\beta_{m-k}| + \sum_{k=m-N}^{m} |a_k|\,|\beta_{m-k}| \\ &\le \epsilon \sum_{k=0}^{m-N-1} |a_k| + \max_{0\le j\le N} |\beta_j| \sum_{k=m-N}^{m} |a_k|.\end{aligned}$$

Since $\sum |a_k|$ converges absolutely, $\sum_{k=m-N}^{m} |a_k| \to 0$ as $m \to \infty$. Therefore, for large m,

$$\sum_{k=0}^{m} |a_k|\,|\beta_{m-k}| \le \epsilon\left(\sum_{k=0}^{\infty} |a_k| + \max_{0\le j\le N} |\beta_j|\right),$$

which shows the first term tends to 0 as $m \to \infty$.

Hence $\sum c_n$ converges and $\sum c_n = AB$. □

Regarding products, the next theorem should be contrasted with Mertens' theorem—note the different hypotheses. We leave the proof as an exercise for the interested reader.

THEOREM 3.31. *If $\sum a_n = A$, $\sum b_n = B$ and the product $\sum c_n = C$, where*

$$c_n := \Sigma_{k=0}^{n} a_k b_{n-k},$$

then $C = AB$.

Motivated by the product theorems, it is natural to restrict to *absolutely convergent* series. With respect to the addition and the Cauchy product (in following, denoted by $*$ for the sequences of terms in a series), the set of *absolutely convergent* series forms a commutative ring with identity $1 + 0 + 0 + \cdots$ and also an (infinite-dimensional) vector space over $\mathbb{R}$; in fact it is a unital Banach algebra over $\mathbb{R}$ with the ℓ^1-norm

$$\|(a_n)\|_{\ell^1} := \sum |a_n|.$$

In particular, this is a complete metric space with respect to the distance function

$$d_{\ell^1}((a_n),(b_n)) := \|(a_n - b_n)\|_{\ell^1} = \sum |a_n - b_n|,$$

Introduction to Analysis Rui Wang Draft

and

$$\|(a_n) * (b_n)\|_{\ell^1} \leq \|(a_n)\|_{\ell^1} \|(b_n)\|_{\ell^1}.$$

3.7. Rearrangement. For a given series $\sum_{n=1}^{\infty} x_n$, a **rearrangement** is a bijective map

$$\phi : \mathbb{Z}^+ \to \mathbb{Z}^+,$$

and the corresponding series

$$\sum_{n=1}^{\infty} x_{\phi(n)}.$$

EXAMPLE 3.32. Consider the alternating convergent series

$$\Sigma_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots.$$

We know it is convergent but not absolutely convergent.

Now take any $s \in \mathbb{R}$, we build a rearrangement inductively:

- Add unused positive terms $\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \cdots$ until the partial sum first exceeds s;
- then add unused negative terms $-1, -\frac{1}{3}, -\frac{1}{5}, \cdots$ until the partial sum first drops below s;
- repeat these two steps forever.

Because the sums of the positive and negative subseries are $+\infty$ and $-\infty$ respectively, and because the term sizes $1/n$ tend to 0, the partial sums overshoot and undershoot s by ever smaller amounts. Hence the rearranged series converges to s.

Slightly generalize the above example, one can prove the following theorem by Riemann, whose proof is omit here.

THEOREM 3.33. *Suppose the series $\sum x_n$ is conditionally convergent. Then for any $a \leq b$ with a, b possibly being $\pm\infty$, there exists a rearrangement $\sum_n x_{\phi(n)}$ so that the partial sequence has b as upper limit and a as lower limit. In particular, given any $a \in \overline{\mathbb{R}}$, there exists a rearrangement that converges to a.*

On the other hand, when the series converges absolutely, there is the following theorem.

THEOREM 3.34. *Assume $\sum_{n=1}^{\infty} x_n$ is absolutely convergent. Then every rearrangement $\sum_{n=1}^{\infty} x_{\phi(n)}$ (with $\phi : \mathbb{Z}^+ \to \mathbb{Z}^+$ a bijection) is absolutely convergent and converges to the same sum.*

PROOF. Since $\sum |x_n|$ converges, for every $\epsilon > 0$ there exists $N \in \mathbb{Z}^+$ such that for all $m \geq n \geq N + 1$,

$$\sum_{k=n}^{m} |x_k| < \epsilon.$$

Fix such N, consider

$$N' := \max\{\phi^{-1}(j) | j = 1, \cdots, N\}.$$

Then for all $n > N'$, there are $\phi(n) \geq N + 1$.

For any $p \geq 0$,

$$\sum_{k=n}^{n+p} |x_{\phi(k)}| \leq \sum_{j=N+1}^{M} |x_j| < \epsilon,$$

where $M = \max\{\phi(n), \phi(n+1), \cdots, \phi(n+p)\}$.

By Cauchy's criterion, $\sum |x_{\phi(n)}|$ converges, and hence the rearranged series is absolutely convergent.

Introduction to Analysis Rui Wang Draft

To see the sums agree, let $S := \sum_{n=1}^{\infty} x_n$. Given $\epsilon > 0$, choose N so that the tail satisfies $T_N := \sum_{k>N} |x_k| < \epsilon$.

Pick m so large that $\{1, 2, \dots, N\} \subseteq \{\phi(1), \dots, \phi(m)\}$. Then

$$\sum_{k=1}^{m} x_{\phi(k)} = \sum_{j=1}^{N} x_j \; + \; R_m, \quad \text{with } |R_m| \le T_N.$$

Hence

$$|\sum_{k=1}^{m} x_{r(k)} - S| \le |R_m| + |\sum_{j>N} x_j| \le T_N + T_N < 2\epsilon.$$

Thus the partial sums of the rearranged series converge to S, as claimed. □

oduction to Analysis
Rui Wang
Draft

CHAPTER 4

Continuity

1. Limits of functions

1.1. Definition. Suppose (X, d_X) and (Y, d_y) are two metric spaces, and $\emptyset \neq U \subseteq X$ is the domain of a function

$$f : U \to Y.$$

In particular, when $Y = \mathbb{R}$, such a function is called a **real-valued** function; and when $Y = \mathbb{C}$, such a function is called a **complex-valued** function.

Continue with the setup above, we introduce the following definition.

DEFINITION 1.1. Suppose $x_0 \in U'$, i.e., a limit point of U, and y_0 is a point in Y. We say the **limit of the function** $f(x)$ **at** x_0 is y_0, denoted as

$$\lim_{x \to x_0} f(x) = y_0 \quad \text{or} \quad f(x) \to y_0 \text{ as } x \to x_0,$$

if the following $\epsilon - \delta$ statement holds:

For any $\epsilon > 0$, there exists some $\delta > 0$, so that any $x \in U$ with $0 < d_X(x, x_0) < \delta$, there is

$$d_Y(f(x), y_0) < \epsilon.$$

If there is no $y_0 \in Y$ so that $\lim_{x \to x_0} f(x) = y_0$, then we say the **limit of** $f(x)$ **at** x_0 **doesn't exist**.

The following remarks should be helpful in understanding the definition of limit of a function.

REMARK 1.2.

(1) It only makes sense to discuss the limit at a *limit point*: this guarantees that for *every* $\delta > 0$ there are points $x \in U$ with $0 < d_X(x, x_0) < \delta$.
(2) The definition does not use the value of f at x_0; in particular, $f(x_0)$ does not affect the limit, and f may even be undefined at x_0.
(3) The requirement $0 < d_X(x, x_0)$ (i.e., excluding $x = x_0$) is essential; otherwise one would merely force $f(x_0) = y_0$, which says nothing about nearby points.
(4) The limit is a *local property*: $\lim_{x \to x_0} f(x)$ depends only on the values of f on some open neighborhood of x_0.
(5) The negation of the limit statement can be written as follows: there exists $\epsilon_0 > 0$ such that for every $\delta > 0$ there exists $x \in U$ with

$$0 < d_X(x, x_0) < \delta \quad \text{and} \quad d_Y(f(x), y_0) \geq \epsilon_0.$$

Next, we introduce an equivalent way to define the limit of a function using sequences. This sequential formulation is often more convenient to use, especially since we have already studied limits of sequences in detail.

PROPOSITION 1.3. *The* $\epsilon - \delta$ *definition of*

$$\lim_{x \to x_0} f(x) = y_0$$

Introduction to Analysis Rui Wang Draft

is equivalent to the following sequence statement:
For any sequence (x_n) in $U \setminus \{x_0\}$ that converges to x_0 in (X, d_X), the sequence

$$f(x_n) \to y_0 \quad \text{as } n \to \infty, \text{ in } (Y, d_Y).$$

PROOF. • Suppose the $\epsilon - \delta$-statement holds. Let's prove the sequence statement.
For any $\epsilon > 0$, there is $\delta > 0$, so that any $x \in U$ with $0 < d_X(x, x_0) < \delta$, there is

$$d_Y(f(x), y_0) < \epsilon.$$

Now for a sequence (x_n) in $U \setminus \{x_0\}$ that converges to x_0, there is $N \in \mathbb{Z}^+$ so that

$$0 < d_X(x_n, x_0) < \delta, \quad \text{for all } n > N.$$

Then if follows

$$d_Y(f(x_n), y_0) < \epsilon, \quad \text{for all } n > N,$$

and so $f(x_n) \to y_0$.

- Suppose the sequence statement holds, but the $\epsilon - \delta$-statement doesn't hold, we will derive contradiction from it and then this shows the the sequence statement implies the $\epsilon-\delta$-statement.
If there is $\epsilon_0 > 0$ so that every $\delta > 0$, there exists $x \in U$ with

$$0 < d_X(x, x_0) < \delta \quad \text{and} \quad d_Y(f(x), y_0) \geq \epsilon_0.$$

We can take a sequence of δ as $\delta = \frac{1}{n}$, $n = 1, 2, \cdots$, and then obtain a sequence (x_n) in in $U \setminus \{x_0\}$ with

$$d_X(x_n, x_0) < \frac{1}{n} \quad \text{and} \quad d_Y(f(x_n), y_0) \geq \epsilon_0.$$

On the hand, by the squeeze theorem, $x_n \to x_0$, and there should follow

$$f(x_n) \to y_0,$$

which contradicts

$$d_Y\big(f(x_n), y_0\big) \geq \epsilon_0.$$

We are done.

□

1.2. Basic properties. So, use the properties for limits of sequences we can obtain some properties for limits of functions easily.

PROPOSITION 1.4. *The limit $\lim_{x \to x_0} f(x)$, if exists, then is unique.*

PROOF. We present two proofs.

(1) (Use $\epsilon - \delta$.) Suppose both $y_1, y_2 \in Y$ are limits $\lim_{x \to x_0} f(x)$. Then any $\epsilon > 0$, there exists some $\delta_i > 0$ so that any $x \in U$ with $0 < d_X(x, x_0) < \delta_i$, there are $d_Y(f(x), y_i) < \frac{\epsilon}{2}$, $i = 1, 2$.
Take $x \in U$ with $0 < d_X(x, x_0) < \delta := \min\{\delta_1, \delta_2\}$. (Notice δ is positive since both δ_1, δ_2 are positive.) There follows

$$d_Y(y_1, y_2) \leq d_Y(y_1, f(x)) + d_Y(f(x), y_2) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Notice this holds for all $\epsilon > 0$. Let ϵ shrink to zero, there follows $d_Y(y_1, y_2) = 0$, and so $y_1 = y_2$.

Introduction to Analysis Rui Wang Draft

(2) (Use sequence statement.) Suppose both $y_1, y_2 \in Y$ are limits $\lim_{x\to x_0} f(x)$. Then for any sequence (x_n) in $U \setminus \{x_0\}$ that converges to x_0, the sequence $(f(x_n))$ in Y converges to both y_1 and y_2.

By the uniqueness of limits of sequences, $y_1 = y_2$.

□

When the target space $Y = \mathbb{R}$ or $\mathbb{C}$, there are the following algebraic properties just as limits of sequences.

PROPOSITION 1.5. *Suppose $Y = \mathbb{R}$ (or $\mathbb{C}$), $\lim_{x\to x_0} f(x) = A$, and $\lim_{x\to x_0} g(x) = B$. There are*

(1) $\lim_{x\to x_0}(f \pm g)(x) = A \pm B$;
(2) $\lim_{x\to x_0}(f \cdot g)(x) = A \cdot B$;
(3) *If $B \neq 0$,* $\lim_{x\to x_0} \frac{f}{g}(x) = \frac{A}{B}$.
(4) *For any $c \in \mathbb{R}$ (or $\mathbb{C}$),* $\lim_{x\to x_0}(cf)(x) = cA$.

PROOF. We prove (3) here and others are left to the reader as exercises. Take an arbitrary sequence (x_n) in $U \setminus \{x_0\}$ that converges to x_0, there are

$$f(x_n) \to A, \quad \text{and} \quad g(x_n) \to B.$$

Then by the product property for sequences,

$$(f \cdot g)(x_n) = f(x_n) \cdot g(x_n) \to A \cdot B.$$

This shows that $\lim_{x\to x_0}(f \cdot g)(x) = A \cdot B$. □

Now, let's see some examples.

EXAMPLE 1.6. (1) (The constant function.) Take $y_0 \in Y$ and consider the function

$$f : U \to Y, \quad f(x) = y_0.$$

Then for any sequence (x_n) in $U \setminus \{x_0\}$ that converges to x_0, there is

$$f(x_n) = y_0.$$

This constant sequence always converges to y_0, we get

$$\lim_{x\to x_0} f(x) = y_0.$$

(2) (The identity map/function.) For any set X, the identity map (also called the identity function)

$$f(x) = x$$

is a *canonical* map (means, it doesn't need any choice to define) associated to it. With respect to the *same* metric d_X, it is easy to see the limit

$$\lim_{x\to x_0} f(x) = x_0,$$

since $x_n \to x_0$ and $f(x_n) \to x_0$ are exactly the same.

We remark that one needs to be careful about the metric here: In case of two *different* metrics are considered, it is possible that

$$\lim_{x\to x_0} f(x) \neq x_0.$$

Introduction to Analysis Rui Wang Draft

We leave the reader to check that when the domain metric is the standard Euclidean metric and the codomain metric is the discrete metric, the identity map has no limit at any point.

(3) (Polynomial functions.) Consider $X = Y = \mathbb{R}$ or $\mathbb{C}$, with the standard Euclidean metric. A polynomial function can be written as

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n, \quad a_0, a_1, \cdots, a_n \in \mathbb{R}(\text{ or } \mathbb{C}).$$

Apply Proposition 1.5 and the limit for the identity map, there is

$$\lim_{x \to x_0} f(x) = a_0 + a_1 x_0 + a_2 x_0^2 + \cdots a_n x_0^n = f(x_0),$$

for all $x_0 \in \mathbb{R}$ (or $\mathbb{C}$). With the definition of continuous function introduced later, this implies, polynomial functions are continuous function on $\mathbb{R}$ (or $\mathbb{C}$).

The set of all polynomials with coefficients in $\mathbb{R}$ (in $\mathbb{C}$) is written as $\mathbb{R}[x]$ (resp. $\mathbb{C}[x]$). In particular, constant functions and the identity function are polynomial functions.

(4) (Rational functions.) Let $f(x)$ and $g(x)$ be are two polynomial functions and $g(x)$ is not the zero function. Then the fraction $\frac{f(x)}{g(x)}$ is called a rational function.

There are at most finite zeros of $g(x)$, which form a closed set, let's denote it by Z, in $\mathbb{R}$. It follows $\mathbb{R} \setminus Z$ is a nonempty open set, and every point in it is a limit point due to the Euclidean metric.

Take $x_0 \in \mathbb{R} \setminus Z$, i.e., $g(x_0) \neq 0$, there is

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f(x_0)}{g(x_0)}.$$

due to the quotient property of limits of functions from Proposition 1.5. (The same statements hold over $\mathbb{C}$.)

With the definition of continuous function introduced later, this implies, rational functions are continuous function over it natural domain $\mathbb{R} \setminus Z$ (resp. $\mathbb{C} \setminus Z$).

The set of all rational functions with coefficients in $\mathbb{R}$ (in $\mathbb{C}$) is written as $\mathbb{R}(x)$ (resp. $\mathbb{C}(x)$). All polynomial functions are rational functions, i.e., $\mathbb{R}[x] \subseteq \mathbb{R}(x)$ (resp. $\mathbb{C}[x] \subseteq \mathbb{C}(x)$).

2. Continuous functions

2.1. Definition and basic examples. Again, we start with the general setup. Let $(X, d_X), (Y, d_Y)$ be two metric spaces and U is a nonempty subset of X as the domain of a function $f : U \to Y$.

DEFINITION 2.1. For a point $x_0 \in U$, we say the function $f(x)$ is **continuous at** x_0, if either

(1) x_0 is an isolated point of U; or
(2) x_0 is a limit point of U and $\lim_{x \to x_0} f(x) = f(x_0)$.

If f is not continuous at x_0, we say f is **discontinous at** x_0.

If f is continuous at every point $x_0 \in U$, we say f is a **continuous function** on U. We use $C^0(U, Y)$ or simply $C^0(U)$, to denote the set of all continuous functions on U.

The following sequential characterization of continuity follows immediately from the sequential description of limits.

PROPOSITION 2.2. *The function f is continuous at $x_0 \in U$ if and only if for every sequence (x_n) in U with $x_n \to x_0$, we have $f(x_n) \to f(x_0)$ as $n \to \infty$.*

From our limit calculations we already have the following examples.

Introduction to Analysis Rui Wang Draft

EXAMPLE 2.3. (1) Constant functions are continuous on their domains.

(2) If the domain and codomain carry the same metric, the identity map is continuous on X.

(3) For $Y = \mathbb{R}$ or $\mathbb{C}$, every polynomial function is continuous, i.e.

$$\mathbb{R}[x] \subseteq C^0(\mathbb{R}, \mathbb{R}) \quad \text{and} \quad \mathbb{C}[x] \subseteq C^0(\mathbb{C}, \mathbb{C}).$$

(4) Again for $Y = \mathbb{R}$ or $\mathbb{C}$, a rational function $\frac{f}{g}$ (with g not the zero polynomial) is continuous on its natural domain, namely

$$\frac{f}{g} \in C^0(\mathbb{R} \setminus Z, \mathbb{R}) \quad \text{and} \quad \frac{f}{g} \in C^0(\mathbb{C} \setminus Z, \mathbb{C}),$$

where $Z := \{x \mid g(x) = 0\}$ (empty or a finite set in either case).

The next statement follows immediately from Proposition 1.5.

PROPOSITION 2.4. *Let $U \subseteq X$ and $f, g : U \to Y$ with $Y = \mathbb{R}$ or $\mathbb{C}$. If f and g are continuous at $x_0 \in U$, then so are*

(1) *$f \pm g$;*
(2) *$f \cdot g$;*
(3) *$\frac{f}{g}$, provided $g(x_0) \neq 0$ (hence $g \neq 0$ in a neighborhood of x_0);*
(4) *cf for every scalar $c \in Y$.*

PROOF. We show (1). If x_0 is isolated in U, continuity is automatic. If x_0 is a limit point of U, then

$$\lim_{x \to x_0} (f + g)(x) = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x) = f(x_0) + g(x_0) = (f + g)(x_0),$$

so $f + g$ is continuous at x_0. The other items are identical applications of Proposition 1.5. □

COROLLARY 2.5. *If $f, g : U \to \mathbb{R}$ (or $\mathbb{C}$) are continuous on U, then $f \pm g$, cf (any $c \in \mathbb{R}$ or $\mathbb{C}$), and fg are continuous on U. Moreover, f/g is continuous on $U \setminus \{x \in U \mid g(x) = 0\}$ (assuming $g \not\equiv 0$).*

REMARK 2.6. As a consequence, $C^0(U, \mathbb{R})$ (resp. $C^0(U, \mathbb{C})$) is a commutative unital $\mathbb{R}$- (resp. $\mathbb{C}$-) algebra: it is a ring under pointwise $+$ and $\cdot$, a (infinite-dimensional) vector space over $\mathbb{R}$ (resp. $\mathbb{C}$), and the operations are compatible:

$$c(f + g) = cf + cg, \qquad c(fg) = (cf)g = f(cg).$$

The polynomial functions form a unital subalgebra: $\mathbb{R}[x] \subseteq C^0(\mathbb{R}, \mathbb{R})$ and $\mathbb{C}[x] \subseteq C^0(\mathbb{C}, \mathbb{C})$.

2.2. Examples from power series. Recall that a power series $\sum_{n=0}^{\infty} a_n x^n$ is a generalization of polynomials, that may contain infinitely many nonzero terms. In its convergence disk $D_R(0)$ with

$$R = \frac{1}{\limsup_{n \to \infty} \sqrt[n]{|a_n|}},$$

the series converges absolutely everywhere, and this defines a function

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \,:\, D_R(0) \to \mathbb{R}.$$

We now ask a question that is this function continuous at $x_0 \in D_R(0)$?

Introduction to Analysis Rui Wang Draft

EXAMPLE 2.7. Is the exponential function a continuous function on $\mathbb{R}$?

Take $x_0 \in \mathbb{R}$, by definition,

$$e^{x_0} = \sum_{n=0}^{\infty} \frac{1}{n!} {x_0}^n.$$

The question becomes to ask if

$$\lim_{x \to x_0} \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \sum_{n=0}^{\infty} \frac{1}{n!} {x_0}^n \quad ?$$

In general, to ask if the power series $\sum_{n=0}^{\infty} a_n x^n$ is continuous at $x_0 \in D_R(0)$ is the same to ask if there is

$$\lim_{x \to x_0} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n {x_0}^n. \tag{2.1}$$

Denote by

$$S_m(x) := \sum_{n=0}^{m} a_n x^n.$$

For each $m \geq 0$, this is a polynomial function, so is continuous everywhere in $\mathbb{R}$. In anther word, there is

$$\lim_{x \to x_0} S_m(x) = S_m(x_0)$$

for every $m \geq 0$.

Take $m \to \infty$, we obtain

$$\lim_{m \to \infty} (\lim_{x \to x_0} S_m(x)) = \lim_{m \to \infty} S_m(x_0) = \sum_{n=0}^{\infty} a_n {x_0}^n, \qquad \text{for every } x_0 \in B_R(0),$$

which is the RHS of (2.1).

