A FIRST COURSE

IN

ANALYSIS

A FIRST COURSE

IN

ANALYSIS

MAT2006 Notebook

Lecturer

Prof. Weiming Ni

The Chinese University of Hongkong, Shenzhen

Tex Written By

Mr. Jie Wang

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is taken notes from the MAT2006 in fall semester, 2018. These lecture notes were taken and compiled in LaTeX by Jie Wang, an undergraduate student in Fall 2018. Prof. Weiming Ni has not edited this document. Students taking this course may use the notes as part of their reading and reference materials. This version of the lecture notes were revised and extended for many times, but may still contain many mistakes and typos, including English grammatical and spelling errors, in the notes. It would be greatly appreciated if those students, who will use the notes as their reading or reference material, tell any mistakes and typos to Jie Wang for improving this notebook.

Notations and Conventions

$\mathbb{R}^n$	n-dimensional real space
$\mathbb{C}^n$	n-dimensional complex space
$\mathbb{R}^{m\times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m\times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $a_{ij} = a_{ji}$ for all i,j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i,j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i,j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i,j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 1

Week1

1.1. Wednesday

Recommended Reading.

1. (Springer-Lehrbuch) V. A. Zorich, J. Schüle-Analysis I-Springer (2006).
2. (The Carus mathematical monographs 13) Ralph P. Boas, Harold P. Boas, A primer of real functions-Mathematical Association of America (1996).
3. (International series in pure and applied mathematics) Walter Rudin, Principles of Mathematical Analysis-McGraw-Hill (1976).
4. Terence Tao, Analysis I,II-Hindustan Book Agency (2006)
5. (Cornerstones) Anthony W. Knapp, Basic real analysis-Birkhäuser (2005)

1.1.1. Introduction to Set

For a set $\mathcal{A} = \{1,2,3\}$, we have $2^3 = 8$ subsets of $\mathcal{A}$. We are interested to study the collection of sets.

Definition 1.1 [Collection of Subsets] Given a set $\mathcal{A}$, the the collection of subsets of $\mathcal{A}$ is denoted as $2^{\mathcal{A}}$. ■

We use Candinal to describe the order of number of elements in a set.

Definition 1.2 Given two sets $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{A}$ and $\mathcal{B}$ are said to be **equivalent** (or have the same**candinal**) if there exists a 1-1 onto mapping from $\mathcal{A}$ to $\mathcal{B}$. ■

Definition 1.3 [Countability] The set $\mathcal{A}$ is said to be **countable** if $\mathcal{A} \sim \mathbb{N} = \{1,2,3,\dots\}$; an infinite set $\mathcal{A}$ is **uncountable** if it is not equivalent to $\mathbb{N}$.

R Note that the set of integers, i.e., $\mathbb{Z} = \{\cdots,-2,-1,0,1,2,\cdots\}$ is also countable; the set of rational numbers, i.e., $\mathbb{Q} = \{p/q \mid p,q \in \mathbb{Z}, q \neq 0\}$ is countable.

We skip the process to define real numbers.

Proposition 1.1 The set of real numbers $\mathbb{R}$ is **uncountable**.

For example, $\sqrt{2} \notin \mathbb{Q}$. Some inrational numbers are the roots of some polynomials, such a number is called **algebraic** numbers. However, some inrational numbers are not, such a number is called **transcendental**. For example, π is **not** algebraic. We will show that the collection of algebraic numbers are countable in the future.

There are two steps for the proof for proposition(1.1):

Proof. 1. $2^{\mathbb{N}}$ is **uncountable**:

Assume $2^{\mathbb{N}}$ is countable, i.e.,

$$2^{\mathbb{N}} = \{A_1, A_2, \dots, A_k, \dots\}$$

Define $B := \{k \in \mathbb{N} \mid k \notin A_k\}$, it is a collection of subscripts such that the subscript k does not belong to the corresponding subsets A_k.