The LHS of (2.1) can be written to

$$\lim_{x \to x_0} \sum_{n=0}^{\infty} a_n x^n = \lim_{x \to x_0} (\lim_{m \to \infty} S_m(x)).$$

Hence, within the disk of convergence, the question of continuity at x_0 reduces to whether we may interchange the two limiting processes. Concretely, that is to ask, if there is

$$\lim_{m \to \infty} (\lim_{x \to x_0} S_m(x)) = \lim_{x \to x_0} (\lim_{m \to \infty} S_m(x)).$$

A central – and subtle – theme in analysis is deciding when two limiting processes may be interchanged – this is not always true: for example

$$\lim_{m \to \infty} \lim_{n \to \infty} \frac{m}{m+n} \neq \lim_{n \to \infty} \lim_{m \to \infty} \frac{m}{m+n}.$$

This question motivates the notions of *uniform* hypotheses: uniform convergence, uniform continuity, and the uniform conditions that justify termwise differentiation and integration for sequence or series of functions. We will come to these concepts later on.

For now, let us prove that these two limiting processes can indeed be interchanged for the exponential series. The same argument generalizes to any power series at points within its disk of convergence.

EXAMPLE 2.8 (Continue). For the exponential function, for every $m \in \mathbb{Z}^+$, there is

$$\begin{aligned} |e^x - e^{x_0}| &= |(e^x - S_m(x)) + (S_m(x) - S_m(x_0)) + (S_m(x_0) - e^{x_0})| \\ &\leq |e^x - S_m(x)| + |S_m(x) - S_m(x_0)| + |S_m(x_0) - e^{x_0}|. \end{aligned}$$

Introduction to Analysis Rui Wang Draft

Notice that, for a fixed $r > 0$ and any $|x| < r$, there is

$$|e^x - S_m(x)| = |\sum_{n=m+1}^{\infty} \frac{1}{n!}x^n| \le |\sum_{n=m+1}^{\infty} \frac{1}{n!}r^n| \to 0$$

as $m \to \infty$. This shows for any $\epsilon > 0$, there is $M \in \mathbb{Z}^+$ so that every $m > M$, there is

$$|e^x - S_m(x)| < \epsilon, \quad \text{for } \textbf{all } |x| < r.$$

The key point is that the choice of M is independent of x, whenever

$$\sum_{n=0}^{\infty} \frac{1}{n!}r^n \text{ is convergent and } |x| < r.$$

(Here since the convergence radius of $\sum_{n=0}^{\infty} \frac{1}{n!}x^n$ is infinity, the series is absolutely convergent for any $r > 0$.)

Then we only need to fix some $r > |x_0|$ and then fix a $m > M$. By the continuity of the polynomial $S_m(x)$, there is $\delta > 0$ so that any $|x - x_0| < \delta$ there is

$$|S_m(x) - S_m(x_0)| < \epsilon.$$

There follows then

$$|e^x - S_m(x)| + |S_m(x) - S_m(x_0)| + |S_m(x_0) - e^{x_0}| < \epsilon + \epsilon + \epsilon = 3\epsilon.$$

All together, this proves

$$\lim_{x \to x_0} e^x = e^{x_0},$$

i.e., the exponential function is continuous at every $x_0 \in \mathbb{R}$.

Similarly, the trigonometric functions

$$\sin x := \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} - \cdots,$$

and

$$\cos x := \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots,$$

both have ∞ as convergence radius, so absolutely convergent and continuous everywhere in $\mathbb{R}$.

The above examples, the exponential function, sine and cosine functions, can actually be define on $\mathbb{C}$. The same argument goes to show they are actually continuous functions on $\mathbb{C}$. Also, with complex notations, they are related by the **Euler's identity**

$$e^{ix} = \cos x + i \sin x.$$

Then

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

for every complex number x.

From these examples, it is not hard to derive the following general theorem.

THEOREM 2.9. *A power series $\sum a_n x^n$ is continuous on its convergence disk $B_R(0)$.*

The details in the proof is left to the reader as a homework problem.

Introduction to Analysis Rui Wang Draft

2.3. More on continuity. We return to continuous maps between two general metric spaces. We summarize the continuity at a point in the following proposition.

PROPOSITION 2.10. *Let (X, d_X) and (Y, d_Y) be metric spaces, $U \subseteq X$, and $f : U \to Y$. Fix $x_0 \in U$. The following are equivalent:*

(1) *f is continuous at x_0.*
(2) *For every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in U$ with $d_X(x, x_0) < \delta$ one has $d_Y\big(f(x), f(x_0)\big) < \epsilon$.*
(3) *For every $\epsilon > 0$ there exists $\delta > 0$ such that*
$$f(B_\delta(x_0) \cap U) \subseteq B_\epsilon(f(x_0)).$$
(4) *For every $\epsilon > 0$, the set $f^{-1}(B_\epsilon(f(x_0)))$ is a neighbourhood of x_0 in U, i.e. it contains $B_\delta(x_0) \cap U$ for some $\delta > 0$.*
(5) *For every sequence (x_n) in U that converges to x_0, the sequence $f(x_n)$ converges to $f(x_0)$.*

PROPOSITION 2.11. *Let $f : (X, d_X) \to (Y, d_Y)$.*

(1) *f is continuous on X if and only if the preimage of every open set $V \subseteq Y$ is open in X.*
(2) *f is continuous on X if and only if the preimage of every closed set $F \subseteq Y$ is closed in X.*

PROOF. (1) Suppose f is continuous. Let $V \subseteq Y$ be open. For any $x \in f^{-1}(V)$, choose $\epsilon > 0$ with $B_\epsilon(f(x)) \subseteq V$. By continuity at x there exists $\delta > 0$ such that $f(B_\delta(x)) \subseteq B_\epsilon(f(x)) \subseteq V$, hence $B_\delta(x) \subseteq f^{-1}(V)$, so $f^{-1}(V)$ is open.

Conversely, if $f^{-1}(V)$ is open in X for every open $V \subseteq Y$, then for any $x \in X$ and $\epsilon > 0$, the set $f^{-1}(B_\epsilon(f(x)))$ is an open neighbourhood of x, i.e. it contains some $B_\delta(x)$. Thus $f(B_\delta(x)) \subseteq B_\epsilon(f(x))$, showing f is continuous at x. (2) is equivalent via complements. □

REMARK 2.12. In general topological spaces (not necessarily metric), the open-preimage condition in Proposition 2.11(1) is taken as the *definition* of continuity.

PROPOSITION 2.13. *Suppose (X, d_X), (Y, d_Y) and (Z, d_Z) are three metric spaces. If $f : X \to Y$ and $g : Y \to Z$ are both continuous functions, then the composition*
$$g \circ f : X \to Z$$
is continuous.

PROOF. Take any open set $W \subseteq Z$. Since g is continuous, by Proposition 2.11, $g^{-1}(W)$ is an open set in Y. Further since f is continuous, by Proposition 2.11, $f^{-1}(g^{-1}(W))$ is an open set in X.

Notice that
$$f^{-1}(g^{-1}(W)) = (g \circ f)^{-1}(W),$$
by Proposition 2.11 again, $g \circ f$ is continuous.

□

REMARK 2.14 (*). Proposition 2.13 actually shows that continuous maps are the morphisms in the category of metric spaces (and, more generally, in the category of topological spaces).

Now we introduce the definition of uniform continuity. We will prove later, a continuous function on a compact set must be uniformly continuous.

Introduction to Analysis Rui Wang Draft

DEFINITION 2.15. Let (X, d_X) and (Y, d_Y) be metric spaces and $U \subseteq X$. A function $f : U \to Y$ is **uniformly continuous** if for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x_1, x_2 \in U$,

$$d_Y\big(f(x_1), f(x_2)\big) < \epsilon \quad \text{whenever} \quad d_X(x_1, x_2) < \delta.$$

In particular, fixing $x_2 = x_0 \in U$ shows that uniform continuity implies ordinary (pointwise) continuity:

PROPOSITION 2.16. *If $f : U \to Y$ is uniformly continuous, then f is continuous at every $x_0 \in U$.*

The converse fails in general.

EXAMPLE 2.17. Let $f(x) = \frac{1}{x}$ on $(0, \infty)$. It is continuous, but not uniformly continuous. Notice

$$|\tfrac{1}{x_1} - \tfrac{1}{x_2}| = \frac{|x_1 - x_2|}{|x_1 x_2|}.$$

Thus even if $|x_1 - x_2| < \delta$, the difference can be made arbitrarily large by taking $x_1 x_2$ very small (i.e.x_1, x_2 close to 0).

3. Continuity and compactness

As before, we assume (X, d_X) and (Y, d_Y) are metric spaces.

THEOREM 3.1. *Let $f : X \to Y$ be continuous. Then the image of every compact set $K \subseteq X$ is compact in Y, i.e. $f(K)$ is compact.*

PROOF. We argue from the definition of compactness. Let $\{V_\alpha \mid \alpha \in \Lambda\}$ be an open cover of $f(K)$. By continuity of f and Proposition 2.11, the family $\{f^{-1}(V_\alpha) \mid \alpha \in \Lambda\}$ is an open cover of K. Since K is compact, there exist indices $\alpha_1, \dots, \alpha_n$ such that

$$K \subseteq \bigcup_{i=1}^{n} f^{-1}(V_{\alpha_i}).$$

Applying f, we obtain

$$f(K) \subseteq \bigcup_{i=1}^{n} V_{\alpha_i},$$

so $\{V_{\alpha_1}, \dots, V_{\alpha_n}\}$ is a finite subcover of $f(K)$. Hence $f(K)$ is compact. □

We record an alternative proof in terms of sequential compactness. For metric spaces compactness is equivalent to sequential compactness, so this yields the same result. In general topological spaces, however, compactness and sequential compactness are independent notions.

ALTERNATIVE PROOF. Assume K is sequentially compact. We show $f(K)$ is sequentially compact. Let $\{f(x_n)\}$ be any sequence in $f(K)$ with $x_n \in K$. Since K is sequentially compact, there is a subsequence $x_{n_k} \to x_0 \in K$. By continuity of f we have

$$f(x_{n_k}) \to f(x_0) \in f(K),$$

so $\{f(x_n)\}$ has a convergent subsequence in $f(K)$. Thus $f(K)$ is sequentially compact. □

The converse of the statement is not true in general.

DEFINITION 3.2. A function $f : X \to Y$ is called **proper** if the preimage of any compact subset of Y is compact in X.

duction to Analysis Rui Wang Draft

EXAMPLE 3.3. Not every continuous function is proper. Consider $f(x) = \frac{1}{x}$ from $(0, +\infty)$ to $\mathbb{R}$. The preimage of $[0, 1]$ is $[1, +\infty)$ which is not compact.

An immediate corollary from the above result and Proposition 4.2 is the following

COROLLARY 3.4. *Assume $f : X \to Y$ is a continuous map. Then for any compact subset $K \subseteq X$, the image set $f(K)$ is bounded and closed in Y.*

In particular, when $Y = \mathbb{R}$, we have the following important result.

THEOREM 3.5. *A real-valued continuous function defined on a compact subset of a metric space can obtain its supremum and infimum, i.e., has maximum and minimum.*

PROOF. Assume $f : K \to \mathbb{R}$ is a continuous function and Y is a compact subset of X. From Corollary 3.4, we know $f(K)$ is a bounded and closed subset of $\mathbb{R}$.

Then by the completeness of $\mathbb{R}$, $\sup f(K), \inf f(K) \in \mathbb{R}$. Since $\sup f(K), \inf f(K)$ are limit points of $f(K)$, further by the closeness of $f(K)$, they live in $f(K)$. This is saying that there are points $x_1, x_2 \in K$ so that

$$f(x_1) = \sup f(K), \quad f(x_2) = \inf f(K).$$

□

EXAMPLE 3.6. The assumption of compactness is important. For example $f(x) = \frac{1}{x}$ has neither max nor min over $(0, +\infty) \subseteq \mathbb{R}$, but it has both over any closed interval $[a, b] \subseteq (0, +\infty)$, for $a \leq b$.

Next, we consider the continuity of the inverse function of a continuous function. We know that if f is an injective function from X to Y, then we can define the inverse function

$$f^{-1} : f(X) \to X,$$

which maps $f(x)$ back to x. Here comes the question that if f is continuous, then is f^{-1} continuous on $f(X)$? The answer is no in general.

EXAMPLE 3.7 (Same set, different metrics). Let $X = (\mathbb{R}, d_{\text{disc}})$ with the discrete metric $d_{\text{disc}}(x, y) = 0$ iff $x = y$ and 1 otherwise, and let $Y = (\mathbb{R}, |\cdot|)$ with the Euclidean metric. Consider the identity map $f : X \to Y$, $f(x) = x$.

Claim. f is continuous, but $f^{-1} : Y \to X$ is not.

Proof. For continuity of f, use the ε–δ definition: if $d_{\text{disc}}(x, x_0) < \frac{1}{2}$ then $x = x_0$, hence $|f(x) - f(x_0)| = 0 < \varepsilon$ for any $\varepsilon > 0$.

To see f^{-1} is not continuous, use the open-set characterization. The singleton $\{0\} \subset X$ is open (discrete topology). But

$$(f^{-1})^{-1}(\{0\}) = \{0\} \subset Y$$

is not open in the Euclidean topology. Hence f^{-1} is not continuous. □

EXAMPLE 3.8 (Wrapping an interval onto the circle). Let $X = [0, 2\pi)$ with the Euclidean metric and let $Y = S^1 = \{(\cos t, \sin t) : t \in \mathbb{R}\} \subset \mathbb{R}^2$ with the subspace metric. Define

$$f : X \to Y, \qquad f(t) = (\cos t, \sin t).$$

Then f is a continuous bijection, but $f^{-1} : Y \to X$ is not continuous.

Introduction to Analysis Rui Wang Draft

Proof. Continuity of f is clear (polynomial combinations of sin, cos). Bijectivity holds because each point of S^1 has a unique preimage in $[0, 2\pi)$. Set $z_n = f(2\pi - \frac{1}{n})$. Then $z_n \to (1, 0)$ in Y, while

$$f^{-1}(z_n) = 2\pi - \frac{1}{n} \longrightarrow 2\pi \neq 0 = f^{-1}((1, 0))$$

in X. Hence f^{-1} is not continuous at $(1, 0)$. □

In fact, we can prove the following result.

THEOREM 3.9. *Assume f is an bijective function from X to Y and f is continuous. Then if X is compact, its inverse is also continuous.*

PROOF. We give two proofs. The first one uses open sets interpretationand the second proof uses sequence interpretation.

(1) This is enough to show f is an open map, i.e., f maps any open subset U of X to an open set of Y.

To show $f(U)$ is open, it is enough to show $f(U)^c$ is closed. Notice that

$$f(U)^c = f(U^c).$$

Since U^c is closed subset of X, which is compact, U^c is compact, and hence $f(U^c)$ is also compact by Theorem 3.1. Then $f(U)^c$ is closed and we are done.

(2) Taking an arbitrary sequence $\{f(x_n) | x_n \in X\}$ in Y that converges to some $f(x_0)$, we prove that the sequence $\{f^{-1}(f(x_n)) = x_n\}$ converges to x_0.

Assume this is not the case, then there exists some $\epsilon_0 > 0$ and a subsequence $\{x_{n_k}\}$ so that

$$d_X(x_{n_k}, x_0) \geq \epsilon_0.$$

Assuming X is sequentially compact, we can further find a subsequence $\{x_{n_{k_\ell}}\}$ so that it converges to some $x_0' \in X$, and

$$d_X(x_0', x_0) \geq \epsilon_0.$$

At the same time, by the continuity of f, $\{f(x_{n_{k_\ell}})\}$ converges to $f(x_0')$. By the uniqueness of limit, there must be

$$f(x_0) = f(x_0'),$$

and thus $x_0 = x_0'$, which contradicts with $d_X(x_0', x_0) \geq \epsilon_0$.

□

REMARK 3.10. Two metric spaces X and Y are called **homeomorphic**, if there exists some bijective map $f : X \to Y$ with both f and f^{-1} continuous. The above result shows that assuming X is compact, for a homeomorphism, it is enough to check f is continuous.

THEOREM 3.11. *For a function defined on a compact domain, it is continuous if and only if it is uniformly continuous.*

PROOF. We only need to prove that under the assumption of compactness, continuity implies uniform continuity. Again we provide two proofs, one is using compactness, the other uses sequential compactness.

Introduction to Analysis Rui Wang Draft

(1) Since f is continuous everywhere on X, for any given $\epsilon > 0$, for each $x \in X$, there exists some $\delta(x) > 0$ (may DEPEND on x) so that

$$f(B_{\delta(x)}(x)) \subseteq B_\epsilon(f(x)).$$

Now notice that

$$\{B_{\frac{1}{2}\delta(x)}(x) | x \in X\}$$

form an open cover of X (here to shrink $\delta(x)$ to $\frac{1}{2}\delta(x)$ is technically important.) and X is compact, we can pick finite points $x_1, \cdots, x_n$ from X so that the balls

$$B_{\frac{1}{2}\delta(x_1)}(x_1), B_{\frac{1}{2}\delta(x_2)}(x_2), \cdots, B_{\frac{1}{2}\delta(x_n)}(x_n)$$

cover X. Define

$$\delta := \min\{\frac{1}{2}\delta(x_1), \cdots, \frac{1}{2}\delta(x_n)\}.$$

Now for any two points $p_1, p_2 \in X$ with distance

$$d_X(p_1, p_2) < \delta,$$

we estimate $d_Y(f(p_1), f(p_2))$.

First, p_1 must live in a ball $B_{\frac{1}{2}\delta(x_i)}(x_i)$ for some $i = 1, \cdots, n$. We claim that $p_2 \in B_{\delta(x_i)}(x_i)$. This is because

$$\begin{aligned} & d_X(x_i, p_2) \\ \leq\ & d_X(x_i, p_1) + d_X(p_1, p_2) \\ \leq\ & \frac{1}{2}\delta(x_i) + \delta \\ \leq\ & \frac{1}{2}\delta(x_i) + \frac{1}{2}\delta(x_i) = \delta(x_i). \end{aligned}$$

Then

$$\begin{aligned} & d_Y(f(p_1), f(p_2)) \\ \leq\ & d_Y(f(p_1), f(x_i)) + d_Y(f(x_i), f(p_2)) \\ \leq\ & \epsilon + \epsilon = 2\epsilon. \end{aligned}$$

This proves f is uniformly continuous.

(2) We now give a proof using sequences.

Assume f is not uniformly continuous. Then there exists some $\epsilon > 0$ and a pair of sequences $\{x_n\}, \{x'_n\}$ with

$$d_X(x_n, x'_n) < \frac{1}{n}$$

but

$$d_Y(f(x_n), f(x'_n)) \geq \epsilon_0.$$

Because X is compact, after passing to subsequences, we can assume $x_n \to x_0$ and $x'_n \to x'_0$ for some $x_0, x'_0 \in X$. (I abuse notations here that I still use $\{x_n\}$ to denote a subsequence of it, but this way is very commonly used in literatures.) By the continuity of the distance function $d_X : X \times X \to \mathbb{R}$ (This is supposed to be proved in your homework.) implies

$$d_X(x_0, x'_0) \leq \lim_{n\to\infty} \frac{1}{n} = 0, \quad \text{thus, } x_0 = x'_0.$$

Introduction to Analysis Rui Wang Draft

On the other hand, apply the continuity of f and the distance function $d_Y : Y \times Y \to \mathbb{R}$, it follows

$$d_Y(f(x_0), f(x_0')) \geq \epsilon_0,$$

which then contradicts with $x_0 = x_0'$.

This proves f must be uniformly continuous.

□

4. Continuity and connectedness

4.1. Two kinds of connectedness.

DEFINITION 4.1. A metric space X is called **connected**, if the following holds:
$X = U \cup V, U \cap V = \emptyset$, both U, V are open $\implies$ either U or V is empty set.
A subset $S \subseteq X$ is called **connected** if it is connected as a metric space with the subspace (induced) metric from X.

REMARK 4.2. (*)

(1) (Open sets relative to a subset) Let (X, d) be a metric space and $S \subseteq X$. A set $O \subseteq S$ is *open in* S (with the subspace/relative topology) iff for every $x \in O$ there exists $r > 0$ such that

$$B_r^X(x) \cap S \subseteq O,$$

equivalently, $O = U \cap S$ for some open $U \subseteq X$. With the restricted metric $d|_{S\times S}$, the open balls in S are $B_r^S(x) = B_r^X(x) \cap S$. Similarly, $C \subseteq S$ is *closed in* S iff $C = F \cap S$ for some closed $F \subseteq X$, and

$$\overline{A}^S = \overline{A}^X \cap S, \qquad \text{int}_S(A) = \text{int}_X(A) \cap S.$$

Here $\text{int}_Y(A)$ denotes the set of interior points in A with respect to the metric space Y.

(2) (Connectedness is defined in any topological space) The notion of connectedness depends only on the open sets of a space, so it makes sense for *any* topological space (not just metric spaces).

PROPOSITION 4.3. *X is connected, if and only if no proper nonempty subset of X which is both open and closed.*

PROOF.
- ($\implies$.) Take any $S \subseteq X$, we can write $X = S \cup S^c$. Suppose S is open and closed, then by the definition of connectedness, we must have either S or S^c is empty, which implies S is either empty or X.
- ($\impliedby$.) When $X = U \cup V$ with both U, V open, then $U = V^c$ is both open and closed. If $U \neq \emptyset$, then $U = X$, and then $V = U^c = \emptyset$.

□

EXAMPLE 4.4. In $\mathbb{R}^2$, the union of two disjoint disks is not connected (clear from the definition). In contrast, if two disks $D_1, D_2 \subset \mathbb{R}^2$ intersect, then $D_1 \cup D_2$ is *path-connected* (defined next) and we will prove every path-connected metric space is connected.

We now introduce the notion of path-connectedness which is often easier to verify than connectedness.

Introduction to Analysis Rui Wang Draft

DEFINITION 4.5. A metric space X is **path-connected** if for every $x, y \in X$ there exists a continuous map $\gamma : [0,1] \to X$ with $\gamma(0) = x$ and $\gamma(1) = y$.

EXAMPLE 4.6. In $\mathbb{R}^n$ with the standard Euclidean metric:

(1) Every (open or closed) ball is path-connected.
(2) The union of two nonempty disjoint balls is neither connected nor path-connected.
(3) A subset $S \subset \mathbb{R}^n$ is **convex** if for any $x_0, x_1 \in S$ the straight line segment

$$[x_0, x_1] = \{(1-t)x_0 + tx_1|, t \in [0,1]\}$$

lies in S. Every convex set is path-connected.
(4) A subset $S \subset \mathbb{R}^n$ is **star-shaped** if there exists $x_0 \in S$ such that for every $x \in S$ the segment $[x_0, x] \subset S$. Every star-shaped set is path-connected.

LEMMA 4.7. *$[0,1] \subseteq \mathbb{R}$ is both connected and path-connected.*

PROOF. Path-connectedness of $[0,1]$ is immediate: for any $x, y \in [0,1]$ the map

$$t \mapsto (1-t)x + ty$$

is a continuous path in $[0,1]$ joining x to y.

We now prove connectedness. Suppose $[0,1] = A \cup B$ where A, B are disjoint, nonempty sets that are open in the subspace topology of $[0,1]$. Pick $a \in A$ and $b \in B$ with $a < b$. Since B is (relatively) open in $[0,1]$, there exists $\varepsilon > 0$ such that $(b-\varepsilon, b] \subset B$. Set

$$s := \sup\big(A \cap [a,b]\big).$$

Then $A \cap [a,b] \neq \emptyset$ (because $a \in A$), so s exists and $a \leq s \leq b$. As $b \in B$ and $A \cap B = \emptyset$, we must have $s < b$.

We claim $s \notin A$ and $s \notin B$, which is impossible since $A \cup B = [0,1]$.

- If $s \in A$, then by relative openness of A there exists $\delta > 0$ with

$$(s-\delta, s+\delta) \cap [0,1] \subset A.$$

Because $s < b$, we may choose $y \in (s, \min\{s+\delta, b\})$. Then $y \in A \cap [a,b]$ and $y > s$, contradicting $s = \sup(A \cap [a,b])$. Thus $s \notin A$.
- If $s \in B$, then by relative openness of B there exists $\delta > 0$ with

$$(s-\delta, s+\delta) \cap [0,1] \subset B.$$

By the definition of supremum, for this δ there exists $x \in A \cap [a,b]$ with $s - \delta < x < s$. Hence $x \in (s-\delta, s) \subset B$, contradicting $A \cap B = \emptyset$. Thus $s \notin B$.

Therefore $s \notin A \cup B = [0,1]$, a contradiction. It follows that $[0,1]$ is connected. □

Using this, we now prove

THEOREM 4.8. *Any path-connected metric space must be connected.*

PROOF. Assume X is a path-connected metric space, we prove it must be connected. For this, assume

$$X = U \cup V$$

with both U, V open and U, V are disjoint, then if neither is empty, we pick $x_0 \in U$ and $x_1 \in V$.

Introduction to Analysis Rui Wang Draft

Since X is path-connected, there exists a continuous function

$$f : [0,1] \to X, \quad f(0) = x_0, f(1) = x_1.$$

Define $A := f^{-1}(U)$ and $B := f^{-1}(V)$. Both A, B are open in $[0,1]$ since f is continuous, but neither is empty since

$$0 \in A, \quad 1 \in B.$$

Hence $[0,1]$ is not connected, which contradicts with the fact we just proved in Theorem 4.7. □

However, the verse vice is not true as we can see from the following topologists' sine curve.

EXAMPLE 4.9. [Topologist's sine curve] Let

$$T := \left\{(x, \sin \tfrac{1}{x}) : x \in (0,1]\right\}, \qquad S := \overline{T} = T \cup \left(\{0\} \times [-1,1]\right) \subset \mathbb{R}^2.$$

The set S is connected (with the metric inherited from $\mathbb{R}^2$). (Hint: first use the next theorems to show T is connected as the continuous image of the connected interval $(0,1]$, then use the closure of a connected set is connected to show S is connected.)

The set S is NOT path-connected: there is no continuous path in S joining a point on the vertical segment $\{0\} \times [-1,1]$ to a point of T.

PROPOSITION 4.10. *(*) Let X be a topological (in particular, metric) space and $S \subset X$ be connected. Then the closure $\overline{S}$ is connected.*

PROOF. Suppose, toward a contradiction, that $\overline{S} = A \cup B$ with A, B disjoint, nonempty, and open in the subspace topology of $\overline{S}$. Then there exist open sets $U, V \subset X$ with $U \cap V = \emptyset$ such that $A = U \cap \overline{S}$ and $B = V \cap \overline{S}$.

Intersecting with S gives a separation

$$S = (U \cap S) \cup (V \cap S),$$

where $U \cap S$ and $V \cap S$ are open in the subspace S and disjoint. Since S is connected, one of them is empty; without loss of generality $U \cap S = \emptyset$.

But then $X \setminus U$ is a closed set containing S, hence $\overline{S} \subset X \setminus U$, so $A = U \cap \overline{S} = \emptyset$, contradiction. Therefore $\overline{S}$ is connected. □

We remark that closure needs not preserve path–connectedness from the example above.

In fact, in $\mathbb{R}^n$ (with the standard Euclidean metric), every *open* connected set is path-connected. In anther word, for *open* sets in $\mathbb{R}^n$, connectedness and path-connectedness are exactly the same. (Notice Example exm:toposine, the topologist's sine curve is not open in $\mathbb{R}^2$.)

THEOREM 4.11. *(*) Every nonempty open connected set $U \subset \mathbb{R}^n$ is path-connected.*

PROOF. Fix $x_0 \in U$ and let

$$E := \{\, y \in U : \text{ there exists a continuous path } \gamma : [0,1] \to U \text{ with } \gamma(0) = x_0,\ \gamma(1) = y \,\}.$$

(Equivalently, one may require γ to be piecewise linear; the proof below shows such paths exist.)

We show E is both open and relatively closed in U.

Step 1: E is open in U. Let $y \in E$. Since U is open in $\mathbb{R}^n$, there exists $r > 0$ with $B(y,r) \subset U$. For any $z \in B(y,r)$, the straight line segment $[y,z] \subset B(y,r) \subset U$. Concatenating a path in U from x_0 to y with the segment $[y,z]$ gives a path from x_0 to z in U. Hence $B(y,r) \subset E$, so E is open.

Introduction to Analysis Rui Wang Draft

Step 2: $U \setminus E$ is open in U. Let $y \in U \setminus E$. Again choose $r > 0$ with $B(y, r) \subset U$. If there were $z \in B(y, r) \cap E$, then by Step 1 there is a path in U from x_0 to z, and the straight segment $[z, y] \subset B(y, r) \subset U$ would extend this to a path from x_0 to y, contradicting $y \notin E$. Therefore $B(y, r) \subset U \setminus E$, so $U \setminus E$ is open.

We have shown that E is both open and closed (relative to U). Since $x_0 \in E$, $E \neq \emptyset$. Because U is connected, the only nonempty clopen subset of U is U itself. Thus $E = U$. Consequently, every $y \in U$ can be joined to x_0 by a path in U; hence U is path-connected. □

REMARK 4.12. (*) The proof uses that open subsets of $\mathbb{R}^n$ are *locally path-connected* (every point has a small convex ball). In any locally path-connected space, connectedness implies path-connectedness of *open* connected sets.

The following two parallel theorems show that a continuous map preserves connectedness.

THEOREM 4.13. *Let (X, d_X) and (Y, d_Y) be metric spaces. If $f : X \to Y$ is continuous and X is connected, then $f(X) \subseteq Y$ (with the subspace metric/topology from Y) is connected.*

PROOF. Suppose, toward a contradiction, that $f(X)$ is not connected. Then there exist disjoint nonempty sets $U, V \subseteq f(X)$ that are open in the subspace $f(X)$ and satisfy $f(X) = U \cup V$. By definition of the subspace topology, there are open sets $O, O' \subseteq Y$ with $U = O \cap f(X)$ and $V = O' \cap f(X)$. Since f is continuous, $f^{-1}(O)$ and $f^{-1}(O')$ are open in X. Moreover,

$$X = f^{-1}(f(X)) = f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V),$$

and $f^{-1}(U) \cap f^{-1}(V) = \emptyset$. Thus X is the union of two disjoint nonempty open sets, contradicting connectedness of X. Hence $f(X)$ must be connected. □

THEOREM 4.14. *Let (X, d_X) and (Y, d_Y) be metric spaces. If $f : X \to Y$ is continuous and X is path-connected, then $f(X) \subseteq Y$ (with the subspace metric/topology) is path-connected.*

PROOF. Take any $y_0, y_1 \in f(X)$. Choose $x_0, x_1 \in X$ with $f(x_i) = y_i$. Since X is path-connected, there exists a continuous path $\gamma : [0, 1] \to X$ with $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Then $f \circ \gamma : [0, 1] \to Y$ is a continuous path entirely contained in $f(X)$ joining y_0 to y_1. Hence $f(X)$ is path-connected. □

REMARK 4.15. Again both theorems remain true for general topological spaces (metric hypotheses are not needed).

4.2. Connectedness in $\mathbb{R}$ and the intermediate value theorem. Though in general, the path-connectedness is stronger than connectedness, for subsets in $\mathbb{R}$, the connectedness and path-connectedness are exactly the same thing.

THEOREM 4.16. *In $\mathbb{R}$, a subset is connected if and only it is path-connected.*

PROOF. We only need to prove any connected subset of $\mathbb{R}$ is path-connected. Take $S \subseteq \mathbb{R}$ which is connected. For any two points $x_0, x_1 \in S$, we show that $[x_0, x_1] \subseteq S$.

If this is not the case, there exists some $x \in [x_0, x_1]$ but not in S. Then define

$$A := S \cap (-\infty, x), \quad B := S \cap (x, +\infty).$$

They are disjoint and both are open in S. Moreover, $S = A \cup B$, but neither A nor B is empty. This contradicts with the connectedness of S. □

Introduction to Analysis
Rui Wang
Draft

Now, let's use it to prove the important result for continuous functions.

THEOREM 4.17. *[Intermediate value theorem]Assume $[a,b] \subseteq \mathbb{R}$ is a closed interval, and $f : [a,b] \to \mathbb{R}$ is a continuous function. If $f(a) < f(b)$, then for each*

$$f(a) < y_0 < f(b),$$

there exists some $x_0 \in (a,b)$ so that $f(x_0) = y_0$.

PROOF. Since $[a,b]$ is connected, using Theorem 4.13, $f([a,b])$ is connected. Moreover, $f(a), f(b) \in f([a,b]) \subseteq \mathbb{R}$. From the proof of Theorem 4.16, the closed interval

$$[f(a), f(b)] \subseteq f([a,b]).$$

Now for any $f(a) < y_0 < f(b)$, there is

$$y_0 \in f([a,b]).$$

It follows there must be some $x_0 \in (a,b)$ with $f(x_0) = y_0$.

□

5. Monotonic functions and their discounuity

5.1. One-side limits, and types of discontinuous points for functions defined on $\mathbb{R}$. Let $U \subseteq \mathbb{R}$ and $f : U \to \mathbb{R}$ be a function. For $x_0 \in \mathbb{R}$, set

$$U^- := U \cap (-\infty, x_0), \qquad U^+ := U \cap (x_0, \infty).$$

We write $f|_{U^\pm}$ for the restriction of f to $U^\pm$.

DEFINITION 5.1. (1) We say the **left-hand limit** of f at x_0 exists and equals L, written $\lim_{x \to x_0^-} f(x) = L$, if x_0 is a limit point of U^- and

$$\lim_{x \to x_0} f|_{U^-}(x) = L.$$

(2) We say the **right-hand limit** of f at x_0 exists and equals R, written $\lim_{x \to x_0^+} f(x) = R$, if x_0 is a limit point of U^+ and

$$\lim_{x \to x_0} f|_{U^+}(x) = R.$$

PROPOSITION 5.2. *Assume x_0 is a limit point of U^-. The following are equivalent:*

(1) $\lim_{x \to x_0^-} f(x) = L$.
(2) *For every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in U$ with $0 < x_0 - x < \delta$, we have $|f(x) - L| < \epsilon$.*
(3) *For every sequence $(x_n) \subset U \setminus \{x_0\}$ with $x_n \to x_0$ and (eventually) $x_n < x_0$, i.e., $x_n \to x_0-$, we have $f(x_n) \to L$.*

Similarly, assuming x_0 is a limit point of U^+, the following are equivalent:

(1) $\lim_{x \to x_0^+} f(x) = R$.
(2) *For every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in U$ with $0 < x - x_0 < \delta$, we have $|f(x) - R| < \epsilon$.*
(3) *For every sequence $(x_n) \subset U \setminus \{x_0\}$ with $x_n \to x_0$ and (eventually) $x_n > x_0$, i.e., $x_n \to x_0+$, we have $f(x_n) \to R$.*

Introduction to Analysis Rui Wang Draft

COROLLARY 5.3. *Let $U \subseteq \mathbb{R}$, $f : U \to \mathbb{R}$, and let x_0 be a limit point of U. Then $\lim_{x\to x_0} f(x)$ exists if and only if one of the following holds:*

(1) *x_0 is a limit point of exactly one of U^- or U^+, and the corresponding one-sided limit exists (in which case $\lim_{x\to x_0} f(x)$ equals that one-sided limit);*
(2) *x_0 is a limit point of both U^- and U^+, and both one-sided limits exist and are equal.*

Now we are ready to discuss the types of discontinuities on $\mathbb{R}$.

DEFINITION 5.4. Let $U \subseteq \mathbb{R}$ and $f : U \to \mathbb{R}$. Assume that x_0 is a limit point of both U^- and U^+. We say that f has a **discontinuity at** x_0 if f is not continuous at x_0. We classify such discontinuities as follows (all limits are finite real limits):

(1) **First kind (simple discontinuity):** both one–sided limits exist:

$$\lim_{x\to x_0^-} f(x) \quad \text{and} \quad \lim_{x\to x_0^+} f(x).$$

It splits into:
- **Removable discontinuity:** the two one–sided limits are equal (so $\lim_{x\to x_0} f(x)$ exists), but $f(x_0)$ is either undefined or different from this common value.
- **Jump discontinuity:** the two one–sided limits exist but are different.

(2) **Second kind (essential discontinuity):** at least one of the one–sided limits $\lim_{x\to x_0^-} f(x)$, $\lim_{x\to x_0^+} f(x)$ does not exist (as a finite real limit).

EXAMPLE 5.5. (1) (First kind - jump)

$$f(x) = \begin{cases} 0, & x \le 0, \\ 1, & x > 0. \end{cases}$$

Discontinuous at $x = 0$ with a jump discontinuity; continuous on $\mathbb{R} \setminus \{0\}$.

(2) (Removable / first kind)

$$f(x) = \begin{cases} 0, & x = 0, \\ 1, & x \neq 0. \end{cases}$$

Here $\lim_{x\to 0} f(x) = 1 \neq f(0)$, so the discontinuity at 0 is removable. The function is continuous on $\mathbb{R} \setminus \{0\}$.

(3) (Second kind - Dirichlet's function)

$$f(x) = \begin{cases} 0, & x \in \mathbb{Q}, \\ 1, & x \notin \mathbb{Q}, \end{cases}$$

is discontinuous at every $x \in \mathbb{R}$; at each point the one-sided limits do not exist, so these are discontinuities of the second kind.

(4) (Continuous only at 0; second kind elsewhere)

$$f(x) = \begin{cases} x, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}. \end{cases}$$

This f is continuous only at $x = 0$ and is discontinuous (of the second kind) at every $x \neq 0$.

Introduction to Analysis Rui Wang Draft

(5) (Continuous at all irrationals; second kind of discontinuity at all rationals – Thomae's (popcorn) function.)

$$f(x) = \begin{cases} 0, & x \notin \mathbb{Q}, \\ \dfrac{1}{q}, & x = \dfrac{p}{q} \in \mathbb{Q} \text{ in lowest terms } (q \geq 1). \end{cases}$$

The proof is a homework problem (Rudin Chapter 4 – 18).

(6) (Second kind - oscillatory)

$$f(x) = \begin{cases} \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is discontinuous at $x = 0$; the left/right limits do not exist due to oscillation, hence a second-kind discontinuity.

(7) (Second kind - infinite discontinuity) Consider $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x}$. The point 0 is a limit point of the domain from both sides, and

$$\lim_{x \to 0^-} f(x) = -\infty, \qquad \lim_{x \to 0^+} f(x) = +\infty,$$

so $x = 0$ is a discontinuity of the second kind (an infinite discontinuity). If one defines $f(0) = 0$ to extend the function to $\mathbb{R}$, it remains discontinuous at 0.

5.2. Monotonic functions over segments in $\mathbb{R}$.

DEFINITION 5.6. A real valued function $f : (a, b) \to \mathbb{R}$ is called

(1) **increasing**, if any $a < x_1 \leq x_2 < b$, there is $f(x_1) \leq f(x_2)$.
(2) **decreasing**, if any $a < x_1 \leq x_2 < b$, there is $f(x_1) \geq f(x_2)$.
(3) **monotonic**, if it is increasing or decreasing.

THEOREM 5.7. *Suppose $f : (a, b) \to \mathbb{R}$ is an increasing function. Then both*

$$f(x_0+) := \lim_{x \to x_0+} f(x) \quad \textit{and} \quad f(x_0-) := \lim_{x \to x_0-} f(x)$$

exist for any $x_0 \in (a, b)$. More precisely,

$$\sup_{a<t<x_0} f(t) = f(x_0-) \leq f(x_0) \leq f(x_0+) = \inf_{x_0<t<b} f(t).$$

Further more, for any x, y with $a < x < y < b$, there is

$$f(x+) \leq f(y-).$$

PROOF. (1) Take any $x_0 \in (a, b)$. Consider the set $\{f(t) | a < t < x_0\}$. Since f is increasing, it has upper bound $f(x_0)$. Hence it must has the l.u.b., which we denote by

$$A := \sup_{a<t<x_0} f(t).$$

Now given any $\epsilon > 0$, there exists some $t_0 \in (a, x_0)$ so that

$$f(t_0) > A - \epsilon.$$

Notice f is increasing, and so any $t \in [t_0, x_0)$ has

$$A - \epsilon < f(t_0) \leq f(t) \leq A,$$

Introduction to Analysis Rui Wang Draft

which exactly states that $f(x_0-) = A$.

For the $f(x_0+)$ part the proof is similar and is left to you as exercise.

Then the inequality

$$\sup_{a<t<x_0} f(t) = f(x_0-) \le f(x_0) \le f(x_0+) = \inf_{x_0<t<b} f(t).$$

follows from the monotonicity of the the function f.

(2) Take some number $z \in (x, y)$, from the above together with f is increasing, there is

$$f(x+) = \inf_{x<t<b} f(t) \le \inf_{x<t<z} f(t) \le f(z) \le \sup_{z<t<y} f(t) \le \sup_{a<t<y} f(t) = f(y-).$$

□

COROLLARY 5.8. *Monotonic functions have no discontinuity of the second kind.*

PROOF. This is because we have proved that for any point $x \in (a, b)$, both $f(x-)$, $f(x+)$ exist. □

THEOREM 5.9. *Assume $f : (a, b) \to \mathbb{R}$ is monotonic. Then the set of discontinuous points is at most countable.*

PROOF. WLOG, assuming f is increasing. For each discontinuous point $x \in (a, b)$, there must be $f(x-) < f(x+)$. We choose some rational number $r_x \in \mathbb{Q}$ so that

$$f(x-) < r_x < f(x+).$$

By this way, we set up an injective (why injective?) map from discontinuous points in (a, b) to $\mathbb{Q}$, and so the set of discontinuous points must be at most countable.

□

EXAMPLE 5.10. Let $F : \mathbb{R} \to \mathbb{R}$ be the floor function $F(t) = \lfloor t \rfloor$, which has a first kind jump discontinuity at every integer.

Using it, we can construct an example on a finite interval with countable first kind of discontinuity. For example, consider the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ define

$$f(x) := F \circ \tan(x) = \lfloor \tan x \rfloor.$$

Since tan is strictly increasing and the floor map is nondecreasing, f is increasing. Its discontinuities are exactly at

$$x_n = \arctan(n), \quad n \in \mathbb{Z},$$

so f has countably many first kind jump discontinuities.

Introduction to Analysis Rui Wang Draft

CHAPTER 5

Differentiation

In this chapter, we focus on real valued functions defined on open or closed intervals in $\mathbb{R}$.

1. The derivative of a real function

DEFINITION 1.1. A function $f : [a, b] \to \mathbb{R}$ is called **differentiable** at a point $x_0 \in [a, b]$, if the limit of the function

$$\frac{f(t) - f(x_0)}{t - x_0}, \quad a < t < b, t \neq x_0$$

exists as $t \to x_0$. For this case, we write

$$f'(x_0) = \frac{d}{dx}|_{x=x_0} f = \lim_{t \to x_0} \frac{f(t) - f(x_0)}{t - x_0}.$$

It is called the **derivative** of f at x_0.

The function f is called **differentiable on** $[a, b]$ if it is differentiable for each $x \in [a, b]$. In this case, the function

$$\frac{df}{dx} = f' : [a, b] \to \mathbb{R},$$

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

PROPOSITION 1.2. *If $f : [a, b] \to \mathbb{R}$ is differentiable at $x_0 \in [a, b]$, then it is continuous at x_0.*

PROOF. Notice that

$$|f(x) - f(x_0)| = \frac{|f(x) - f(x_0)|}{|x - x_0|} \cdot |x - x_0|, \quad a \leq x \leq b, x \neq x_0.$$

The limit exists and it is zero as $x \to x_0$ when $f'(x_0)$ exists. It follows

$$\lim_{x \to x_0} f(x) = f(x_0),$$

and thus f is continuous at x_0. □

On an interval $[a, b]$, we write $C^k([a, b])$ for the class of functions $f : [a, b] \to \mathbb{R}$ whose derivatives $f^{(j)}$ exist on $[a, b]$ for $1 \leq j \leq k$. In particular, $C^1([a, b])$ the **continuously differentiable** functions, i.e., the derivative function is in $C^0([a, b])$, the class of continuous functions. We use $C^\infty([a, b])$ to denote the class of **smooth functions**, which have any order of derivatives.