It follows that $B \in 2^{\mathbb{N}} \implies B = A_n$ for some n. Then it follows two cases:

- If $n \in A_n$, then $n \notin B = A_n$, which is a contradiction
- Otherwise, $n \in B = A_n$, which is also a contradiction.

The proof for the claim $2^{\mathbb{N}}$ is **uncountable** is complete.

2. $\mathbb{R} \sim 2^{\mathbb{N}}$:

Firstly we have $\mathbb{R} \sim (0,1)$. This can be shown by constructing a one-to-one mapping:

$$f : \mathbb{R} \mapsto (0,1) \quad f(x) = \frac{1}{\pi}\arctan x + \frac{1}{2}, \forall x \in \mathbb{R}$$

Secondly, we show that $2^{\mathbb{N}} \sim (0,1)$**.** We construct a mapping f such that

$$f: 2^{\mathbb{N}} \mapsto (0,1),$$

where for $\forall A \in 2^{\mathbb{N}}$,

$$f(A) = 0.a_1a_2a_3\dots, \quad a_j = \begin{cases} 2, & \text{if } j \in A \\ 4, & \text{if } j \notin A \end{cases}$$

This function is only 1-1 mapping but not onto mapping.

Reversely, we construct a 1-1 mapping from $(0,1)$ to $2^{\mathbb{N}}$. We construct a mapping g such that

$$g: (0,1) \mapsto 2^{\mathbb{N}}$$

where for any real number from $(0,1)$, we can write it into binary expansion:

$$\text{binary form:} \quad 0.a_1a_2\dots \quad \text{where } a_j = 0 \text{ or } 1.$$

Hence, we construct $g(0.a_1a_2\dots) = \{j \in \mathbb{N} \mid a_j = 0\} \subseteq \mathbb{N}$, which implies $g(\cdot) \in 2^{\mathbb{N}}$.

R Our intuition is that two 1-1 mappings in the reverse direction will lead to a 1-1 **onto** mapping. If this is true, then we complete the proof. This intuition is the **Schroder-Bernstein Theorem**.

∎

Defining Binary Form. However, during this proof, we must be careful about the binary form of a real number from $(0,1)$. Now we give a clear definition of Binary Form:

For a real number a, to construct its binary form, we define

$$a_1 = \begin{cases} 0, & \text{if } a \in (0, \frac{1}{2}) \\ 1, & \text{if } a \in [\frac{1}{2}, 1). \end{cases}$$

After having chosen $a_1, a_2, \ldots, a_{j-1}$, we define a_j to be the largest integer such that

$$\frac{1}{2}a_1 + \frac{1}{2^2}a_2 + \cdots + \frac{a_j}{2^j} \leq a$$

Then the binary form of a is $a := 0.a_1 a_2 \ldots$.

Theorem 1.1 — **Schroder-Bernstein Theorem.** If $f : A \mapsto B$ and $g : B \mapsto A$ are both 1-1 mapping, then there exists a 1-1 onto mapping from A to B, i.e., card $\#A$ equals to card $\#B$.

Exercise: Show that $(0,1)$ and $[0,1]$ have 1-1 onto mapping without applying Schroder-Bernstein Theorem.

The next lecture we will take a deeper study into the proof of Schroder-Bernstein Theorem and the real number.

1.2. Quiz

1. Show that the sequence $\{x_n\}$ is convergent, where

$$x_n = \frac{\sin 1}{2} + \frac{\sin 2}{2^2} + \cdots + \frac{\sin n}{2^n}.$$

2. Compute the following limits:

   (a)
   $$\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^{1/(1-\cos x)}$$

   (b)
   $$\lim_{n\to\infty}\int_0^1 \frac{x^n}{1+\sqrt{x}}\,\mathrm{d}x$$

3. Justify that the natural number e is irrational, where

$$e := \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

4. Every rational x can be written in the form $x = p/q$, where $q > 0$ and p and q are integers without any common divisors. When $x = 0$, we take $q = 1$. Consider the function f defined on $\mathbb{R}^1$ by

$$f(x) = \begin{cases} 0, & x \text{ is irrational} \\ \dfrac{1}{q}, & x = \frac{p}{q}. \end{cases}$$

   Find:

   (a) all continuities of $f(x)$;

   (b) all discontinuities of $f(x)$

   and prove your results.