EXAMPLE 1.3. Using the definition, one checks easily that

(1) the derivative of a constant function is 0;
(2) the derivative of the identity function is the constant 1.

THEOREM 1.4. *Suppose $f, g : [a, b] \to \mathbb{R}$ are differentiable at $x_0 \in [a, b]$. Then $f \pm g$, fg, and f/g (when $g(x_0) \neq 0$) are differentiable at x_0. Moreover,*

(1) $(f \pm g)'(x_0) = f'(x_0) \pm g'(x_0)$;
(2) *(Product rule).* $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$;

Introduction to Analysis
Rui Wang
Draft

(3) *(Quotient rule).* $(\frac{f}{g})'(x_0) = \frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g(x_0)^2}$.

PROOF. We give the proof of (2) and (3) as examples.

For (2), compute

$$\begin{aligned} & \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ = & \frac{(f(x) - f(x_0))g(x) + f(x_0)((g(x) - g(x_0))}{x - x_0} \\ = & \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0}. \end{aligned}$$

As $x \to x_0$, differentiability of f, g at x_0 gives

$$\frac{f(x) - f(x_0)}{x - x_0} \to f'(x_0), \quad \frac{g(x) - g(x_0)}{x - x_0} \to g'(x_0),$$

and by Proposition 1.2 (differentiability implies continuity) we also have $g(x) \to g(x_0)$. Hence the right-hand side tends to $f'(x_0)g(x_0) + f(x_0)g'(x_0)$, proving (2).

For (3), we first calculate $(\frac{1}{g(x)})'|_{x=x_0}$. By the assumption, $g(x_0) \neq 0$. By Proposition 1.2, differentiability implies continuity, $g(x) \neq 0$ for x in an open neighborhood of x_0. Compute

$$\begin{aligned} & \frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} \\ = & -\frac{g(x) - g(x_0)}{g(x)g(x_0)(x - x_0)} \end{aligned}$$

As $x \to x_0$, differentiability of g at x_0 gives

$$g(x) \to g(x_0), \quad \frac{g(x) - g(x_0)}{x - x_0} \to g'(x_0).$$

Hence there follows

$$(\frac{1}{g(x)})'|_{x=x_0} = -\frac{g'(x_0)}{g(x_0)^2}.$$

Then the quotient rule follows from the product rule to the function

$$\frac{f(x)}{g(x)} = f(x) \cdot \frac{1}{g(x)}.$$

□

THEOREM 1.5 (The Chain Rule). *Let $f : [a, b] \to \mathbb{R}$ be a real-valued function that is differentiable at $x_0 \in [a, b]$. Let g be a real-valued function defined on an interval that contains $f([a, b])$, and g is differentiable at $f(x_0)$. Then, the composition*

$$h(x) := g \circ f(x) := g(f(x)) : [a, b] \to \mathbb{R}$$

is differentiable at x_0, and the derivative at x_0 can be calculated as

$$h'(x_0) = g'(f(x_0))f'(x_0).$$

PROOF. We introduce some useful method for proofs regarding limits.

In general, for a function say $p(x)$, if

$$\lim_{x \to x_0} p(x) = 0,$$

Introduction to Analysis Rui Wang Draft

we write $p(x) = o(|x-x_0|)$. Here you can think $o(|x-x_0|)$ denotes a function that defined in a sufficiently small neighborhood of x_0 excluding x_0, which goes to 0 as $|x - x_0|$ goes to 0.

Using this notation, we can write

$$\begin{aligned} h(x) - h(x_0) &= g(f(x)) - g(f(x_0)) \\ &= (g'(f(x_0)) + o(|f(x) - f(x_0)|))(f(x) - f(x_0)) \\ &= (g'(f(x_0)) + o(|f(x) - f(x_0)|))(f'(x_0) + o(|x - x_0|))(x - x_0), \end{aligned}$$

and so

$$\frac{h(x) - h(x_0)}{x - x_0} = (g'(f(x_0)) + o(|f(x) - f(x_0)|))(f'(x_0) + o(|x - x_0|)).$$

By Proposition 1.2, the differentiability of $f(x)$ at x_0 implies $f(x)$ is continuous at x_0, i.e.,

$$|f(x) - f(x_0)| = o(|x - x_0|).$$

We can then write

$$\frac{h(x) - h(x_0)}{x - x_0} = (g'(f(x_0)) + o(|x - x_0|))(f'(x_0) + o(|x - x_0|)).$$

Take limit $x \to x_0$, it follows

$$\begin{aligned} \lim_{x\to x_0} \frac{h(x) - h(x_0)}{x - x_0} &= \lim_{x\to x_0} (g'(f(x_0)) + o(|x - x_0|))(f'(x_0) + o(|x - x_0|)) \\ &= g'(f(x_0))f'(x_0). \end{aligned}$$

□

REMARK 1.6. Rudin's book (Thm 5.5) states the chain rule assuming f is continuous on an interval. This is stronger than needed at a point: differentiability of f at x_0 already implies continuity at x_0, which is sufficient for the proof above.

EXAMPLE 1.7. (1) For any polynomial

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x],$$

from Theorem 1.4,

$$f'(x) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1} \in \mathbb{R}[x].$$

In particular, $f(x) \in C^\infty(\mathbb{R})$.

(2)

$$f(x) = \begin{cases} x^2 & x \le 0 \\ 0 & x > 0 \end{cases}$$

The function $f(x)$ is differentiable over $\mathbb{R}$ with

$$f'(x) = \begin{cases} 2x & x < 0 \\ 0 & x \ge 0 \end{cases}$$

but its derivative $f'(x)$ is not differentiable but only continuous at $x = 0$. (Notice: To get $f'(0)$, we need to use the definition of derivatives.)

Introduction to Analysis Rui Wang Draft

(3) The function

$$f(x)=\begin{cases}x^2\sin\frac{1}{x} & x\neq 0\\ 0 & x=0\end{cases}$$

has derivative function as

$$f(x)=\begin{cases}2x\sin\frac{1}{x}-\cos\frac{1}{x} & x\neq 0\\ 0 & x=0\end{cases}.$$

It is continuous at $x\neq 0$ and has $x=0$ as second kind discontinuous point.

Later, we will prove that if $f(x)$ is differentiable over $[a,b]$, its derivative $f'(x)$ has no first kind discontinuity.

(4) From this example, Volterra (1881) constructed functions which are differentiable everywhere but nowhere continuous. https://en.wikipedia.org/wiki/Volterra%27s_function.

(5) A continuous function may not be differentiable. For example, the function

$$f(x)=\begin{cases}x\sin\frac{1}{x} & x\neq 0\\ 0 & x=0\end{cases}$$

In fact, there are functions which are continuous everywhere but not differentiable anywhere. A family of famous examples is the Weierstrass functions. https://en.wikipedia.org/wiki/Weierstrass_function

2. Mean value theorem

DEFINITION 2.1. Let (X,d) be a metric space and $f: X\to\mathbb{R}$ a real-valued function.

- We say that f has a **local maximum** at $x_0\in X$ if there exists $\delta>0$ such that
$$f(x)\leq f(x_0)\quad \text{for all } x\in B_\delta(x_0).$$
- We say that f has a **local minimum** at $x_0\in X$ if there exists $\delta>0$ such that
$$f(x)\geq f(x_0)\quad \text{for all } x\in B_\delta(x_0).$$

DEFINITION 2.2. For a function $f:(a,b)\to\mathbb{R}$, a point $x_0\in(a,b])$ is called a **critical point** if f is not differentiable at x_0 or $f'(x_0)=0$.

THEOREM 2.3. *Assume f is defined over $[a,b]$. If f has a local maximum or local minimum at some $x_0\in(a,b)$, then x_0 is a critical point of f.*

PROOF. If f is not differentiable at x_0, we are done. Assume now f is differentiable at x_0 and x_0 is a local maximum.

Then there exists some $\delta>0$ so that

$$f(x_0)\geq f(x),\qquad \text{for any } x_0-\delta<x<x_0+\delta.$$

It follows

$$\frac{f(x)-f(x_0)}{x-x_0}\begin{cases}\geq 0 & x_0-\delta<x<x_0\\ \leq 0 & x_0<x<x_0+\delta.\end{cases}$$

Introduction to Analysis Rui Wang Draft

Further because $f'(x_0)$ exists, both one-side limits exist and

$$\lim_{x\to x_0-}\frac{f(x)-f(x_0)}{x-x_0}\geq 0,\quad \lim_{x\to x_0+}\frac{f(x)-f(x_0)}{x-x_0}\leq 0.$$

They must be equal, and then the common value, which is $f'(x_0)$, must be zero. □

REMARK 2.4. If a local maximum or minimum is attained at an endpoint a or b of the interval, then the derivative may NOT vanish at the corresponding endpoint.

Next, we are going to introduce several versions of mean value theorems. The first one is the following Rolle's theorem.

THEOREM 2.5 (Rolle's theorem). *Suppose the function $f(x)$ is continuous on $[a,b]$ and differentiable on (a,b). If there is further $f(a)=f(b)$, then there exists some $x_0\in(a,b)$ so that*

$$f'(x_0)=0.$$

PROOF. Since $f(x)$ is continuous on the closed interval $[a,b]$, by Theorem 3.5, both $\max_{[a,b]} f$ and $\min_{[a,b]} f$ exist.

- If $\max_{[a,b]} f=\min_{[a,b]} f$, then f is constant on $[a,b]$ and $f'(x)=0$ for all $x\in[a,b]$.
- If $\max_{[a,b]} f>\min_{[a,b]} f$, since $f(a)=f(b)$, there exists $x_0\in(a,b)$ that is a local maximum or a local minimum of f. By the differentiability of f and Theorem 2.3, $f'(x_0)=0$.

We are done. □

THEOREM 2.6 (The Mean Value Theorem). *If $f(x)$ is continuous over $[a,b]$, differentiable over (a,b), then there exists some $x_0\in(a,b)$ so that*

$$f(b)-f(a)=f'(x_0)(b-a).$$

PROOF. Consider the function

$$h(x)=f(x)-\frac{f(b)-f(a))}{b-a}\cdot x.$$

It is continuous over $[a,b]$, differentiable over (a,b) and

$$h(a)=h(b).$$

Then From the Rolle's theorem, there exists some $x_0\in(a,b)$ so that

$$h'(x_0)=f'(x_0)-\frac{f(b)-f(a))}{b-a}=0,$$

which is equivalent to

$$f(b)-f(a)=f'(x_0)(b-a).$$

□

We can further generalize the mean value theorem to the following form.

THEOREM 2.7 (Cauchy's Mean Value Theorem). *Assume f,g are two real-valued functions that are continuous over $[a,b]$ and differentiable over (a,b). Then there exists some $x_0\in(a,b)$ so that*

$$(f(b)-f(a))g'(x_0)=(g(b)-g(a))f'(x_0).$$

Introduction to Analysis Rui Wang Draft

PROOF. Consider the function

$$h(x) := (f(b) - f(a))g(x) - (g(b) - g(a))f(x).$$

Then $h \in C^0([a, b])$, differentiable over (a, b) and $h(a) = h(b)$. Apply the Rolle's theorem, there exists some $x_0 \in (a, b)$ so that

$$h'(x_0) = (f(b) - f(a))g'(x_0) - (g(b) - g(a))f'(x_0) = 0.$$

We are done. □

REMARK 2.8. Geometrically, Rolle's theorem, the Mean Value Theorem, and Cauchy's Mean Value Theorem all assert the existence of a point $c \in (a, b)$ where the tangent is parallel to the secant joining the endpoints. For Cauchy's Mean Value Theorem (applied to f, g),

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)},$$

so the tangent vector $(g'(c), f'(c))$ is parallel to the chord $(g(b) - g(a),\ f(b) - f(a))$ of the parametrized curve $x \mapsto (g(x), f(x))$.

In particular, taking $f(x) = x$ (so $f'(x) = 1$ and $f(b) - f(a) = b - a$), Cauchy's Mean Value Theorem reduces to the classical Mean Value Theorem for g:

$$g'(c) = \frac{g(b) - g(a)}{b - a}.$$

Now let's see two typical applications of the Mean Value Theorem.

THEOREM 2.9. *If $f(x)$ is differentiable over (a, b), then*

(1) *$f'(x) \geq 0$ for all x implies $f(x)$ is increasing;*
(2) *$f'(x) \leq 0$ for all x implies $f(x)$ is decreasing;*
(3) *$f'(x) = 0$ for all x implies $f(x)$ is constant.*

PROOF. Take any $a < x_1 < x_2 < b$, and apply the mean value theorem over $[x_1, x_2]$. There exists some $x_1 < x_0 < x_2$ so that

$$f(x_2) - f(x_1) = f'(x_0)(x_2 - x_1).$$

Then

(1) if $f' \geq 0$, then $f(x_2) \geq f(x_1)$. This shows f is increasing.
(2) if $f' \leq 0$, then $f(x_2) \leq f(x_1)$. This shows f is decreasing.
(3) if $f' = 0$, then $f(x_2) = f(x_1)$. This shows f is constant.

□

We know that a continuous function on a closed interval (compact) is uniformly continuous. On a non-compact open interval one useful sufficient criterion is:

PROPOSITION 2.10. *If $f : (a, b) \to \mathbb{R}$ is differentiable and its derivative is bounded on (a, b), i.e. there exists $M > 0$ with $|f'(x)| \leq M$ for all $x \in (a, b)$, then f is uniformly continuous on (a, b).*

PROOF. Fix $x, y \in (a, b)$ with $x \neq y$. By the Mean Value Theorem there exists x_0 between x and y such that

$$|f(x) - f(y)| = |f'(x_0)||x - y| \leq M|x - y|.$$

Thus for any $\epsilon > 0$, choosing $\delta = \epsilon/M$ gives $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$, proving uniform continuity. □

Introduction to Analysis Rui Wang Draft

REMARK 2.11. The converse need not hold: a function can be uniformly continuous on (a, b) without having a bounded derivative (or even being differentiable) on (a, b). Concrete examples are left to the reader.

3. The intermediate value property of derivatives

Recall from Theorem 4.17 that a continuous function on a closed interval attains every value between its endpoint values (the *intermediate value property*).

On the other hand, a derivative need not be continuous (e.g. $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$), but every derivative still enjoys the intermediate value property:

THEOREM 3.1 (The Intermediate Value Property for Derivatives). *Assume $f(x)$ is differentiable over $[a, b]$. Then for every μ strictly between $f'(a)$ and $f'(b)$, there exists $x_0 \in (a, b)$ such that $f'(x_0) = \mu$.*

PROOF. WLOG, we assume

$$f'(a) < \mu < f'(b).$$

Consider the function

$$g(x) = f(x) - \mu x,$$

which is differentiable and so continuous on $[a, b]$. It follows from Theorem 3.5 that $\min_{[a,b]} g$ exists in $[a, b]$.

Notice at endpoints,

$$g'(a) = f'(a) - \mu < 0, \quad g'(b) = f'(b) - \mu > 0.$$

Further, $g'(a) < 0$ implies that there exists some $a < x_1 < b$ so that $g(x_1) < g(a)$; and $g'(b) > 0$ implies that there exists some $a < x_2 < b$ so that $g(x_2) < g(b)$.

Thus neither a, b is the global minimum of g over $[a, b]$. In anther word, there is some $x_0 \in (a, b)$ so that

$$g(x_0) = \min_{[a,b]} g.$$

By Theorem 2.3,

$$g'(x_0) = 0,$$

which implies $f'(x_0) = \mu$.

We are done with the proof. □

COROLLARY 3.2. *Assume $f(x)$ is differentiable over $[a, b]$. Then $f'(x)$ have no discontinuities of the first kind. In another word, if $f'(x)$ is discontinuous at $x_0 \in [a, b]$, then x_0 is a discontinuous point of the second kind.*

PROOF. Assume the conclusion is not true, and $f'(x)$ is discontinuous at $x_0 \in (a, b)$ with both one-side limits exist. Then there are two possibilities:

(1) $\lim_{x \to x_0-} f'(x) \neq \lim_{x \to x_0+} f'(x)$;
(2) $\lim_{x \to x_0-} f'(x) = \lim_{x \to x_0+} f'(x) \neq f'(x_0)$.

We are going to obtain contradiction from both.

First for the case (1), WLOG, let's assume

$$\lim_{x \to x_0-} f'(x) < \lim_{x \to x_0+} f'(x), \quad \text{and write } \ell := \lim_{x \to x_0+} f'(x) - \lim_{x \to x_0-} f'(x)$$

which is a positive number.

duction to Analysis Rui Wang Draft

Then there exists some $\delta > 0$ so that

- for any $x_0 - \delta < x < x_0$, there is
$$f'(x) < \lim_{x \to x_0-} f'(x) + \frac{1}{3}\ell.$$
- for any $x_0 < x < x_0 + \delta$,
$$f'(x) > \lim_{x \to x_0+} f'(x) - \frac{1}{3}\ell.$$

Then in the interval $[x_0 - \frac{1}{2}\delta, x_0 + \frac{1}{2}\delta]$, there is at most one value in
$$(\lim_{x \to x_0-} f'(x) + \frac{1}{3}\ell, \lim_{x \to x_0-} f'(x) - \frac{1}{3}\ell)$$
which is taken by $f'(x_0)$. This contradicts with the intermediate value property Theorem 3.1.

For the case (2) and the endpoint cases are handled by the same arguments with one–sided versions of the statements and details are left to the reader.

□

REMARK 3.3. We have seen that a monotone function has only first–kind discontinuities; therefore, if $f'(x)$ is discontinuous at some point, $f'(x)$ cannot be monotone on any interval containing that point.

4. L'Hospital's Rule

In calculus we learned a powerful tool for evaluating limits of quotients in the indeterminate forms $0/0$ and ∞/∞: L'Hospital's[1] rule. Here are some elementary examples.

EXAMPLE 4.1. (1) $\lim_{x \to 0} \frac{\sin x}{x}$. Since $\sin x \to 0$ and $x \to 0$, this is of type $0/0$. By L'Hospital's rule,
$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{(\sin x)'}{(x)'} = \lim_{x \to 0} \frac{\cos x}{1} = 1.$$
(2) $\lim_{x \to +\infty} \frac{\log x}{x}$. As $x \to +\infty$, both $\log x$ and x tend to $+\infty$ (type ∞/∞). Hence,
$$\lim_{x \to +\infty} \frac{\log x}{x} = \lim_{x \to +\infty} \frac{(\log x)'}{(x)'} = \lim_{x \to +\infty} \frac{1/x}{1} = 0.$$

We now prove L'Hospital's rule. Since a two-sided limit exists iff both one-sided limits exist, it suffices to establish the one-sided version.

THEOREM 4.2. *Assume f, g are differentiable over (a, b) with $g(x) \neq 0$. If either*

(1) $\lim_{x \to a} f(x) = 0$ *and* $\lim_{x \to a} g(x) = 0$*; or*
(2) $\lim_{x \to a} |g(x)| = +\infty$,

and
$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = A \in [-\infty, +\infty], \quad \text{assuming } g'(x) \neq 0 \text{ over } (a, b),$$
then
$$\lim_{x \to a} \frac{f(x)}{g(x)} = A.$$

[1]The rule is named after Guillaume de L'Hospital (1661–1704), who published it in his 1696 textbook *Analyse des infiniment petits (pour l'intelligence des lignes courbes)*. However, surviving correspondence shows that the key ideas and proofs were due to Johann Bernoulli (1667–1748). In 1694, L'Hospital and Bernoulli made a private agreement under which Bernoulli sent him results that L'Hospital was free to use in his book.

Introduction to Analysis
Rui Wang
Draft

PROOF. We prove the case when $a, A \in \mathbb{R}$. Other cases are left to the reader as exercise.
Since $\lim_{x\to a} \frac{f'(x)}{g'(x)} = A$, then for any $\epsilon > 0$, there exists some $\delta > 0$ so that any

$$a < x < a + \delta,$$

there is

$$A - \epsilon < \frac{f'(x)}{g'(x)} < A + \epsilon.$$

Now we take any $x, y \in (a, a+\delta)$, by Cauchy's mean value theorem, there exists some $\xi \in (a, a+\delta)$, which may depend on x, y, so that

$$(f(x) - f(y))g'(\xi) = (g(x) - g(y))f'(\xi).$$

It follows

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(x) - f(y)}{g(x) - g(y)} \in (A - \epsilon, A + \epsilon). \tag{4.1}$$

(1) If $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, for each fixed y, we take $x \to a$, and it follows

$$\frac{f(y)}{g(y)} = \lim_{x\to a} \frac{f(x) - f(y)}{g(x) - g(y)} \in [A - \epsilon, A + \epsilon].$$

By definition of limit, this says

$$\lim_{y\to a} \frac{f(y)}{g(y)} = A.$$

(2) If $\lim_{x\to a} |g(x)| = +\infty$, then for each fixed y, we can make x be close enough to a so that $a < x < y < b$ and

$$\frac{g(x) - g(y)}{g(x)} > 0.$$

Multiplying it to (4.1), we obtain

$$(A - \epsilon) \cdot \frac{g(x) - g(y)}{g(x)} < \frac{f(x) - f(y)}{g(x)} < (A + \epsilon) \cdot \frac{g(x) - g(y)}{g(x)}.$$

Take $x \to a$, we obtain

$$A - \epsilon \le \liminf_{x\to a} \frac{f(x)}{g(x)} \le \limsup_{x\to a} \frac{f(x)}{g(x)} \le A + \epsilon.$$

At last, take $\epsilon \to 0$, we are done.

□

Here are more examples of L'Hospital's rule.

EXAMPLE 4.3. (1) $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$.

(2) $\lim_{x\to+\infty} \frac{x^2}{e^{3x}} = 0$.