1.3. Friday

Before we give a proof of Schroder-Bernstein theorem, we'd better review the definitions for one-to-one mapping and onto mapping.

Definition 1.4 [One-to-One/Onto Mapping] If $f : A \mapsto B$, then

- f is said to be **onto** mapping if

$$\forall b \in B, \exists a \in A \text{ s.t. } f(a) = b;$$

- f is said to be **one-to-one** mapping if

$$\forall a, b, \in A, f(a) = f(b) \implies a = b.$$

The Fig.(1.1) shows the examples of one-to-one/onto mappings.

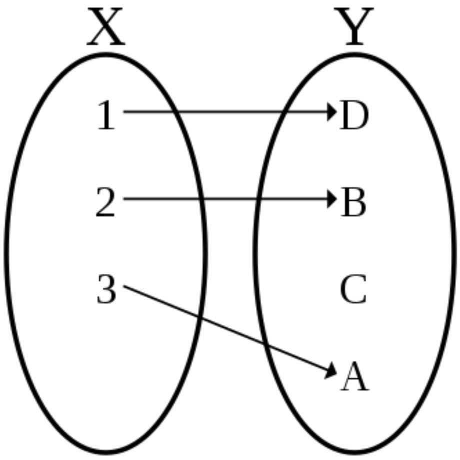
1.3.1. Proof of Schroder-Bernstein Theorem

Before the proof, note that in this lecture we abuse the notation fg to denote the composite function $f \circ g$, but in the future fg will refer to other meanings.

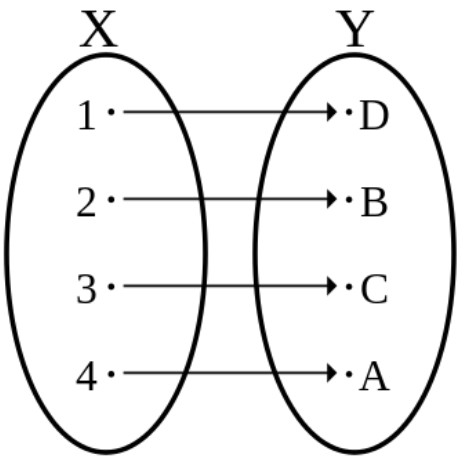
Intuition from Fig.(1.2). The proof for this theorem is constructive. Firstly Fig.(1.2) gives us the intuition of the proof for this theorem. Let $f : A \mapsto B$ and $g : B \mapsto A$ be two one-to-one mappings, and D, C are the image from A, B respectively. Note that

> if the set $B \setminus D$ is empty, then $D = B = f(A)$ with f being the one-to-one mapping, which implies f is one-to-one onto mapping. In this case the proof is complete.

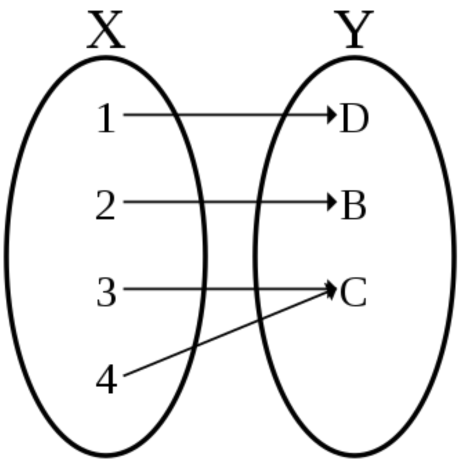
Hence it suffices to consider the case $B \setminus D$ is non-empty. Thus $B \setminus D$ is the "**trouble-maker**". To construct a one-to-one onto mapping from A, we should study the subset $g(B \setminus D)$ of A (which can also be viewed as a *trouble-maker*). Moreove, we should study



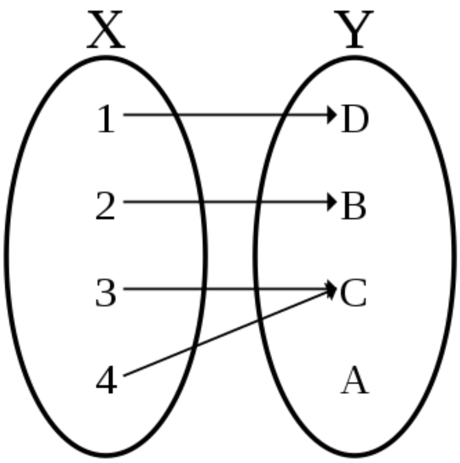
(a) A one-to-one but not onto mapping



(b) A one-to-one onto mapping



(c) A onto but not one-to-one mapping



(d) Neither a one-to-one nor onto mapping

Figure 1.1: Illustrations of one-to-one/onto mappings

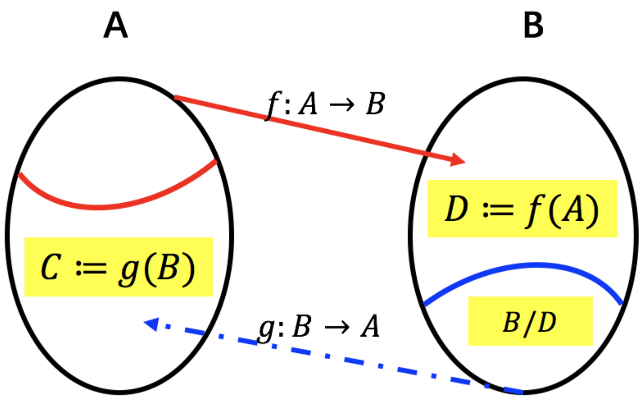


Figure 1.2: Illustration of Schroder-Bernstein Theorem

the subset $gf[g(B \setminus D)]$ (which is also a *trouble-maker*)... so on and so forth. Therefore, we should study the *union of these trouble makers*, i.e., we define

$$A_1 := g(B \setminus D), \quad A_2 := gf(A_1), \quad \cdots, \quad A_n := gf(A_{n-1}),$$

Then we study the union of infinite sets

$$S := A_1 \bigcup A_2 \bigcup \cdots \bigcup A_n \bigcup \cdots$$

Define

$$F(a) = \begin{cases} f(a), & a \in A \setminus S \\ g^{-1}(a), & a \in S \end{cases}$$

We claim that $F : A \mapsto B$ is one-to-one onto mapping.

F is onto mapping. Given any element $b \in B$, it follows two cases:

1. $g(b) \in S$. It implies $F(g(b)) = g^{-1}(g(b)) = b$.
2. $g(b) \notin S$. It implies $b \in D$, since otherwise $b \in B \setminus D \implies g(b) \in g(B \setminus D) \subseteq S$, which is a contradiction. $b \in D$ implies that $\exists a \in A$ s.t. $f(a) = b$.

   Then we study the relationship between $gf(S)$ and S. Verify by yourself that

   $$S = g(B \setminus D) \bigcup gf(S)$$

   With this relationship, we claim $a \notin S$, since otherwise $a \in S \implies gf(a) \in S$, but $gf(a) = g(b) \notin S$, which is a contradiction.

   Hence, $F(a) = f(a) = b$.

Hence, for any element $b \in B$, we can find a element from A such that the mapping for which is equal to b, i.e., F is onto mapping.

F is one-to-one mapping. Assume not, verify by yourself that the only possibility is that $\exists a_1 \in A \setminus S$ and $a_2 \in S$ such that $F(a_1) = F(a_2)$, i.e., $f(a_1) = g^{-1}(a_2)$, which follows

$$gf(a_1) = a_2 \in S = A_1 \bigcup A_2 \bigcup \cdots \tag{1.1}$$

We claim that Eq.(1.1) is false. Note that $gf(a_1) \notin A_1 := g(B \setminus D)$, since otherwise $f(a_1) \in B \setminus D$, which is a contradiction; note that $gf(a_1) \notin A_2$, since otherwise $gf(a_1) \in gfg(B \setminus D) \implies a_1 \in g(B \setminus D) = A_1 \subseteq S$, which is a contradiction.