(3) $\lim_{x\to 0+} x \log x = 0$. (Write $x \log x$ as $\frac{\log x}{\frac{1}{x}}$.)

(4) $\lim_{x\to 0+} \frac{\log x}{x} = -\infty$. Notice: L'Hospital Rule doesn't work for this example.

(5) $\lim_{x\to 0+} x^x = 1$. Use $x^x = e^{x \log x}$. (This recovers $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$.)

(6) $\lim_{x\to\infty}(1 + \frac{1}{x})^x = e$. Use $(1 + \frac{1}{x})^x = e^{x \log(1+\frac{1}{x})}$ and then use L'Hospital rule to calculate the limit of $x \log(1 + \frac{1}{x}) = \frac{\log(1+\frac{1}{x})}{1/x}$. (This recovers $\lim_{n\to\infty}(1 + \frac{1}{n})^n = e$.)

Introduction to Analysis Rui Wang Draft

5. Taylor expansion

5.1. The statement of Taylor theorem. The Taylor expansion can be viewed as a higher–order generalization of the mean value theorem.

EXAMPLE 5.1 (Zero order approximation). Consider a function $f : [a, b] \to \mathbb{R}$. We can regard the constant function

$$f_0(x) = f(a)$$

as the zero–order approximation of $f(x)$. Then we ask how well this approximation captures $f(x)$, i.e. we study the remainder

$$R(x) := f(x) - f(a), \quad x \in [a, b].$$

If we assume $f \in C^0([a, b])$ and f' exists on (a, b), then the mean value theorem tells us that for each $x \in (a, b]$ there exists some $a < \xi_x < x$ (here ξ_x emphasizes that ξ depends on x) such that

$$R(x) = f'(\xi_x)(x - a).$$

This says that the derivative of f controls the remainder $R(x)$ by a monomial of order 1.

For example, if on $[a, b]$, $|f'(x)| \le M$, then the remainder

$$|R(x)| \le M|x - a|.$$

EXAMPLE 5.2 (First order (linear) approximation). Again, consider a function $f : [a, b] \to \mathbb{R}$. The first–order approximation, i.e. the linear approximation at a is

$$f_1(x) = f(a) + f'(a)(x - a), \quad x \in [a, b].$$

Notice that the first order approximation shares the same value and derivative value at a with the original function $f(x)$, i.e.,

$$f_1(a) = f(a), \quad f_1'(a) = f(a).$$

The remainder now is

$$R(x) := f(x) - f_1(x), \quad x \in [a, b],$$

which satisfies

$$R(a) = 0, \quad R'(a) = 0.$$

We want to understand the relation of $R(x)$ and $(x - a)^2$.

Motivated by the proof of the mean value theorem from Rolle's theorem, we construct the function

$$h(t) := R(t) - \frac{R(x)}{(x - a)^2}(t - a)^2.$$

so that h vanishes both at a and x. Moreover, $h'(a) = 0$.

Apply Rolle's theorem to $h(t)$ on the interval $[a, x]$, there exists $a < \xi_x < x$ so that

$$h'(\xi_x) = 0.$$

Apply Rolle's theorem to $h'(t)$ on the interval $[a, \xi_x]$, there exists $a < \xi_x' < \xi_x$ so that

$$h''(\xi_x') = 0.$$

Notice that $R''(t) = f''(t)$,

$$h''(t) = f''(t) - 2\frac{R(x)}{(x - a)^2},$$

Introduction to Analysis Rui Wang Draft

and then evaluate at ξ_x'

$$0 = f''(\xi_x') - 2\frac{R(x)}{(x-a)^2}.$$

This shows

$$R(x) = \frac{f''(\xi_x')}{2}(x-a)^2.$$

This says that the second derivative of f controls the remainder $R(x)$ by a monomial of order 2.

For example, if on $[a,b]$, $|f''(x)| \le M$, then the remainder

$$|R(x)| \le \frac{M}{2}|x-a|^2.$$

Now we generalize the above two examples to higher orders. Assume $f : [a,b] \to \mathbb{R}$ has derivatives up to order n at a, i.e. $f'(a), \dots, f^{(n)}(a)$ exist. Define

$$\begin{aligned} f_n(x) &:= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad x \in [a,b]. \end{aligned}$$

This is a polynomial of order n and it is called the ***n*-th order approximation** of $f(x)$ at a.

The key property of the n-th order approximation is that

$$f^{(k)}(a) = f_n^{(k)}(a), \quad \text{for all} \quad k = 0, 1, \cdots, n. \tag{5.1}$$

THEOREM 5.3 (Taylor's Theorem). *Assume $f \in C^n([a,b])$ and $f^{(n+1)}$ exists on (a,b), where $n \in \{0,1,2,\dots\}$. Then for each $x \in (a,b]$ there exists some $a < \xi_x < x$ such that the remainder*

$$R(x) := f(x) - f_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x-a)^{n+1}.$$

PROOF. For each fixed $x \in (a,b)$, consider the function

$$h(t) := R(t) - \frac{R(x)}{(x-a)^{n+1}}(t-a)^{n+1},$$

defined for $t \in [a,x]$. It is in $C^n([a,b])$ and $n+1$-th order derivative exists on (a,b).

By (5.1),

$$h(a) = h(x) = 0, \quad h'(a) = 0, \cdots, h^{(n)}(a) = 0.$$

Apply Rolle's theorem to h on $[a,x]$: there exists $\xi_{1,x} \in (a,x)$ with $h'(\xi_{1,x}) = 0$.

Then apply Rolle's theorem to h' on $[a,\xi_{1,x}]$ to obtain $\xi_{2,x} \in (a,\xi_{1,x})$ with $h''(\xi_{2,x}) = 0$.

Repeating this procedure, after applying Rolle's theorem to $h^{(n)}$ on $[a,\xi_{n,x}]$ we obtain

$$\xi_{n+1,x} \in (a,\xi_{n,x}) \subset (a,x)$$

such that

$$h^{(n+1)}(\xi_{n+1,x}) = f^{(n+1)}(\xi_{n+1,x}) - (n+1)!\frac{R(x)}{(x-a)^{n+1}} = 0.$$

Denote $\xi_x := \xi_{n+1,x}$. Then

$$R(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x-a)^{n+1}.$$

□

Introduction to Analysis Rui Wang Draft

5.2. Taylor series.

DEFINITION 5.4. Assume $f \in C^\infty((a,b))$ (including $a=-\infty$, $b=+\infty$) and $c \in (a,b)$. We call the power series

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

the **Taylor series** of f centered at c.

Denote by

$$f_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

the n-th order approximation centered at x, and by

$$R_n(x) := f(x) - f_n(x)$$

the error term.

From Taylor's theorem, $R_n(x)$ can be controlled by $(n+1)$-th derivative of f by a monomial of order $(n+1)$.

In general, f may differ from its Taylor series, and by definition we have: for any fixed $x \in (a,b)$,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n \quad \Longleftrightarrow \quad \lim_{n\to\infty} R_n(x) = 0.$$

Taylor's theorem gives us a way to estimate $R_n(x)$, and hence a concrete tool to decide when $f(x)$ is equal to its Taylor series.

EXAMPLE 5.5 (Taylor series converging to f everywhere). Consider

$$f(x) = e^x, \qquad x \in \mathbb{R}.$$

For every $n \geq 0$ we have

$$f^{(n)}(x) = e^x, \quad \text{so} \quad f^{(n)}(0) = 1.$$

Thus the Taylor series of f about 0 is

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

This power series has radius of convergence $R = +\infty$, and for every $x \in \mathbb{R}$,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

So in this example, the Taylor series converges everywhere and agrees with f everywhere.

In fact, in this example we have used that the derivative of e^x is again e^x, which has not been proved yet.

In general, if

$$f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$$

is a power series with radius of convergence $R > 0$, then on $(c-R, c+R)$ the function f is infinitely differentiable, and its Taylor series about c is exactly this same power series:

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

Introduction to Analysis Rui Wang Draft

Equivalently, on the interval of convergence one may differentiate term by term:

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x-c)^{n-1}.$$

We will prove this general fact later in Chapter 7, by studying when differentiation can be interchanged with an infinite sum under appropriate uniform convergence assumptions.

EXAMPLE 5.6. Consider a polynomial

$$f(x) = 2 + 2x + 3x^2 + x^3.$$

Since a polynomial has only finitely many nonzero terms, its Taylor series centered at any point always converges to the original function.

For example, we can compute its Taylor expansion at $c = 1$ by rewriting in powers of $(x-1)$:

$$\begin{aligned} f(x) &= 2 + 2\big((x-1)+1\big) + 3\big((x-1)+1\big)^2 + \big((x-1)+1\big)^3 \\ &= 8 + 11(x-1) + 6(x-1)^2 + (x-1)^3. \end{aligned}$$

From this, we read off

$$f(1) = 8, \qquad f'(1) = 11, \qquad f''(1) = 12, \qquad f^{(3)}(1) = 6.$$

This example extends directly to arbitrary power series; the general proof relies on changing the order of summation when re-expanding about a new center (we will return to this point later).

EXAMPLE 5.7 (Taylor series convergent but not equal to f). Define

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0, \end{cases} \qquad x \in \mathbb{R}.$$

One can show that $f \in C^\infty(\mathbb{R})$ and, moreover, for every $n \geq 1$,

$$f^{(n)}(0) = 0.$$

Hence the Taylor series of f about 0 is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} 0 \cdot x^n \equiv 0,$$

which converges for all $x \in \mathbb{R}$. Clearly

$$f(x) \neq \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \quad \text{for } x \neq 0.$$

Thus the Taylor series converges everywhere, but it agrees with f only at the expansion point $x = 0$.

EXAMPLE 5.8 (Taylor series diverging at some points of the domain). Consider

$$f(x) = \frac{1}{1-x}, \qquad x \in \mathbb{R} \setminus \{1\}.$$

For $|x| < 1$, we have the geometric series expansion

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n,$$

which is the Taylor series of f about 0. This power series has radius of convergence $R = 1$.

Introduction to Analysis
Rui Wang
Draft

Thus, for every x with $|x| < 1$, the series

$$\sum_{n=0}^{\infty} x^n$$

converges and equals $f(x)$. But for any x with $|x| \geq 1$ (for example $x = 2$), the series

$$\sum_{n=0}^{\infty} 2^n$$

diverges, since its terms do not tend to 0.

So in this example, the Taylor series of f about 0 diverges at some points of the domain of f.

oduction to Analysis
Rui Wang
Draft

CHAPTER 6

The Riemann–Stieltjes Integral

The Riemann–Stieltjes integral is a mild but natural generalization of the Riemann integral: it gains significantly more flexibility while requiring only a small amount of extra machinery, so it is worthwhile to develop this more general notion first and we will start from here.

1. Definition of Riemann–Stieltjes Integral

Assume $[a, b]$ is a closed interval in $\mathbb{R}$. By a **partition** $\mathcal{P}$, we mean a finite set of points

$$a = x_0 \leq x_1 \leq \cdots \leq x_{n-1} \leq x_n = b.$$

Assume f is a bounded real-valued function over $[a, b]$ and α is an increasing function over $[a, b]$. Denote by

$$M_i = \sup_{[x_{i-1}, x_i]} f(x), \quad m_i = \inf_{[x_{i-1}, x_i]} f(x),$$

and by

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

Define the **upper sum** of f with respect to the partition and α as

$$U(f, \alpha; \mathcal{P}) := \Sigma_{i=1}^{n} M_i \Delta\alpha_i,$$

and the **lower sum** of f with respect to the partition and α as

$$L(f, \alpha; \mathcal{P}) := \Sigma_{i=1}^{n} m_i \Delta\alpha_i.$$

Define the **upper Riemann–Stieltjes integral** as

$$\overline{\int_a^b} f(x) d\alpha(x) := \inf_{\mathcal{P}} U(f, \alpha; \mathcal{P})$$

and the **lower Riemann–Stieltjes integral** as

$$\underline{\int_a^b} f(x) d\alpha(x) := \sup_{\mathcal{P}} L(f, \alpha; \mathcal{P}).$$

It is easy to see from definition that

$$\underline{\int_a^b} f(x) d\alpha(x) \leq \overline{\int_a^b} f(x) d\alpha(x).$$

DEFINITION 1.1. Call a function f is **Riemann–Stieltjes integrable with respect to α over** $[a, b]$, if

$$\overline{\int_a^b} f(x) d\alpha(x) = \underline{\int_a^b} f(x) d\alpha(x).$$

We use $\int_a^b f(x) d\alpha(x)$ to denote the common value and call it the **Riemann–Stieltjes integral** of f **with respect to** α over $[a, b]$.

We use the notation $R(\alpha)([a, b])$ to denote the set of Riemann–Stieltjes integrable functions with respect to α over $[a, b]$.

Introduction to Analysis Rui Wang Draft

In particular, when $\alpha(x) = x$, we call the corresponding Riemann–Stieltjes integration the **Riemann integration** and use $R([a,b])$ to denote the set of **Riemann integrable functions**.

EXAMPLE 1.2 (Constant functions). Consider the constant function $f(x) = c$ for all $x \in [a,b]$, where $c \in \mathbb{R}$. For any partition $\mathcal{P}$, we have $M_i = m_i = c$, so

$$U(f,\alpha;\mathcal{P}) = L(f,\alpha;\mathcal{P}) = c\sum_{i=1}^{n}\Delta\alpha_i = c\big(\alpha(b)-\alpha(a)\big),$$

independent of $\mathcal{P}$. Hence $f \in R(\alpha)([a,b])$ and

$$\int_a^b f(x)\,d\alpha(x) = c\big(\alpha(b)-\alpha(a)\big).$$

EXAMPLE 1.3 (The function $f(x) = x$). Let $f(x) = x$ on $[a,b]$, and let $\alpha : [a,b] \to \mathbb{R}$ be increasing. For an arbitrary partition

$$\mathcal{P} : a = x_0 \le x_1 \le \cdots \le x_{n-1} \le x_n = b,$$

we have

$$M_i = \sup_{[x_{i-1},x_i]} x = x_i, \qquad m_i = \inf_{[x_{i-1},x_i]} x = x_{i-1},$$

and

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

(1) In the special case $\alpha(x) = x$ we have $\Delta\alpha_i = x_i - x_{i-1}$. The upper and lower sums are

$$U(f,\alpha;\mathcal{P}) = \sum_{i=1}^{n} x_i\,(x_i - x_{i-1}), \qquad L(f,\alpha;\mathcal{P}) = \sum_{i=1}^{n} x_{i-1}\,(x_i - x_{i-1}).$$

Now take a sequence of equally spaced partitions $\mathcal{P}_n$, $n = 1,2,\dots$, given by

$$x_{n,i} = a + \frac{b-a}{n}\,i, \qquad i = 0,1,\dots,n.$$

Then for each n,

$$U(\mathcal{P}_n) = \sum_{i=1}^{n}\Big(a + \frac{b-a}{n}i\Big)\frac{b-a}{n}, \qquad L(\mathcal{P}_n) = \sum_{i=0}^{n-1}\Big(a + \frac{b-a}{n}i\Big)\frac{b-a}{n}.$$

By definition of upper and lower integrals,

$$L(\mathcal{P}_n) \;\le\; \underline{\int_a^b} x\,dx \;\le\; \overline{\int_a^b} x\,dx \;\le\; U(\mathcal{P}_n).$$

Set $h := \frac{b-a}{n}$. We calculate

$$L(\mathcal{P}_n) = \sum_{i=0}^{n-1}(a+ih)\,h = h\Big(na + h\sum_{i=0}^{n-1} i\Big) = h\Big(na + h\frac{(n-1)n}{2}\Big).$$

Substituting $h = \frac{b-a}{n}$ gives

$$L(\mathcal{P}_n) = a(b-a) + \frac{(b-a)^2}{n^2}\cdot\frac{(n-1)n}{2} = a(b-a) + \frac{(b-a)^2(n-1)}{2n}.$$

A short algebraic simplification shows

$$L(\mathcal{P}_n) = \frac{b^2-a^2}{2} - \frac{(b-a)^2}{2n}.$$

Introduction to Analysis Rui Wang Draft

Similarly,

$$U(\mathcal{P}_n) = \sum_{i=1}^{n} (a + ih)\, h = h\Big(na + h \sum_{i=1}^{n} i\Big) = h\Big(na + h\frac{n(n+1)}{2}\Big),$$

hence

$$U(\mathcal{P}_n) = a(b-a) + \frac{(b-a)^2}{n^2} \cdot \frac{n(n+1)}{2} = a(b-a) + \frac{(b-a)^2(n+1)}{2n} = \frac{b^2-a^2}{2} + \frac{(b-a)^2}{2n}.$$

For every n,

$$L(\mathcal{P}_n) \;\le\; \underline{\int_a^b} x\,dx \;\le\; \overline{\int_a^b} x\,dx \;\le\; U(\mathcal{P}_n),$$

and as $n \to \infty$,

$$\lim_{n\to\infty} L(\mathcal{P}_n) = \lim_{n\to\infty} U(\mathcal{P}_n) = \frac{b^2-a^2}{2}.$$

By the squeeze argument above, this implies

$$\int_a^b x\,dx = \frac{b^2-a^2}{2}.$$

(2) In general, the Riemann–Stieltjes integral depends on α. For example, if $\alpha(x) \equiv c$ is constant, then $\Delta\alpha_i = 0$ for all i, so every upper and lower sum is 0, and hence

$$\int_a^b x\,d\alpha(x) = 0.$$

(3) The following case shows the most interesting part of the Riemann–Stieltjes integral over the usual Riemann integral.

Fix $c \in (a, b)$ and constants $c_1 < c_2$. Define

$$\alpha(x) = \begin{cases} c_1, & a \le x < c, \\ c_2, & c \le x \le b, \end{cases}$$

which is increasing and has a single jump at $x = c$.

For any partition that includes c in some $[x_{i-1}, x_i]$, there are

$$U(\mathcal{P}) = x_i(c_2 - c_1), \qquad L(\mathcal{P}) = x_{i-1}(c_2 - c_1).$$

Then

$$L(\mathcal{P}) \;\le\; \underline{\int_a^b} x\,dx \;\le\; \overline{\int_a^b} x\,dx \;\le\; U(\mathcal{P}).$$

Shrinking the length of each small interval in the partitions, by the continuity of $f(x) = x$, both the upper sum and lower sum converge to $c(c_2 - c_1)$, this shows

$$\int_a^b x\,d\alpha(x) = c\,(c_2 - c_1).$$

(Exercise: Generalize it to any function which is continuous at the jumping point c.)

Introduction to Analysis Rui Wang Draft

2. Riemann–Stieltjes integrable functions

The first problem we need to understand is what kind of functions are Riemann–Stieltjes integrable.

THEOREM 2.1. *$f \in R(\alpha)([a,b])$ if and only if for each $\epsilon > 0$, there exists some partition $\mathcal{P}$ so that*

$$U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) < \epsilon.$$

PROOF. (1) Assume $f \in R(\alpha)([a,b])$, by definition

$$\inf_{\mathcal{P}} U(f,\alpha;\mathcal{P}) = \int_a^b f(x)d\alpha(x) = \sup_{\mathcal{P}} L(f,\alpha;\mathcal{P}).$$

For any $\epsilon > 0$, there exists partitions $\mathcal{P}_1, \mathcal{P}_2$ so that

$$U(f,\alpha;\mathcal{P}_1) < \int_a^b f(x)d\alpha(x) + \frac{1}{2}\epsilon$$

and

$$L(f,\alpha;\mathcal{P}_2) > \int_a^b f(x)d\alpha(x) - \frac{1}{2}\epsilon.$$

Consider the **common refinement** $\mathcal{P}$ by making union of points in $\mathcal{P}_1$ and $\mathcal{P}_2$. It follows

$$L(f,\alpha;\mathcal{P}_2) \le L(f,\alpha;\mathcal{P}) \le U(f,\alpha;\mathcal{P}) \le U(f,\alpha;\mathcal{P}_1),$$

and then

$$\begin{aligned} & U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) \\ \le\ & U(f,\alpha;\mathcal{P}_1) - L(f,\alpha;\mathcal{P}_2) \\ \le\ & (\int_a^b f(x)d\alpha(x) + \frac{1}{2}\epsilon) - (\int_a^b f(x)d\alpha(x) - \frac{1}{2}\epsilon) \\ =\ & \epsilon. \end{aligned}$$

(2) Assume that for each $\epsilon > 0$, there exists some partition $\mathcal{P}$ so that

$$U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) < \epsilon,$$

then it follows

$$0 \le \overline{\int_a^b} f(x)d\alpha(x) - \underline{\int_a^b} f(x)d\alpha(x) \le \epsilon.$$

Take $\epsilon \to 0$, we are done.

□

Now we use this criterion to prove the following several theorems.

THEOREM 2.2. *$C^0([a,b]) \subseteq R(\alpha)([a,b])$.*

PROOF. If α is constant, then by definition, $\int_a^b f(x)d\alpha(x) = 0$. In the following, we assume $\alpha(a) < \alpha(b)$.

For any $f \in C^0([a,b])$, it is uniformly continuous over $[a,b]$. Hence for any $\epsilon > 0$, there exists some $\delta > 0$ so that

$$|f(t_1) - f(t_2)| < \frac{1}{2(\alpha(b) - \alpha(a))}\epsilon$$

whenever $|t_1 - t_2| < \delta$.

Now we take a partition $\mathcal{P}$ so that $\Delta_i = x_i - x_{i-1} < \delta$. Then it follows

$$M_i - m_i < \frac{1}{\alpha(b) - \alpha(a)}\epsilon.$$

Introduction to Analysis Rui Wang Draft

For this partition $\mathcal{P}$,

$$\begin{aligned} & U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) \\ = \ & \Sigma_i (M_i - m_i)\Delta\alpha_i \\ \leq \ & \Sigma_i \frac{1}{\alpha(b)-\alpha(a)}\epsilon\Delta\alpha_i \\ = \ & \epsilon. \end{aligned}$$

Apply Theorem 2.1, we proved $f \in R(\alpha)([a,b])$.