Applying the similar trick, we wil show that $gf(a_1) \notin A_k$ for $k \geq 1$. Hence, Eq.(1.1) is false, the proof is complete.

Example 1.1 Given two sets $A := (0,1]$ and $B := [0,1)$. Now we apply the idea in the proof above to construct a one-to-one onto mapping from A to B:

- Firstly we construct two one-to-one mappings:

$$\begin{aligned} f : & A \mapsto B & \qquad g : & B \mapsto A \\ f(x) & = \frac{1}{2}x & \qquad g(x) & = x \end{aligned}$$

- It follows that $B \setminus D = (\frac{1}{2}, 1)$, $gf(B \setminus D) = (\frac{1}{4}, 1)$, so on and so forth.

$$S = (\frac{1}{2}, 1) \bigcup (\frac{1}{4}, 1) \bigcup \cdots$$

- Hence, the one-to-one onto mapping we construct is

$$F(x) = \begin{cases} \frac{1}{2}x, & x \in A \setminus S \\ x, & x \in S \end{cases}$$

- Conversely, to construct the inverse mapping, we define

$$f(x) = x \quad g(x) = \tfrac{1}{2}x$$

- It follows that $D = (0,1)$, $B \setminus D = \{1\}$. Then

$$S = \left\{ \frac{1}{2} \right\} \bigcup \cdots = \left\{ \frac{1}{2}, \frac{1}{4}, \cdots \right\}$$

- Hence, the function we construct for inverse mapping is

$$F(x) = \begin{cases} x, & x \neq \dfrac{1}{2^m} \\ 2x, & x = \dfrac{1}{2^m} \end{cases} \qquad (m = 1,2,3,\dots)$$

1.3.2. Connectedness of Real Numbers

There are two approaches to construct real numbers. Let's take $\sqrt{2}$ as an example.

1. The first way is to use **Dedekind Cut**, i.e., every non-empty subset has a least upper bound. Therefore, $\sqrt{2}$ is actually the least upper bound of a non-empty subset
$$\{x \in \mathbb{Q} \mid x^2 < 2\}.$$
2. Another way is to use **Cauchy Sequence**, i.e., every Cauchy sequence is convergent. Therefore, $\sqrt{2}$ is actually the limit of the given sequence of decimal approximations below:
$$\{1, 1.4, 1.41, 1.414, 1.4142, \dots\}$$

We will use the second approach to define real numbers. Every real number r essentially represents a collection of cauchy sequences with limit r, i.e.,

$$r \in \mathbb{R} \implies \left\{ \{x_n\}_{n=1}^{\infty} \middle| \lim_{n\to\infty} x_n = r \right\}$$

Let's give a formal definition for cauchy sequence and a formal definition for real number.

Definition 1.5 [Cauchy Sequence]

- Any sequence of rational numbers $\{x_1, x_2, \cdots\}$ is said to be a **cauchy sequence** if for every $\epsilon > 0$, $\exists N$ s.t. $|x_n - x_m| < \epsilon$, $\forall m,n \geq N$

- Two cauchy sequences $\{x_1, x_2, \dots\}$ and $\{y_1, y_2, \dots\}$ are said to be **equivalent** if for every $\epsilon > 0$, there $\exists N$ s.t. $|x_n - y_n| < \epsilon$ for $\forall n \geq N$.
- A real number is a **collection** of **equivalent** cauchy sequences. It can be represented by a cauchy sequence:

$$x \in \mathbb{R} \sim \{x_1, x_2, \dots, x_n, \dots\},$$

where x_j is a rational number.

R Let $\xi_{\mathbb{Q}}$ denote a collection of any cauchy sequences. Then once we have equivalence relation, the whole collection $\xi_{\mathbb{Q}}$ is partitioned into several disjoint subsets, i.e., equivalence classes. Hence, the real number space $\mathbb{R}$ are the equivalence classes of $\xi_{\mathbb{Q}}$.

The real numbers are well-defined, i.e., given two real numbers $x \sim \{x_1, x_2, \dots\}$ $y \sim \{y_1, y_2, \dots\}$, we can define add and multiplication operator.

$$x + y \sim \{x_1 + y_1, x_2 + y_2, \dots\}$$
$$x \cdot y \sim \{x_1 \cdot y_1, x_2 \cdot y_2, \dots\}$$

We will show how to define $x > 0$ in next lecture, this construction essentially leads to the lemma below:

Proposition 1.2 $\mathbb{Q}$ are dense in $\mathbb{R}$.

In the next lecture we will also show the completeness of $\mathbb{R}$:

Theorem 1.2 $\mathbb{R}$ is complete, i.e., every cauchy sequence of real numbers converges.

Recommended Reading:

Prof. Katrin Wehrheim, MIT Open Course, Fall 2010, Analysis I Course Notes, Online avaiable:

https://ocw.mit.edu/courses/mathematics

/18-100b-analysis-i-fall-2010/readings-notes/MIT18_100BF10_Const_of_R.pdf

Chapter 2

Week2

2.1. Wednesday

2.1.1. Review and Announcement

The quiz results will not be posted.

In this lecture we study the number theories.

The office hour is 2 - 4pm, TC606 on Wednesday

2.1.2. Irrational Number Analysis

Definition 2.1 [Algebraic Number] A number $x \in \mathbb{R}$ is said to be an **algebraic number** if it satisfies the following equation:

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{2.1}$$

where $a_n, a_{n-1}, \ldots, a_0$ are integers and not all zero. We say x is **of degree** n if $a_n \neq 0$ and x is not the root of any polynomial with lower degree.

Definition 2.2 A number $x \in \mathbb{R}$ is **transcendental** if it is not an algebraic number.

The first example is that all rational numbers are algebraic, since rational number $\frac{p}{q}$ satisfies $qx - p = 0$. Also, $\sqrt{2}$ is algebraic. We leave an exericse: show that e and π are all transcendental. In history Joseph Liouville (1844) have constructed the first transcendental number. Let's look at the insights of his construction in this lecture:

Proposition 2.1 The set of all algebraic numbers is countable.

Proof.

1. Let $\mathcal{P}_n$ denote the set of all polynomials of degree n (Here we assert polynomials have all integer coefficients by default.), i.e.,

$$\mathcal{P}_n = \{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \mid a_j \in \mathbb{Z}\}$$

   The set $\mathcal{P}_n$ have the one-to-one onto mapping to the set $\{(a_n, a_{n-1}, \ldots, a_0) \mid a_j \in \mathbb{Z}\} \subseteq \mathbb{Z}^{n+1}$, which implies $\mathcal{P}_n$ is **countable**.
2. Let $\mathcal{R}_n$ denote the set of all real roots of polynomials in $\mathcal{P}_n$. Since each polynomial of degree n has at most n real roots, the set $\mathcal{R}_n$ is a countable union of finite sets, which is at most countable. It is easy to show $\mathcal{R}_n$ is infinite, and thus countable.
3. Hence, we construct the set of all algebraic numbers $\bigcup_{n=1}^{\infty} \mathcal{R}_n$, which is countable since countably union of countable sets is also countable.

■

How fast to approximate rational numbers using rational numbers? How fast to approximate irrational numbers using rational numbers? How fast to approximate transcendental numbers using rational numbers? We need the definition for the rate of approximation first to answer these questions.

Definition 2.3 A real number ξ is **approximable** by **rational numbers to order** n if $\exists$ a constant $K = K(\xi)$ such that the inequality

$$\left|\frac{p}{q} - \xi\right| \leq \frac{K}{q^n}$$

has **infinitely many** solutions $\frac{p}{q} \in \mathbb{Q}$ with $q > 0$ and p, q are integers without any common divisors.