□

THEOREM 2.3. *Assume f is monotonic and α is continuous on $[a,b]$, then $f \in R(\alpha)([a,b])$.*

PROOF. If f is constant C, then by definition, we can see $\int_a^b f(x)d\alpha(x) = C(\alpha(b)-\alpha(a))$. Otherwise, WLOG, we can assume

$$f(b) - f(a) > 0.$$

Sine $\alpha \in C^0([a,b])$ implies α is uniformly continuous over $[a,b]$. Then for any $\epsilon > 0$, there exists some $\delta > 0$ so that

$$|\alpha(t_1) - \alpha(t_2)| < \frac{1}{2(f(b)-f(a))}\epsilon$$

whenever $|t_1 - t_2| < \delta$.

Now we take a partition $\mathcal{P}$ so that $\Delta_i = x_i - x_{i-1} < \delta$. Then it follows

$$\Delta\alpha_i < \frac{1}{f(b)-f(a)}\epsilon.$$

For this partition $\mathcal{P}$, there is

$$\begin{aligned} & U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) \\ = \ & \Sigma_i (f(x_i) - f(x_{i-1}))\Delta\alpha_i \\ \leq \ & \Sigma_i (f(x_i) - f(x_{i-1}))\frac{1}{f(b)-f(a)}\epsilon \\ = \ & \epsilon. \end{aligned}$$

Apply Theorem 2.1, we have $f \in R(\alpha)([a,b])$.

□

THEOREM 2.4. *Assume f is bounded over $[a,b]$ with only finitely many discontinuous points and α is continuous on these points. Then $f \in R(\alpha)([a,b])$.*

PROOF. Assume f discontinuous at $p_1 < p_2 < \cdots < p_m \in [a,b]$. For each $i = 1, \cdots, m$, since α is continuous at p_i, for any $\epsilon > 0$ there exists $\delta_i > 0$, so that any $|x - p_i| < \delta_i$,

$$|\alpha(x) - \alpha(p_i)| < \epsilon.$$

Moreover, we can shrink δ_i's so that these intervals $[p_i - \delta_i, p_i + \delta_i]$ have no intersection.

On the other hand, f is continuous over the complement of the union of interiors of these intervals, which we denote by $K \subseteq [a,b]$. In fact, K is the union of finite closed intervals as

$$K = [a, p_1 - \delta_1] \cup [p_1 + \delta_1, p_2 - \delta_2] \cup \cdots \cup [p_{m-1} + \delta_{m-1}, p_m - \delta_m] \cup [p_m + \delta_m, b].$$

We denote them by $K_0, K_1, \cdots, K_{m-1}, K_m$ one by one from left to right.

Introduction to Analysis Rui Wang Draft

It follows from Theorem 2.2 and Theorem 2.1 that, for each K_j, $j = 0, 1, \cdots, m$, there exists a partition $\mathcal{P}_j$ so that

$$U(f|_{K_j}, \alpha; \mathcal{P}_j) - L(f|_{K_j}, \alpha; \mathcal{P}_j) < \epsilon.$$

Now we consider a partition $\mathcal{P}$ for $[a, b]$ whose points are the union of partitions $\mathcal{P}_0, \cdots, \mathcal{P}_m$. It follows from our construction that

$$\begin{aligned} & U(f, \alpha; \mathcal{P}) - L(f, \alpha; \mathcal{P}) \\ = \ & \Sigma_{j=0}^{m}(U(f|_{K_j}, \alpha; \mathcal{P}_j) - L(f|_{K_j}, \alpha; \mathcal{P}_j)) + \Sigma_{i=1}^{m}(\sup_{[p_i-\delta_i, p_i+\delta_i]} f - \inf_{[p_i-\delta_i, p_i+\delta_i]} f)(\alpha(p_i+\delta_i) - \alpha(p_i-\delta_i)) \\ \leq \ & (m+1)\epsilon + 2Mm\epsilon \\ = \ & (m+1+2Mm)\epsilon. \end{aligned}$$

Here M is a fixed number so that $|f(x)| \leq M$ for any $x \in [a, b]$.

Apply Theorem 2.1, we have $f \in R(\alpha)([a, b])$.

□

THEOREM 2.5. *Assume $f \in R(\alpha)([a, b])$ with $m \leq f \leq M$ and $g \in C^0([m, M])$. Then $g \circ f \in R(\alpha)([a, b])$.*

PROOF. First, $g \in C^0([m, M])$ implies g is uniformly continuous over $[m, M]$. Hence for any $\epsilon > 0$, there exists $\delta > 0$ so that any

$$|g(y_1) - g(y_2)| < \epsilon \quad \text{for any} \quad |y_1 - y_2| < \delta.$$

Next, use $f \in R(\alpha)([a, b])$, there exists a partition $\mathcal{P}$ for $[a, b]$ so that

$$U(f, \alpha; \mathcal{P}) - L(f, \alpha; \mathcal{P}) < \delta^2.$$

Each $[x_{i-1}, x_i]$ determined by the partition $\mathcal{P}$ belongs to either one of the following two cases:

(1) $\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f < \delta$. We denote by $\mathcal{P}_1$ the sub-partition that contains such intervals. Over such intervals,

$$\sup_{[x_{i-1}, x_i]} g \circ f - \inf_{[x_{i-1}, x_i]} g \circ f < \epsilon.$$

(2) $\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \geq \delta$. We denote by $\mathcal{P}_2$ the sub-partition that contains such intervals. Over $\mathcal{P}_2$, we have

$$\delta \Sigma_{i \in \mathcal{P}_2} \Delta\alpha_i \leq U(f, \alpha; \mathcal{P}_2) - L(f, \alpha; \mathcal{P}_2) \leq U(f, \alpha; \mathcal{P}) - L(f, \alpha; \mathcal{P}) < \delta^2,$$

and hence

$$\Sigma_{i \in \mathcal{P}_2} \Delta\alpha_i < \delta.$$

We can further shrink δ so that $\delta < \epsilon$.

Now we consider $\mathcal{P}_1, \mathcal{P}_2$ together and obtain

$$\begin{aligned} & U(g \circ f, \alpha; \mathcal{P}) - L(g \circ f, \alpha; \mathcal{P}) \\ = \ & \Sigma_{i \in \mathcal{P}_1}(\sup_{[x_{i-1}, x_i]} g \circ f - \inf_{[x_{i-1}, x_i]} g \circ f)\Delta\alpha_i + \Sigma_{i \in \mathcal{P}_2}(\sup_{[x_{i-1}, x_i]} g \circ f - \inf_{[x_{i-1}, x_i]} g \circ f)\Delta\alpha_i \\ \leq \ & \epsilon \Sigma_{i \in \mathcal{P}_1} \Delta\alpha_i + 2C \Sigma_{i \in \mathcal{P}_2} \Delta\alpha_i \\ \leq \ & \epsilon(\Sigma_{i \in \mathcal{P}_1} \Delta\alpha_i + 2C) \\ \leq \ & \epsilon(\alpha(b) - \alpha(a) + 2C). \end{aligned}$$

Introduction to Analysis Rui Wang Draft

Here C is an upper bound of $|g|$ over $[m, M]$.

At last, apply Theorem 2.1, we obtain $g \circ f \in R(\alpha)([a, b])$.

□

EXAMPLE 2.6. If $f \in R(\alpha)([a, b])$, then $f^2 \in R(\alpha)([a, b])$.

3. Properties of the integral

THEOREM 3.1. (1) *Linearity of f:*

(a) *If $f_1, f_2 \in R(\alpha)([a, b])$, then $f_1 + f_2 \in R(\alpha)([a, b])$ and*

$$\int_a^b (f_1 + f_2)d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha.$$

(b) *If $f \in R(\alpha)([a, b])$ and $c \in \mathbb{R}$, then $cf \in R(\alpha)([a, b])$ and*

$$\int_a^b (cf)d\alpha = c\int_a^b f\, d\alpha.$$

(2) *Linearity of α:*

(a) *If $f \in R(\alpha_1)([a, b]) \cap R(\alpha_2)([a, b])$, then $f \in R(\alpha_1 + \alpha_2)([a, b])$ and*

$$\int_a^b f\, d(\alpha_1 + \alpha_2) = \int_a^b f\, d\alpha_1 + \int_a^b f\, d\alpha_2.$$

(b) *If $f \in R(\alpha)([a, b])$ and $c \geq 0$, then $f \in R(c\alpha)([a, b])$ and*

$$\int_a^b f\, d(c\alpha) = c\int_a^b f\, d\alpha.$$

(3) *If $f_1, f_2 \in R(\alpha)([a, b])$ and $f_1 \leq f_2$, then*

$$\int_a^b f_1 d\alpha \leq \int_a^b f_2 d\alpha.$$

(4) *If $f \in R(\alpha)([a, b])$ and $c \in [a, b]$, then $f \in R(\alpha)([a, c]) \cap R(\alpha)([c, b])$ and*

$$\int_a^b f\, d\alpha = \int_a^c f\, d\alpha + \int_c^b f\, d\alpha.$$

(5) *If $f \in R(\alpha)([a, b])$ and $|f| \leq M$, then*

$$\left|\int_a^b f\, d\alpha\right| \leq M(\alpha(b) - \alpha(a)).$$

(6) *If $f, g \in R(\alpha)([a, b])$, then $fg \in R(\alpha)([a, b])$.*

(7) *If $f \in R(\alpha)([a, b])$, then $|f| \in R(\alpha)([a, b])$ and*

$$\left|\int_a^b f\, d\alpha\right| \leq \int_a^b |f| d\alpha.$$

PROOF. (1) (a) First we notice that for any partition $\mathcal{P}$ over $[a, b]$, we have the following inequality

$$\begin{aligned} L(f_1, \alpha; \mathcal{P}) + L(f_2, \alpha; \mathcal{P}) &\leq L(f_1 + f_2, \alpha; \mathcal{P}) \\ &\leq U(f_1 + f_2, \alpha; \mathcal{P}) \leq U(f_1, \alpha; \mathcal{P}) + U(f_2, \alpha; \mathcal{P}). \end{aligned}$$

Hence

$$\begin{aligned} &U(f_1 + f_2, \alpha; \mathcal{P}) - L(f_1 + f_2, \alpha; \mathcal{P}) \\ \leq\ &(U(f_1, \alpha; \mathcal{P}) - L(f_1, \alpha; \mathcal{P})) + (U(f_2, \alpha; \mathcal{P}) - L(f_2, \alpha; \mathcal{P})). \end{aligned} \tag{3.1}$$

Introduction to Analysis Rui Wang Draft

Now since $f_1, f_2 \in R(\alpha)([a,b])$, using Theorem 2.1, for any $\epsilon > 0$, there exists partitions $\mathcal{P}_1, \mathcal{P}_2$ for $[a,b]$ so that

$$\begin{aligned} U(f_1,\alpha;\mathcal{P}_1) - L(f_1,\alpha;\mathcal{P}_1) &< \epsilon/2 \\ U(f_2,\alpha;\mathcal{P}_2) - L(f_2,\alpha;\mathcal{P}_2) &< \epsilon/2. \end{aligned}$$

Take $\mathcal{P}$ as a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$, it follows

$$\begin{aligned} U(f_1,\alpha;\mathcal{P}) - L(f_1,\alpha;\mathcal{P}) \le U(f_1,\alpha;\mathcal{P}_1) - L(f_1,\alpha;\mathcal{P}_1) < \epsilon/2 \\ U(f_2,\alpha;\mathcal{P}) - L(f_2,\alpha;\mathcal{P}) \le U(f_2,\alpha;\mathcal{P}_2) - L(f_2,\alpha;\mathcal{P}_2) < \epsilon/2. \end{aligned}$$

Connect it with (3.1), we get

$$U(f_1+f_2,\alpha;\mathcal{P}) - L(f_1+f_2,\alpha;\mathcal{P}) < \epsilon$$

and then Theorem 2.1 implies $f_1 + f_2 \in R(\alpha)([a,b])$.

Now for any $\epsilon > 0$, take a partition $\mathcal{P}_i$, $i = 1, 2$, so that

$$\int_a^b f_i d\alpha \le U(f_i,\alpha;\mathcal{P}_i) \le \int_a^b f_i d\alpha + \epsilon.$$

Assume $\mathcal{P}$ is a common refinement of $\mathcal{P}_i$. It follows the following inequality

$$U(f_1+f_2,\alpha;\mathcal{P}) \le U(f_1,\alpha;\mathcal{P}) + U(f_2,\alpha;\mathcal{P}) \le U(f_1,\alpha;\mathcal{P}_1) + U(f_2,\alpha;\mathcal{P}_2),$$

and then

$$\int_a^b (f_1+f_2)d\alpha \le U(f_1+f_2,\alpha;\mathcal{P}) \le \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha + 2\epsilon.$$

Since $\epsilon > 0$ can be arbitrarily small, this proves

$$\int_a^b (f_1+f_2)d\alpha \le \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha.$$

Similarly, using lower sums, we will obtain

$$\int_a^b (f_1+f_2)d\alpha \ge \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha.$$

(b) Similar to (a) and is left to you.

(2) Homework problem.

(3) Homework problem.

(4) Homework problem.

(5) Now for any partition $\mathcal{P}$,

$$U(f,\alpha;\mathcal{P}) \le U(M,\alpha;\mathcal{P}) = M(\alpha(b)-\alpha(a)).$$

Since $f \in R(\alpha)([a,b])$, we have

$$\int_a^b f\,d\alpha = \overline{\int_a^b} f\,d\alpha = \inf_{\mathcal{P}} U(f,\alpha;\mathcal{P}) \le M(\alpha(b)-\alpha(a)).$$

Similarly,

$$L(-f,\alpha;\mathcal{P}) \ge L(-M,\alpha;\mathcal{P}) = -M(\alpha(b)-\alpha(a)).$$

Since $-f \in R(\alpha)([a,b])$ (from 1(b)), we have

$$-\int_a^b f\,d\alpha = \int_a^b (-f)d\alpha = \underline{\int_a^b} (-f)d\alpha = \sup_{\mathcal{P}} L(-f,\alpha;\mathcal{P}) \ge -M(\alpha(b)-\alpha(a)).$$

Introduction to Analysis Rui Wang Draft

Together this proves

$$|\int_a^b f d\alpha| \leq M(\alpha(b) - \alpha(a)).$$

(6) Notice the following equality

$$fg = \frac{1}{4}((f+g)^2 - (f-g)^2).$$

Then $fg \in R(\alpha)([a,b])$ follows from Property 1 and Theorem 2.5.

(7) $|f| \in R(\alpha)([a,b])$ follows from the continuity of the absolute value function and Theorem 2.5. Then inequality can be proved similarly as for (5).

□

REMARK 3.2. For the property (7) above, notice that $|f| \in R(\alpha)([a,b])$ doesn't imply $f \in R(\alpha)([a,b])$. Consider the following example

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \cap [0,1] \\ -1 & x \in \mathbb{Q}^c \cap [0,1]. \end{cases}$$

It is not Riemann integrable over $[0,1]$ (Excise: prove this) but $|f|$ is Riemann integrable.

The following three theorems are related to the useful formula of "substitution" for integration in calculus.

THEOREM 3.3. *Assume α' exists and $\alpha' \in R([a,b])$. Assume f is bounded over $[a,b]$. Then $f \in R(\alpha)([a,b])$ if and only if $f\alpha' \in R([a,b])$. In that case,*

$$\int_a^b f(x)d\alpha(x) = \int_a^b f(x)\alpha'(x)dx.$$

PROOF. Since $\alpha' \in R([a,b])$, for any $\epsilon > 0$, there exists a partition $\mathcal{P}$ for $[a,b]$ so that

$$U(\alpha';\mathcal{P}) - L(\alpha';\mathcal{P}) < \epsilon. \tag{3.2}$$

At the same time, apply the mean value theorem over each interval $[x_{i-1}, x_i]$ from the partition $\mathcal{P}$, there exist points $t_i \in (x_{i-1}, x_i)$ so that

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(t_i)(x_i - x_{i-1}) = \alpha'(t_i)\Delta_i.$$

If follows

$$U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) = \Sigma_i(M_i - m_i)\Delta\alpha_i = \Sigma_i(M_i - m_i)\alpha'(t_i)\Delta_i = \Sigma_i(M_i\alpha'(t_i) - m_i\alpha'(t_i))\Delta_i.$$

On the other hand, notice that there exist points $s_i, s_i' \in (x_{i-1}, x_i)$

$$\begin{aligned} & \Sigma_i(\sup_{[x_{i-1},x_i]}(f\alpha') - \inf_{[x_{i-1},x_i]}(f\alpha'))\Delta_i \\ \leq\ & \Sigma_i(M_i \sup_{[x_{i-1},x_i]}\alpha' - m_i \inf_{[x_{i-1},x_i]}\alpha')\Delta_i \\ \leq\ & \Sigma_i(M_i\alpha'(s_i) - m_i\alpha'(s_i') + \frac{\epsilon}{b-a})\Delta_i \\ =\ & \Sigma_i(M_i\alpha'(s_i) - m_i\alpha'(s_i'))\Delta_i + \epsilon. \end{aligned}$$

Introduction to Analysis Rui Wang Draft

Now if $f \in R(\alpha)([a,b])$, then $|f|$ is bounded by some $C > 0$ and

$$\begin{aligned}
& \Sigma_i(\sup_{[x_{i-1},x_i]}(f\alpha') - \inf_{[x_{i-1},x_i]}(f\alpha'))\Delta_i \\
\leq\ & \Sigma_i(M_i\alpha'(s_i) - m_i\alpha'(s_i'))\Delta_i + \epsilon \\
\leq\ & \Sigma_i(M_i\alpha'(t_i) - m_i\alpha'(t_i))\Delta_i + \epsilon + \Sigma_i(M_i|\alpha'(s_i) - \alpha'(t_i)| - m_i|\alpha'(s_i') - \alpha'(t_i)|)\Delta_i \\
\leq\ & \Sigma_i(M_i\alpha'(t_i) - m_i\alpha'(t_i))\Delta_i + \epsilon + 2C(U(\alpha';\mathcal{P}) - L(\alpha';\mathcal{P})) \\
=\ & U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) + \epsilon + 2C\epsilon.
\end{aligned}$$

By taking a refinement of $\mathcal{P}$ which we still denote by $\mathcal{P}$, we can make $U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) < \epsilon$. Then using Theorem 2.1, we proved $f\alpha' \in R([a,b])$.

Now assume $f\alpha' \in R([a,b])$. There exist $r_i, r_i' \in [x_{i-1}, x_i]$ so that the following estimates hold

$$\begin{aligned}
& U(f,\alpha;\mathcal{P}) - L(f,\alpha;\mathcal{P}) \\
=\ & \Sigma_i(M_i - m_i)\Delta\alpha_i \\
\leq\ & \Sigma_i(f(r_i) - f(r_i'))\Delta\alpha_i + \epsilon(\alpha(b) - \alpha(a)) \\
=\ & \Sigma_i(f(r_i) - f(r_i'))\alpha'(t_i)\Delta_i + \epsilon(\alpha(b) - \alpha(a)) \\
=\ & \Sigma_i(f(r_i)\alpha'(t_i) - f(r_i')\alpha'(t_i))\Delta_i + \epsilon(\alpha(b) - \alpha(a)) \\
\leq\ & \Sigma_i(f(r_i)\alpha'(r_i) - f(r_i')\alpha'(r_i'))\Delta_i + \Sigma_i(f(r_i)|\alpha'(r_i) - \alpha'(t_i)| - f(r_i')|\alpha'(r_i') - \alpha'(t_i)|)\Delta_i + \epsilon(\alpha(b) - \alpha(a)) \\
\leq\ & \Sigma_i(\sup_{[x_{i-1},x_i]}(f\alpha') - \inf_{[x_{i-1},x_i]}(f\alpha'))\Delta_i + 2C\epsilon + \epsilon(\alpha(b) - \alpha(a)).
\end{aligned}$$

If then follows from Theorem 2.1 that $f \in R(\alpha)([a,b])$.

At last, the equality follows from the same estimates as above but for U and L separately: For example, we have

$$\begin{aligned}
& U(f,\alpha;\mathcal{P}) \\
=\ & \Sigma_i M_i\Delta\alpha_i \\
\leq\ & \Sigma_i f(r_i)\Delta\alpha_i + \epsilon(\alpha(b) - \alpha(a)) \\
=\ & \Sigma_i f(r_i)\alpha'(t_i)\Delta_i + \epsilon(\alpha(b) - \alpha(a)) \\
=\ & \Sigma_i f(r_i)\alpha'(t_i)\Delta_i + \epsilon(\alpha(b) - \alpha(a)) \\
\leq\ & \Sigma_i f(r_i)\alpha'(r_i)\Delta_i + \Sigma_i f(r_i)|\alpha'(r_i) - \alpha'(t_i)|\Delta_i + \epsilon(\alpha(b) - \alpha(a)) \\
\leq\ & \Sigma_i \sup_{[x_{i-1},x_i]}(f\alpha')\Delta_i + C\epsilon + \epsilon(\alpha(b) - \alpha(a)).
\end{aligned}$$

Take infimum of $\mathcal{P}$ and let $\epsilon \to 0$, it follows

$$\int_a^b f(x)d\alpha(x) \leq \int_a^b f(x)\alpha'(x)dx.$$

The other direction can be obtained similarly and we skip details. □

THEOREM 3.4 (Change of variable). *Assume $f \in R(\alpha)([a,b])$. Assume φ is strictly increasing and continuous that maps interval $[A,B]$ to $[a,b]$. Define*

$$\beta = \alpha \circ \varphi$$

Introduction to Analysis Rui Wang Draft

which is increasing on $[A, B]$*, and define*

$$g = f \circ \varphi.$$

Then $g \in R(\beta)([A, B])$ *and*

$$\int_A^B g d\beta = \int_a^b f d\alpha.$$

PROOF. By the strictly increasing property of φ, each partition $\mathcal{P} = \{x_i\}$ for $[a, b]$ corresponds to a partition $\mathcal{P}' = \{y_i\}$ for $[A, B]$ with $x_i = \varphi(y_i)$, and

$$U(f, \alpha; \mathcal{P}) = U(g, \beta; \mathcal{P}'), \quad L(f, \alpha; \mathcal{P}) = L(g, \beta; \mathcal{P}').$$

The conclusion then immediately follow from definition of integration. □

Using it, we obtain the following important formula for change of variables.

THEOREM 3.5. *Assume* $f \in R(\alpha)([a, b])$*. Assume* φ *is strictly increasing that maps interval* $[A, B]$ *to* $[a, b]$ *and* $\varphi' \in R([A, B])$*. Then*

$$\int_a^b f(x)dx = \int_A^B f(\varphi(y))\varphi'(y)dy.$$

4. Fundamental theorem of calculus

THEOREM 4.1. *Assume* $f \in R([a, b])$*. For* $a \le x \le b$*, define*

$$F(x) := \int_a^x f(t)dt.$$

Then $F \in C^0([a, b])$*. Furthermore, if* f *is continuous at a point* $x_0 \in [a, b]$*, then* F *is differentiable at* x_0*, and*

$$F'(x_0) = f(x_0).$$

PROOF. (1) Because $f \in R([a, b])$, it must be bounded. (Why?) Assume $|f| \le M$. Consider any $x, y \in [a, b]$,

$$|F(y) - F(x)| = |\int_a^y f(t)dt - \int_a^x f(t)dt| = |\int_x^y f(t)dt| \le M|y - x|.$$

This proves that F is uniformly continuous over $[a, b]$, hence continuous.

(2) Assume f is continuous at a point $x_0 \in [a, b]$, then for any $\epsilon > 0$, there exists $\delta > 0$ so that any $|x - x_0| < \delta$, there is

$$|f(x) - f(x_0)| < \epsilon.$$

Now consider any $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$ and $x \ne x_0$. We have

$$\begin{aligned}
|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)| &= |\frac{1}{x - x_0}\int_{x_0}^x f(t)dt - f(x_0)| \\
&= |\frac{1}{x - x_0}\int_{x_0}^x f(t)dt - \frac{1}{x - x_0}\int_{x_0}^x f(x_0)dt| \\
&= |\frac{1}{x - x_0}\int_{x_0}^x (f(t) - f(x_0))dt| \\
&\le \frac{\int_{x_0}^x |f(t) - f(x_0)|dt}{|x - x_0|} \\
&< \frac{\epsilon|x - x_0|}{|x - x_0|} = \epsilon.
\end{aligned}$$

Introduction to Analysis Rui Wang Draft

□

Usually, we call such F an antiderivative of f.

THEOREM 4.2 (The fundamental theorem of calculus). *If $f \in R([a,b])$ and if there is a differentiable function F on $[a,b]$ so that*

$$F' = f,$$

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

PROOF. Since $f \in R([a,b])$, for any $\epsilon > 0$, there exists a partition $\mathcal{P}$ for $[a,b]$ so that

$$U(f;\mathcal{P}) - L(f;\mathcal{P}) < \epsilon$$

and

$$|U(f;\mathcal{P}) - \int_a^b f(x)dx| < \epsilon.$$

In each interval $[x_{i-1}, x_i]$, apply the mean value theorem to F: There exists some $t_i \in (x_{i-1}, x_i)$ so that

$$F(x_i) - F(x_{i-1}) = F'(t_i)(x_i - x_{x-1}) = f(t_i)\Delta_i.$$

Take summation over all such intervals, we obtain

$$F(b) - F(a) = \Sigma_i f(t_i)\Delta_i.$$

It follows

$$\begin{aligned} |F(b) - F(a) - \int_a^b f(x)dx| &= |\Sigma_i f(t_i)\Delta_i - \int_a^b f(x)dx| \\ &\leq |\Sigma_i f(t_i)\Delta_i - U(f;\mathcal{P})| + |U(f;\mathcal{P}) - \int_a^b f(x)dx| \\ &\leq \epsilon + \epsilon = 2\epsilon. \end{aligned}$$

Take $\epsilon \to 0$, we obtain $\int_a^b f(x)dx = F(b) - F(a)$.

□

A useful corollary for calculation is the following formula of integration by parts.

THEOREM 4.3 (Integration by parts). *Assume F, G are differentiable on $[a,b]$ with $F' = f \in R([a,b])$ and $G' = g \in R([a,b])$. Then*

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

PROOF. Consider the function FG over $[a,b]$. It is differentiable and

$$(FG)' = F'G + FG' = fG + Fg.$$

Notice that $fG + Fg$ is Riemann integrable since both $f, g \in R([a,b])$ and $F, G \in C^0([a,b])$. Apply the Fundamental Theorem of Calculus 4.2,

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

□

Introduction to Analysis Rui Wang Draft

CHAPTER 7

Sequence and series of functions

1. Sequence of functions

Suppose X is a nonempty set and (Y, d_Y) is a metric space. Consider a sequence of functions

$$f_n : X \to Y, \qquad n = 1, 2, \dots.$$

For each fixed $x \in X$, the values $(f_n(x))_{n\in\mathbb{Z}^+}$ form a sequence in Y. If this sequence is convergent in Y, then there is a point $y \in Y$, depending on x, which is its limit.

We now introduce the following definition.

DEFINITION 1.1. Let $(f_n)_{n\in\mathbb{Z}^+}$ be a sequence of functions from X to Y, where (Y, d_Y) is a metric space, and let $f : X \to Y$ be a function. We say that $(f_n)_{n\in\mathbb{Z}^+}$ **converges pointwise** to f if, for every $x \in X$, the sequence

$$f_n(x) \to f(x) \quad \text{as } n \to \infty.$$

A natural question is: do properties of the sequence $(f_n)_{n\in\mathbb{Z}^+}$, such as continuity, differentiability, or integrability, pass to the limit function f?

Let's examine some examples.

EXAMPLE 1.2. (1) Let $X = [0, 1]$ and define

$$f_n(x) = x^n, \qquad n = 1, 2, \dots.$$

Each f_n is continuous on $[0, 1]$. For any $x \in [0, 1)$ we have $x^n \to 0$ as $n \to \infty$, while $f_n(1) = 1$ for all n. Thus

$$f_n(x) \to f(x) := \begin{cases} 0, & 0 \le x < 1, \\ 1, & x = 1, \end{cases}$$

pointwise on $[0, 1]$. The limit function f is not continuous at $x = 1$. This shows that pointwise limits of continuous functions need not be continuous.

(2) Let $X = [0, 1]$ and define

$$f_n(x) = \begin{cases} n, & 0 < x \le \frac{1}{n}, \\ 0, & \frac{1}{n} < x \le 1, \\ 0, & x = 0. \end{cases}$$

Each f_n is Riemann integrable on $[0, 1]$, and

$$\int_0^1 f_n(x)\, dx = n \cdot \frac{1}{n} = 1.$$

Introduction to Analysis
Rui Wang
Draft

For every fixed $x \in [0,1]$ we have $f_n(x) \to 0$ as $n \to \infty$ (eventually $x > 1/n$), so f_n converges pointwise to the zero function $f \equiv 0$. Then

$$\int_0^1 f(x)\,dx = 0 \neq 1 = \lim_{n\to\infty} \int_0^1 f_n(x)\,dx.$$

This shows that, under mere pointwise convergence, we cannot in general interchange limit and integral.

We will now introduce a stronger notion of convergence, which is sufficient to guarantee that continuity and integrability pass to the limit.

2. Uniform Convergence

2.1. Definition of uniform convergence.

DEFINITION 2.1. Let $(f_n)_{n\in\mathbb{Z}^+}$ be a sequence of functions from X to Y, where (Y, d_Y) is a metric space, and let $f : X \to Y$ be a function. We say that $(f_n)_{n\in\mathbb{Z}^+}$ **converges uniformly** to f if, for every $\epsilon > 0$, there exists $N \in \mathbb{Z}^+$ such that for every $x \in X$ and every $n \geq N$ we have

$$d_Y\big(f_n(x), f(x)\big) < \epsilon.$$

We use the notation $f_n \overset{X}{\rightrightarrows} f$ to denote this uniform convergence on X.

Clearly, by definition, uniform convergence implies pointwise convergence. The converse is not true in general. In both examples above the convergence is pointwise but NOT uniform. (Why?)

Recall that in a complete metric space, a sequence converges if and only if it is Cauchy. When we study uniform convergence of functions, the analogous notion is that of a uniformly Cauchy sequence.

In particular, when $Y = \mathbb{R}$ (or $\mathbb{C}$), a sequence (f_n) converges uniformly on X if and only if it is uniformly Cauchy on X. We will prove this later.

DEFINITION 2.2. Let $(f_n)_{n\in\mathbb{Z}^+}$ be a sequence of functions from X to Y, where (Y, d_Y) is a metric space. We say that $(f_n)_{n\in\mathbb{Z}^+}$ is a **uniformly Cauchy sequence** on X if, for every $\epsilon > 0$, there exists $N \in \mathbb{Z}^+$ such that for all $m, n \geq N$ and all $x \in X$,

$$d_Y\big(f_m(x), f_n(x)\big) < \epsilon.$$

2.2. Uniform convergence from a metric space viewpoint. When the target space Y is $\mathbb{R}$ (or $\mathbb{C}$), one can construct an infinite-dimensional vector space with a norm, such that convergence with respect to this norm is equivalent to uniform convergence.

Let S be a nonempty set. Denote by $B(S)$ the set of all bounded functions from S to $\mathbb{R}$; that is, $f \in B(S)$ if there exists $M > 0$ such that

$$|f(x)| \leq M, \quad \text{for all } x \in S.$$

For $f \in B(S)$, define

$$\|f\|_\infty = \|f\|_{\infty,S} := \sup_{x\in S} |f(x)|.$$

This is a real number by the least upper bound property of $\mathbb{R}$.

LEMMA 2.3. *$B(S)$ is an (infinite-)dimensional vector space equipped with the norm $\|\cdot\|_\infty$. This norm $\|\cdot\|_\infty$ is called the **infinity norm** or **supremum norm**.*

Introduction to Analysis Rui Wang Draft

REMARK 2.4. In fact, $B(S)$ is an algebra over $\mathbb{R}$, and the norm satisfies

$$\|fg\|_\infty \le \|f\|_\infty \|g\|_\infty$$

for all $f, g \in B(S)$. Once we prove that $B(S)$ is complete with respect to $\|\cdot\|_\infty$, it follows that $(B(S), \|\cdot\|_\infty)$ is a Banach algebra.

Define

$$d_\infty(f,g) = \|f-g\|_\infty, \qquad \text{for all } f, g \in B(S).$$

Then d_∞ is a metric on $B(S)$; in other words, $(B(S), d_\infty)$ is a metric space. This normed space is often denoted by $\ell^\infty(S)$.

We have the following characterization of uniform convergence in terms of metric spaces.

LEMMA 2.5. *Let S be a nonempty set. A sequence of real-valued bounded functions $f_n : S \to \mathbb{R}$ converges uniformly to $f : S \to \mathbb{R}$ if and only if the sequence $(f_n)_{n\in\mathbb{Z}^+}$ converges to f in the metric space $(B(S), d_\infty)$.*

PROOF. ($\Rightarrow$) Suppose (f_n) converges uniformly to f on S. First we show f is bounded, so that $f \in B(S)$.

Take $\epsilon = 1$. By uniform convergence, there exists N such that for all $x \in S$,

$$|f_N(x) - f(x)| < 1.$$

Since f_N is bounded, there exists $M > 0$ such that $|f_N(x)| \le M$ for all $x \in S$. Then

$$|f(x)| \le |f(x) - f_N(x)| + |f_N(x)| \le 1 + M, \quad \text{for all } x \in S.$$

Hence f is bounded, so $f \in B(S)$.

Now let $\epsilon > 0$ be arbitrary. By uniform convergence, there exists N such that for all $n \ge N$ and all $x \in S$,

$$|f_n(x) - f(x)| < \epsilon.$$

Taking the supremum over $x \in S$ gives

$$\|f_n - f\|_\infty = \sup_{x\in S} |f_n(x) - f(x)| \le \epsilon,$$

so $d_\infty(f_n, f) \to 0$. Thus (f_n) converges to f in $(B(S), d_\infty)$.

($\Leftarrow$) Conversely, suppose $f_n \to f$ in $(B(S), d_\infty)$, i.e.

$$\|f_n - f\|_\infty = \sup_{x\in S} |f_n(x) - f(x)| \to 0.$$

Then for every $\epsilon > 0$ there exists N such that for all $n \ge N$,

$$\sup_{x\in S} |f_n(x) - f(x)| < \epsilon,$$

which is equivalent to

$$|f_n(x) - f(x)| < \epsilon, \quad \text{for all } x \in S.$$

This is exactly uniform convergence of (f_n) to f on S. □

Introduction to Analysis Rui Wang Draft

EXAMPLE 2.6. Consider the sequence of functions

$$f_n(x) = \frac{nx + \sin(nx^3)}{n}, \quad x \in [0,1].$$

For every fixed $x \in [0,1]$,

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{nx + \sin(nx^3)}{n} = x + \lim_{n\to\infty} \frac{\sin(nx^3)}{n} = x.$$

This shows the sequence (f_n) pointwise converges to $f(x) = x$.

Next we compute

$$\|f_n - f\|_\infty = \|\frac{\sin(nx^3)}{n}\|_\infty = \frac{1}{n}\|\sin(nx^3)\|_\infty \to 0,$$

as $n \to \infty$. Hence by Lemma 2.5, $f_n \overset{[0,1]}{\rightrightarrows} f$.

THEOREM 2.7. *The metric space $(B(S), d_\infty)$ is complete.*

PROOF. By definition of completeness, we need to show that every Cauchy sequence $(f_n)_{n\in\mathbb{Z}^+}$ in $B(S)$ converges to some $f \in B(S)$ with respect to d_∞.

Let (f_n) be a Cauchy sequence in $(B(S), d_\infty)$. This means: for every $\epsilon > 0$ there exists N such that for all $m, n \geq N$,

$$d_\infty(f_m, f_n) = \|f_m - f_n\|_\infty = \sup_{x\in S} |f_m(x) - f_n(x)| < \epsilon.$$

Step 1: Pointwise limits. Fix $x \in S$. Then for all $m, n \geq N$,

$$|f_m(x) - f_n(x)| \leq \sup_{y\in S} |f_m(y) - f_n(y)| < \epsilon,$$

so the sequence $(f_n(x))_{n\in\mathbb{Z}^+}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, there exists a limit

$$f(x) := \lim_{n\to\infty} f_n(x) \in \mathbb{R}.$$

Thus we obtain a function $f : S \to \mathbb{R}$ defined pointwise by this limit.

Step 2: Uniform convergence. We now show that $f_n \overset{S}{\rightrightarrows} f$.

For any $\epsilon > 0$, choose N such that for all $m, n \geq N$,

$$\sup_{x\in S} |f_m(x) - f_n(x)| < \epsilon.$$

Then

$$|f_m(x) - f_n(x)| < \epsilon \quad \text{for all } x \in S.$$

Fix $n \geq N$ and $x \in S$. Letting $m \to \infty$ and using $f_m(x) \to f(x)$, and the continuity of absolute value function, we get

$$|f_n(x) - f(x)| \leq \epsilon.$$

Since this holds for every $x \in S$, this proves $f_n \overset{S}{\rightrightarrows} f$.

Apply Lemma 2.5, $f \in B(S)$ and $f_n \to f$ in the metric space $(B(S), d_\infty)$.

□

REMARK 2.8. A complete normed vector space is called a **Banach space**. This theorem states that $(B(S), \|\cdot\|_\infty)$, or denoted by $\ell^\infty(S)$, is a Banach space.

Introduction to Analysis Rui Wang Draft

With minor modifications of the proof above (removing the boundedness assumption on f_n and generalize to a complete metric space), we obtain the following theorem, whose proof is left to the reader.

THEOREM 2.9. *Let (Y, d_Y) be a complete metric space, let S be a nonempty set, and let $(f_n)_{n\in\mathbb{Z}^+}$ be a sequence of functions $f_n : S \to Y$. Then (f_n) converges uniformly on S if and only if it is uniformly Cauchy on S.*

2.3. Uniform convergence of series of functions. As we have seen in Chapter 3, sequences and series carry essentially the same information: every series corresponds to the sequence of its partial sums. Hence the results we have obtained for sequences of functions can be applied to the partial sum sequences of series of functions. One of the most important applications is the following Weierstrass M-test for series of functions.

THEOREM 2.10 (Weierstrass M-test). *Let $(a_n)_{n\in\mathbb{Z}^+}$ be a sequence of real-valued functions on S, where S is a nonempty set. Assume there exists a sequence $(M_n)_{n\in\mathbb{Z}^+}$ of nonnegative real numbers such that*

$$|a_n(x)| \le M_n, \qquad \text{for all } x \in S,\ n \in \mathbb{Z}^+.$$

If the numerical series $\sum_{n=1}^{\infty} M_n$ converges, then the series

$$\sum_{n=1}^{\infty} a_n(x)$$

converges uniformly on S.

PROOF. Let $f_N(x) = \sum_{n=1}^{N} a_n(x)$ be the N-th partial sum. By Theorem 2.7, it is enough to show (f_N) is a Cauchy sequence in $(B(S), \|\cdot\|_\infty)$.

Given $\epsilon > 0$, since $\sum M_n$ converges, there exists N such that for all $m \ge N$ and $p > 0$,

$$\sum_{n=m+1}^{m+p} M_n < \epsilon.$$

Then for any $x \in S$,

$$|f_{m+p}(x) - f_m(x)| = |\sum_{n=m+1}^{m+p} a_n(x)| \le \sum_{n=m+1}^{m+p} |a_n(x)| \le \sum_{n=m+1}^{m+p} M_n < \epsilon.$$

Taking the supremum over $x \in S$ gives

$$\|f_{m+p} - f_m\|_\infty = \sup_{x\in S} |f_{m+p}(x) - f_m(x)| \le \epsilon,$$

so (f_N) is a Cauchy sequence in $(B(X), \|\cdot\|_\infty)$.

Then uniform convergence follows by Theorem 2.7. □

As an application, we have the following uniform convergence result for power series.

THEOREM 2.11. *Suppose $\sum_{n=0}^{\infty} c_n x^n$ is a power series with radius of convergence $R \in (0, \infty]$. Then for any $0 < r < R$, the series converges uniformly on the interval $[-r, r]$.*

PROOF. Fix $0 < r < R$. Choose a constant r' such that

$$r < r' < R.$$

duction to Analysis Rui Wang Draft

For any $x \in [-r, r]$ and any $n \geq 0$, we have $|x| \leq r < r'$, hence

$$|c_n x^n| \leq |c_n|(r')^n.$$

Define

$$M_n := |c_n|(r')^n.$$

By the definition of the radius of convergence R, the power series

$$\sum_{n=0}^{\infty} c_n z^n$$

converges absolutely for every z with $|z| < R$, in particular at $z = r'$. Thus

$$\sum_{n=0}^{\infty} M_n = \sum_{n=0}^{\infty} |c_n|(r')^n$$

is a convergent numerical series.

Now, for all $x \in [-r, r]$ and all n,

$$|c_n x^n| \leq M_n,$$

and $\sum M_n$ converges. By the Weierstrass M-test, the series

$$\sum_{n=0}^{\infty} c_n x^n$$

converges uniformly on $[-r, r]$. □

In general, one cannot expect uniform convergence on the whole interval of convergence $(-R, R)$. For example, the power series $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ does not converge *uniformly* on the whole $(-\infty, +\infty)$. Here is an example with finite radius of convergence R.

EXAMPLE 2.12. Consider the geometric series

$$\sum_{n=0}^{\infty} x^n.$$

Its radius of convergence is $R = 1$, and for each $x \in (-1, 1)$ the series converges to $\dfrac{1}{1-x}$.

Let

$$S_N(x) = \sum_{n=0}^{N} x^n, \qquad R_N(x) = \frac{1}{1-x} - S_N(x) = \frac{x^{N+1}}{1-x}.$$

Uniform convergence of S_N to $\frac{1}{1-x}$ on $(-1, 1)$ would mean

$$\|R_N\|_{\infty,(-1,1)} = \sup_{x \in (-1,1)} |R_N(x)| \longrightarrow 0 \quad \text{as } N \to \infty.$$

However, for each fixed N and for x close to 1 we have $|x^{N+1}| \approx 1$ while $1 - x$ is very small, so

$$|R_N(x)| = \frac{|x^{N+1}|}{1-x}$$

can be made arbitrarily large by taking $x \to 1^-$. Hence

$$\sup_{x \in (-1,1)} |R_N(x)| = +\infty$$

for every N, so in particular $\|R_N\|_{\infty,(-1,1)}$ does not tend to 0.

(One can also see trouble from the other side: as $x \to -1^+$ we have $|R_N(x)| \to \frac{1}{2}$, so on any interval of the form $(-1, r]$ the supremum is bounded below by a positive constant and again cannot go to 0.)

Introduction to Analysis
Rui Wang
Draft

Therefore the geometric series $\sum_{n=0}^{\infty} x^n$ does *not* converge *uniformly* on $(-1, 1)$; it only converges uniformly on intervals that stay a positive distance away from ± 1 (for example, on each closed interval $[a, b]$ with $-1 < a < b < 1$).

3. Uniform convergence and interchange of limits

We have seen an example of non-commuting limits in Section 2.2; let us recall it here.

EXAMPLE 3.1. Consider a function of two variables $m, n \in \mathbb{Z}^+$ defined by

$$f(m, n) = \frac{m}{m + n}.$$

Then, for each fixed n,

$$\lim_{m\to\infty} f(m, n) = 1,$$

so

$$\lim_{n\to\infty} \lim_{m\to\infty} f(m, n) = \lim_{n\to\infty} 1 = 1.$$

On the other hand, for each fixed m,

$$\lim_{n\to\infty} f(m, n) = 0,$$

so

$$\lim_{m\to\infty} \lim_{n\to\infty} f(m, n) = \lim_{m\to\infty} 0 = 0.$$

Thus the two iterated limits give different values.