Intuitively, a rational number is approximable by rational numbers. Now we study its rate of approximation by applying this definition.

■ **Example 2.1** Suppose a rational number is approximable to oder α (which is a parameter). To calculate the value of α, it suffices to choose (p_k, q_k) such that

$$\left|\frac{p_k}{q_k} - \frac{p}{q}\right| \leq \frac{K}{q^\alpha}$$

Note that

$$\left|\frac{p_k}{q_k} - \frac{p}{q}\right| = \left|\frac{p_k q - p q_k}{q_k q}\right| \geq \frac{1}{q q_k} = \frac{1/q}{q_k},$$

- $\frac{p}{q}$ is approximable by rational numbers to order 1:

  If we construct $(p_k, q_k) = (kp - 1, kq)$, it follows that

  $$\left|\frac{p_k}{q_k} - \frac{p}{q}\right| = \frac{1}{kq} = \frac{1}{q_k^1}$$

- $\frac{p}{q}$ is approximable by rational numbers to order no higher than 1:

  Otherwise suppose it is approximable to order $n > 1$. The inequality holds for infinitely many (p_k, q_k):

  $$\frac{1/q}{q_k} \leq \left|\frac{p_k}{q_k} - \frac{p}{q}\right| \leq \frac{K}{q_k^n} \implies \frac{1}{q} q_k^{n-1} \leq K \tag{2.2}$$

  Since infinite (p_k, q_k) satisfy the inequality (2.2), we can choose a solution such that q_k is arbitrarily large, which falsify (2.2).

In summary, any rational number $\frac{p}{q}$ is approximable by rational numbers to order 1 and no higher than 1. ■

Liouville had shown that the transcendental number has the higher approxiamtion rate than rational and algebraic numbers, which is counter-intuitive. Let's review his process of proof:

Theorem 2.1 — Liouville, 1844. A real algebraic number ξ of degree n is not approximable by rational numbers to any order greater than n.

We can show some numbers is not algebraic, i.e., transcendental by applying this theorem:

■ **Example 2.2** [1st Constructed Transcendental Number] Given a number

$$\xi := \frac{1}{10^{1!}} + \frac{1}{10^{2!}} + \cdots,$$

we aim to show it is transcendental. Assume that it is an algebraic number of order n, then we construct the first n tails of ξ:

$$\xi_n = \frac{1}{10^{1!}} + \frac{1}{10^{2!}} + \cdots + \frac{1}{10^{n!}}$$

It follows that

$$\begin{aligned}
|\xi_n - \xi| &= \frac{1}{10^{(n+1)!}} + \frac{1}{10^{(n+2)!}} + \cdots \\
&= \frac{1}{10^{(n+1)!}} \left[1 + \frac{1}{10^{n+2}} + \frac{1}{10^{(n+2)(n+3)}} + \cdots\right] \\
&\leq \frac{1}{10^{(n+1)!}} \cdot 2 = \frac{2}{(10^{n!})^{n+1}} = \frac{2}{q^{n+1}}
\end{aligned}$$

which implies $|\xi - \frac{p}{q}| \leq \frac{K}{q^{n+1}}$ has one solution ξ_n.

We can construct infinitely many solutions from this solution:

$$\xi_{n,1} = \xi_n + \frac{1}{10^{n+2}}, \quad \xi_{n,2} = \xi_n + \frac{1}{10^{(n+2)(n+3)}}, \quad \cdots,$$

Hence, this number is approximable by rational numbers to order $n+1$, which contradicts the fact that it is an algebraic number of degree n. ■

Proof. Given an algebraic number ξ of degree n, there exists a polynomial whose roots contain ξ:

$$f(x) \equiv a_n x^n + \cdots + a_1 x + a_0 = 0.$$

We fix an interval around ξ, i.e, $I_\lambda = [\xi - \lambda, \xi + \lambda]$ (with $\lambda = \lambda(\xi) \in (0,1)$) such that I_λ contains no other root of f except ξ.