We can rewrite this example in terms of a sequence of functions. For $n \in \mathbb{Z}^+$ and $x \neq 0$, define

$$f_n(x) := \frac{1/x}{1/x + n} = \frac{1}{1 + nx}.$$

Then for each fixed n,

$$\lim_{x\to 0} f_n(x) = 1,$$

so

$$\lim_{n\to\infty} \lim_{x\to 0} f_n(x) = \lim_{n\to\infty} 1 = 1.$$

On the other hand, for each fixed $x \neq 0$,

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{1}{1 + nx} = 0,$$

and hence

$$\lim_{x\to 0} \lim_{n\to\infty} f_n(x) = \lim_{x\to 0} 0 = 0.$$

This example shows that, in general, interchanging limit processes does not preserve the limiting value.

Fortunately, uniform convergence is strong enough to guarantee that certain limits can be interchanged as we will discuss later.

Introduction to Analysis Rui Wang Draft

3.1. Uniform convergence and continuity.

THEOREM 3.2. *Suppose X, Y are both metric spaces and Y is complete. Let S be a subset of X with x_0 as a limit point. Let $f_n : S \to Y$ be a sequence of functions on S that uniformly converges to $f : S \to Y$, and for each n, the limit*

$$\lim_{x\to x_0} f_n(x) =: A_n$$

exists.

Then the sequence $(A_n)_{n\in\mathbb{Z}^+}$ in Y converges and

$$\lim_{n\to\infty}\lim_{x\to x_0} f_n(x) = \lim_{n\to\infty} A_n = \lim_{x\to x_0} f(x).$$

In another word, the two limiting process: $\lim_{n\to\infty}$, $\lim_{x\to x_0}$, interchange for $f_n(x)$.

An immediate corollary is that a uniformly convergent sequence can pass the continuity to the limiting function.

COROLLARY 3.3. *Suppose X, Y are both metric spaces and Y is complete. Let $f_n : X \to Y$ be a sequence of functions on X that uniformly converges to $f : X \to Y$. If every f_n is continuous at a point $x_0 \in X$, then the limit function f is also continuous at x_0.*

PROOF. It immediately follows by taking $A_n = f_n(x_0)$. □

PROOF OF THEOREM 3.2. (1) $(A_n)_{n\in\mathbb{Z}^+}$ **is a Cauchy sequence.**

Take any $\epsilon > 0$. Since $f_n \overset{S}{\rightrightarrows} f$, by Theorem 2.9 the sequence $(f_n)_{n\in\mathbb{Z}^+}$ is uniformly Cauchy on S. So there exists $N \in \mathbb{Z}^+$ such that for all $m, n \geq N$ and all $x \in S$,

$$d_Y(f_m(x), f_n(x)) < \epsilon.$$

Now fix $m, n \geq N$, by the meaning of A_m, A_n, there exists $\delta > 0$ (depending on m, n) such that for all $x \in S$ with $0 < d_X(x, x_0) < \delta$,

$$d_Y(f_m(x), A_m) < \epsilon, \quad d_Y(f_n(x), A_n) < \epsilon.$$

Fix such an x. Then

$$\begin{aligned} d_Y(A_m - A_n) &\leq d_Y(A_m, f_m(x)) + d_Y(f_m(x), f_n(x)) + d_Y(f_n(x), A_n) \\ &< \epsilon + \epsilon + \epsilon = 3\epsilon. \end{aligned}$$

Thus for all $m, n \geq N$ we have $d_Y(A_m, A_n) \leq 3\epsilon$, so (A_n) is a Cauchy sequence in Y and hence converges by assuming Y is complete. Denote

$$A := \lim_{n\to\infty} A_n.$$

(2) $\lim_{x\to x_0} f(x) = A$**.**

Let $\epsilon > 0$ be given. Since $A_n \to A$, there exists N_1 such that

$$d_Y(A_n, A) < \epsilon \quad \text{for all } n \geq N_1.$$

Since $f_n \overset{S}{\rightrightarrows} f$, there exists N_2 such that for all $n \geq N_2$ and all $x \in S$,

$$d_Y(f_n(x), f(x)) < \epsilon.$$

Choose $N \geq \max\{N_1, N_2\}$.

Introduction to Analysis Rui Wang Draft

By assumption, $\lim_{x\to x_0} f_N(x) = A_N$, so there exists $\delta > 0$ such that for all $x \in S$ with $0 < d_X(x, x_0) < \delta$,

$$d_Y(f_N(x), A_N) < \epsilon.$$

Now fix $x \in S$ with $0 < d_X(x, x_0) < \delta$. We have

$$d_Y(f(x), f_N(x)) < \epsilon, \quad d_Y(f_N(x), A_N) < \epsilon, \quad d_Y(A_N, A) < \epsilon.$$

Therefore

$$d_Y(f(x), A) \le d_Y(f(x), f_N(x)) + d_Y(f_N(x), A_N) + d_Y(A_N, A) < 3\epsilon.$$

Since $\epsilon > 0$ is arbitrary, this shows

$$\lim_{x\to x_0} f(x) = A.$$

Combining (1) and (2) we obtain

$$\lim_{x\to x_0} f(x) = A = \lim_{n\to\infty} A_n.$$

□

If we go back to the metric space $(B(S), d_\infty)$ of bounded functions, the set $C_B^0(S)$ of both continuous and bounced functions on S, where S is a metric space is a subset of $B(S)$. What we have proved can be interpreted into the following statement.

THEOREM 3.4. *Suppose X is a metric space. Then the subset $C_B^0(X)$ is a closed subset of the metric space $(B(X), d_\infty)$. As a consequence, $C_B^0(X)$ is complete metric space itself.*

PROOF. Assume $(f_n)_{n\in\mathbb{Z}^+}$ is a convergent sequence in $C_B^0(X)$ with limit f in the metric space $(B(X), d_\infty)$. Notice the convergence in d_∞ is exactly the uniform convergence

$$f_n \overset{X}{\rightrightarrows} f.$$

Apply Corollary 3.3, $f_n \in C^0$ implies $f \in C^0$.

This shows $C_B^0(X)$ is a closed subset in $B(X)$.

The completeness follows from the general fact in a metric space: a closed subset S of a complete metric space X is complete. To see this, take a Cauchy sequence in S, then it is a Cauchy sequence in X. By the completeness of X, this sequence converges in X. The limit lives in S because S is closed. This shows the convergence of any Cauchy sequence in S.

□

3.2. Uniform convergence and integration. In this subsection we show that, for Riemann–Stieltjes integrals, uniform convergence is sufficient to interchange limits (and hence infinite series) with integration.

THEOREM 3.5. *Assume $\{f_n\}$ is a sequence of real valued functions defined on $[a, b]$ and each $f_n \in R(\alpha)([a, b])$. If $f_n \rightrightarrows f$ on $[a, b]$, then $f \in R(\alpha)([a, b])$ and*

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

Introduction to Analysis Rui Wang Draft

PROOF. For each n, set

$$\epsilon_n := \sup_{x\in[a,b]} |f_n(x) - f(x)|.$$

Then for all $x \in [a, b]$ we have

$$f_n(x) - \epsilon_n \le f(x) \le f_n(x) + \epsilon_n. \tag{3.1}$$

Since $f_n \rightrightarrows f$ on $[a, b]$, it follows that

$$\lim_{n\to\infty} \epsilon_n = 0.$$

(1) We first show that $f \in R(\alpha)([a, b])$. It is enough to prove

$$\underline{\int_a^b} f\, d\alpha = \overline{\int_a^b} f\, d\alpha.$$

From (3.1) and the monotonicity of the Riemann–Stieltjes integral, we obtain

$$\int_a^b (f_n - \epsilon_n)\, d\alpha \le \underline{\int_a^b} f\, d\alpha \le \overline{\int_a^b} f\, d\alpha \le \int_a^b (f_n + \epsilon_n)\, d\alpha. \tag{3.2}$$

Hence

$$0 \le \overline{\int_a^b} f\, d\alpha - \underline{\int_a^b} f\, d\alpha \le 2\epsilon_n(\alpha(b) - \alpha(a)).$$

Letting $n \to \infty$ and using $\epsilon_n \to 0$, we conclude that $\underline{\int_a^b} f\, d\alpha = \overline{\int_a^b} f\, d\alpha$, so $f \in R(\alpha)([a, b])$.

(2) Going back to (3.2), we now have

$$\int_a^b (f_n - \epsilon_n)\, d\alpha \le \int_a^b f\, d\alpha \le \int_a^b (f_n + \epsilon_n)\, d\alpha,$$

and therefore

$$\left| \int_a^b f\, d\alpha - \int_a^b f_n\, d\alpha \right| \le \epsilon_n(\alpha(b) - \alpha(a)).$$

Taking $n \to \infty$ and using again $\epsilon_n \to 0$, we obtain

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

□

REMARK 3.6. (*) Viewed in the normed space $(B([a, b]), \|\cdot\|_\infty)$ or metric space $(B([a, b]), d_\infty)$, Theorem 3.5 says that the set of Riemann–Stieltjes integrable functions $R(\alpha)([a, b])$ with respect to a fixed monotone α is *closed*: if $f_n \in R(\alpha)([a, b])$ and $f_n \to f$ uniformly in $B([a, b])$, then $f \in R(\alpha)([a, b])$.

Moreover, the integration map

$$\int_a^b : R(\alpha)([a, b]) \to \mathbb{R}, \quad \int_a^b (f) := \int_a^b f\, d\alpha$$

which is linear as Theorem 3.1 (1), and Theorem 3.5 shows that it is a *continuous* map with respect to the metric (since $f_n \to f$ in $\|\cdot\|_\infty$ implies $\int_a^b (f_n) \to \int_a^b (f)$).

In finite-dimensional normed spaces, every linear map is automatically continuous. Here $B([a, b])$ is infinite-dimensional, so continuity of $\int_a^b$ is a nontrivial property.

Introduction to Analysis Rui Wang Draft

COROLLARY 3.7. *Assume $a_n \in R(\alpha)([a,b])$ for all n and that*

$$f(x) := \sum_{n=0}^{\infty} a_n(x)$$

converges uniformly on $[a,b]$. Then

$$\int_a^b f\,d\alpha = \sum_{n=0}^{\infty} \int_a^b a_n\,d\alpha.$$

PROOF. Consider the sequence of partial sums

$$f_n(x) := \sum_{k=0}^{n} a_k(x), \qquad n = 0, 1, 2, \ldots.$$

Each f_n belongs to $R(\alpha)([a,b])$, and $f_n \rightrightarrows f$ on $[a,b]$. Applying Theorem 3.5 to $\{f_n\}$ yields the desired conclusion. □

EXAMPLE 3.8. Prove that $\int_0^a e^x\,dx = e^a - 1$ for $a \geq 0$, starting from the definition

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

PROOF. The power series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ has radius of convergence $R = +\infty$. By Theorem 2.11, it converges uniformly on $[0,a]$. Each term $\frac{x^n}{n!}$ is Riemann integrable on $[0,a]$, so by direct computation term by term and Corollary 3.7, we obtain

$$\int_0^a e^x\,dx = \sum_{n=0}^{\infty} \int_0^a \frac{x^n}{n!}\,dx = \sum_{n=0}^{\infty} \frac{a^{n+1}}{(n+1)!} = e^a - 1.$$

□

4. Uniform convergence and differentiation

We now turn to the differentiability of the limit function. The following example shows that differentiability of each f_n together with uniform convergence $f_n \rightrightarrows f$ is NOT enough to interchange limit and derivative.

EXAMPLE 4.1. Let

$$f_n(x) := \frac{\sin(nx)}{n}, \qquad x \in \mathbb{R}.$$

Then

$$|f_n(x)| = \left|\frac{\sin(nx)}{n}\right| \leq \frac{1}{n}$$

for all x, so $f_n \overset{\mathbb{R}}{\rightrightarrows} 0$. Each f_n is differentiable and

$$f_n'(x) = \cos(nx).$$

The limit function $f \equiv 0$ is differentiable with $f'(x) = 0$, but for any fixed $x \in \mathbb{R}$ the sequence $(\cos(nx))$ does *not* converge. Thus

$$f_n \rightrightarrows f, \quad f_n \text{ differentiable}, \quad f \text{ differentiable}$$

do not imply $\lim_{n\to\infty} f_n'(x) = f'(x)$.

Introduction to Analysis Rui Wang Draft

This shows that we need additional control on the derivatives. One natural condition is that (f_n') converges uniformly. In fact, one can assume slightly less as in the following theorem: uniform convergence of (f_n') together with convergence of (f_n) at a single point will be enough to guarantee uniform convergence of (f_n) on $[a,b]$ and allow us to pass the limit through differentiation.

THEOREM 4.2. *Assume $(f_n)_{n\in\mathbb{Z}^+}$ is a sequence of real-valued differentiable functions on $[a,b]$. If the sequence of derivative functions $(f_n')_{n\in\mathbb{Z}^+}$ converges uniformly to a function g on $[a,b]$ and $(f_n(x_0))_{n\in\mathbb{Z}^+}$ converges at some point $x_0\in[a,b]$ to a number $y_0\in\mathbb{R}$, then (f_n) converges uniformly on $[a,b]$ to some function f (with $f(x_0)=y_0$). Moreover, f is differentiable on $[a,b]$ and*

$$f'(x)=\lim_{n\to\infty} f_n'(x)=g(x) \quad \textit{for every } x\in[a,b].$$

PROOF. (1) **Uniform convergence of (f_n).**

We first show that (f_n) is a uniformly Cauchy sequence on $[a,b]$.

Since $(f_n(x_0))$ converges to y_0, it is a Cauchy sequence in $\mathbb{R}$. Thus for any $\epsilon>0$ there exists $N_1\in\mathbb{Z}^+$ such that

$$|f_n(x_0)-f_m(x_0)|<\epsilon \quad \text{for all } m,n\geq N_1.$$

Since (f_n') converges uniformly on $[a,b]$ to g, it is a uniformly Cauchy sequence of functions (or equivalently, a Cauchy sequence in $B([a,b])$). Hence there exists $N_2\in\mathbb{Z}^+$ such that for all $m,n\geq N_2$ and all $t\in[a,b]$,

$$|f_n'(t)-f_m'(t)|<\epsilon.$$

Fix $m,n\geq N:=\max\{N_1,N_2\}$ and $x\in[a,b]$. Apply the mean value theorem to f_n-f_m on the interval between x_0 and x: there exists some ξ between x_0 and x such that

$$(f_n(x)-f_m(x))-(f_n(x_0)-f_m(x_0))=(f_n'(\xi)-f_m'(\xi))(x-x_0).$$

Therefore

$$|(f_n(x)-f_m(x))-(f_n(x_0)-f_m(x_0))|=|f_n'(\xi)-f_m'(\xi)|\,|x-x_0|\leq\epsilon\,|x-x_0|.$$

Using the triangle inequality,

$$\begin{aligned}|f_n(x)-f_m(x)| &\leq |(f_n(x)-f_m(x))-(f_n(x_0)-f_m(x_0))|+|f_n(x_0)-f_m(x_0)|\\ &\leq \epsilon|x-x_0|+\epsilon\leq\epsilon((b-a)+1).\end{aligned}$$

Since this bound is independent of x, it shows (f_n) is a uniform Cauchy sequence of functions on $[a,b]$. (Equivalently,

$$\|f_n-f_m\|_\infty=\sup_{x\in[a,b]}|f_n(x)-f_m(x)|\leq\epsilon((b-a)+1).$$

Thus (f_n) is uniformly Cauchy on $[a,b]$.) By Theorem 2.7 (or equivalently, completeness of $(B([a,b]),\|\cdot\|_\infty)$ in Theorem2.7), there exists a function $f:[a,b]\to\mathbb{R}$ such that

$$f_n\overset{[a,b]}{\rightrightarrows} f.$$

In particular, $f_n(x_0)\to f(x_0)$, so $f(x_0)=y_0$.

Introduction to Analysis Rui Wang Draft

(2) **Differentiability of f and $f' = g$.**

Let g be the uniform limit of (f_n'):

$$g(x) := \lim_{n\to\infty} f_n'(x), \qquad x \in [a,b].$$

Fix $x \in [a,b]$. For $y \in [a,b] \setminus \{x\}$ define

$$\phi_n(y) := \frac{f_n(y) - f_n(x)}{y - x}, \qquad \phi(y) := \frac{f(y) - f(x)}{y - x}.$$

First, for each fixed $y \neq x$, the uniform convergence $f_n \to f$ on $[a,b]$ implies

$$f_n(y) \to f(y), \qquad f_n(x) \to f(x),$$

hence

$$\phi_n(y) = \frac{f_n(y) - f_n(x)}{y - x} \longrightarrow \frac{f(y) - f(x)}{y - x} = \phi(y).$$

So (ϕ_n) converges pointwise to ϕ on $[a,b] \setminus \{x\}$.

We will improve the pointwise convergence to uniform convergence by showing (ϕ_n) is uniformly Cauchy on $[a,b] \setminus \{x\}$.

For any $y \neq x$, by the mean value theorem applied to $f_n - f_m$ on the interval between x and y, there exists some ξ between x and y such that

$$(f_n(y) - f_m(y)) - (f_n(x) - f_m(x)) = (f_n'(\xi) - f_m'(\xi))(y - x),$$

so

$$|\phi_n(y) - \phi_m(y)| = \frac{|(f_n(y) - f_m(y)) - (f_n(x) - f_m(x))|}{|y - x|} = |f_n'(\xi) - f_m'(\xi)|.$$

Since (f_n') converges uniformly on $[a,b]$, it is uniformly Cauchy; thus, for every $\varepsilon > 0$ there exists N such that for all $m, n \geq N$ and all $t \in [a,b]$,

$$|f_n'(t) - f_m'(t)| < \varepsilon.$$

In particular, this holds for $t = \xi$, so $|\phi_n(y) - \phi_m(y)| < \varepsilon$ for all $y \neq x$. Hence (ϕ_n) is uniformly Cauchy on $[a,b] \setminus \{x\}$.

By Theorem 2.9 (with target space $\mathbb{R}$ complete), the fact that (ϕ_n) is uniformly Cauchy on $[a,b] \setminus \{x\}$ and $\phi_n(y) \to \phi(y)$ for each fixed $y \neq x$ implies

$$\phi_n \overset{[a,b]\setminus\{x\}}{\rightrightarrows} \phi.$$

For each n, differentiability of f_n at x gives

$$\lim_{y\to x} \phi_n(y) = f_n'(x),$$

and the uniform convergence $f_n' \overset{[a,b]}{\rightrightarrows} g$ implies

$$f_n'(x) \to g(x) \quad \text{as } n \to \infty.$$

Now apply Theorem 3.2 to the sequence (ϕ_n) on $S = [a,b] \setminus \{x\}$ with limit point x: we have

$$\phi_n \overset{S}{\rightrightarrows} \phi, \qquad \lim_{y\to x} \phi_n(y) = f_n'(x), \qquad f_n'(x) \to g(x),$$

so

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

duction to Analysis Rui Wang Draft

By definition of ϕ, this means

$$\lim_{y\to x}\frac{f(y)-f(x)}{y-x}=g(x),$$

so f is differentiable at x and $f'(x)=g(x)$. Since $x\in[a,b]$ was arbitrary, f is differentiable on $[a,b]$ and

$$f'(x)=g(x)=\lim_{n\to\infty}f_n'(x)\quad\text{for all } x\in[a,b].$$

□

Again, applying this theorem to series, we obtain:

COROLLARY 4.3. *Assume $(a_n)_{n\in\mathbb{Z}^+}$ is a sequence of real valued differentiable functions on $[a,b]$. If the series $\sum_{n=1}^{\infty}a_n'(x)$ converges uniformly on $[a,b]$ and the series $\sum_{n=1}^{\infty}a_n(x_0)$ converges at some point $x_0\in[a,b]$, then the series $\sum_{n=1}^{\infty}a_n(x)$ converges uniformly on $[a,b]$ to some function f. Moreover, f is differentiable on $[a,b]$ and*

$$f'(x)=\sum_{n=1}^{\infty}a_n'(x)\quad\textit{for all } x\in[a,b].$$

PROOF. Apply the previous theorem to the sequence of partial sums

$$f_N(x):=\sum_{n=1}^{N}a_n(x),\qquad N=1,2,\dots$$

and note that $f_N'(x)=\sum_{n=1}^{N}a_n'(x)$. □

Next, we study power series.

Let

$$\sum_{n=0}^{\infty}c_nx^n$$

be a power series with radius of convergence $R\in(0,\infty]$.

For following lemma follows from a direct calculation for the convergence radius R.

LEMMA 4.4. *The derivative series*

$$\sum_{n=1}^{\infty}nc_nx^{n-1}$$

has the same radius of convergence R as the original power series $\sum_{n=0}^{\infty}c_nx^n$.

Combining this with Theorem 2.11 and Corollary 4.3, we obtain:

THEOREM 4.5. *The power series $\sum_{n=0}^{\infty}c_nx^n$ with radius of convergence $R\in(0,\infty]$ is C^∞ on $(-R,R)$ and all derivatives can be obtained by term-by-term differentiation of the power series.*

EXAMPLE 4.6. We list three important power series (all with radius of convergence $R=\infty$):

(1) The exponential function:

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!},\qquad x\in\mathbb{R}.$$

Term-by-term differentiation gives

$$\frac{d}{dx}\left(\sum_{n=0}^{\infty}\frac{x^n}{n!}\right)=\sum_{n=1}^{\infty}\frac{nx^{n-1}}{n!}=\sum_{n=1}^{\infty}\frac{x^{n-1}}{(n-1)!}=\sum_{k=0}^{\infty}\frac{x^k}{k!}=e^x.$$

Introduction to Analysis Rui Wang Draft

So every derivative of e^x is again e^x.

(2) The sine function:

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots.$$

Term-by-term differentiation gives

$$\frac{d}{dx} \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{(2n+1)x^{2n}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos x.$$

(3) The cosine function:

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots.$$

Term-by-term differentiation gives

$$\frac{d}{dx} \cos x = \sum_{n=1}^{\infty} (-1)^n \frac{2n\, x^{2n-1}}{(2n)!} = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n-1}}{(2n-1)!} = -\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = -\sin x.$$

In particular, by the corollary above, each of these functions is C^∞ on $\mathbb{R}$, and all higher derivatives can be computed term-by-term from their power series.

duction to Analysis
Rui Wang
Draft