Hence, the value of f at any rational number $\frac{p}{q}$ inside I_λ is given by:

$$\left|f(\frac{p}{q})\right| = \left|a_n\frac{p^n}{q^n} + \cdots + a_1\frac{p}{q} + a_0\right| = \left|\frac{a_np^n + a_{n-1}p^{n-1}q + \cdots + a_0q^n}{q^n}\right| \neq 0$$

Hence, $\left|f(\frac{p}{q})\right| \geq \frac{1}{q^n}$, which implies that

$$\frac{1}{q^n} \leq \left|f(\frac{p}{q})\right| \tag{2.3a}$$
$$= \left|f(\frac{p}{q}) - f(\xi)\right| \tag{2.3b}$$
$$\leq |f'(\eta)||\xi - \frac{p}{q}| \tag{2.3c}$$
$$\leq M|\xi - \frac{p}{q}| \tag{2.3d}$$

with $M := \max_{\eta \in I_\lambda} f(\eta)$. Note that from (2.3b) to (2.3c) is due to mean value theorem.

Or equivalently, $|\xi - \frac{p}{q}| \geq \frac{1/M}{q^n}$ applies for any rational number $\frac{p}{q}$ inside the interval I_λ.

- Verify by yourself that ξ is not approximable by rational numbers inside the interval I_λ to any order greater than n.

- For any rational number $\frac{p}{q} \notin I_\lambda$, we have

  $$\left|\frac{p}{q} - \xi\right| \geq \lambda(\xi) \geq \frac{\lambda(\xi)}{q^n}$$

  for $q \geq 1, n \geq 1$. It is obvious that ξ is not approximable by rational numbers outside the interval I_λ to any order greater than n.

The two cases above complete the proof. ■

It's hard to determine which order the transcendental number is approximable by rational numbers. However, we can assrt that there is a "fast" approximation to transcendental numbers by applying countinued faction expansion.

Continued Fraction Expansion. Let x be irrational, then intuitively x could be represented as an infinite continued fraction as below:

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} \tag{2.4}$$

We denote the continued fraction (2.4) as $[a_0; a_1, a_2, \dots]$. Let's define the rigorous process of continued fraction expansion, i.e., how to find a_i:

- We set $a_0 = \lfloor x \rfloor$, which implies that

$$x := a_0 + \xi_0 = a_0 + \frac{1}{\frac{1}{\xi_0}} \qquad \text{for } 0 < \xi_0 < 1.$$

- We set $a_1 = \lfloor \frac{1}{\xi_0} \rfloor$, which implies that

$$x := a_0 + \frac{1}{a_1 + \xi_1} = a_0 + \frac{1}{a_1 + \frac{1}{\frac{1}{\xi_1}}}$$

- After $n+1$ steps we obtain the continued fraction of x:

$$[a_0; a_1, a_2, \dots, a_n + \xi_n]$$

  We continue this process iteratively with

$$\frac{1}{\xi_n} = a_{n+1} + \xi_{n+1}$$

  with $\xi_{n+1} \in (0,1)$.

Such a process will continue without end as x is irrational. After $n+1$ steps alternatively, we write

$$x = [a_0; a_1, \dots, a_n, a'_{n+1}] \quad \text{with } a'_{n+1} := a_{n+1} + \xi_{n+1}.$$

Observations from continued fraction expansion.

1. For $x = [a_0; a_1, \ldots] \notin \mathbb{Q}$, consider its nth convergent term

$$\tfrac{p_n}{q_n} := [a_0; a_1, \ldots, a_n], \quad n \geq 0$$

note that p_n and q_n can be computed iteratively:

$$\begin{aligned}
p_0 &= a & q_0 &= 1 \\
p_1 &= a_1 a_0 + 1 & q_1 &= a_1 \\
&\vdots & &\vdots \\
p_n &= a_n p_{n-1} + p_{n-2} & q_n &= a_n q_{n-1} + q_{n-2}
\end{aligned}$$

Note that (p_n, q_n) have no common divisors. (exercise)

Corollary 2.1 $q_n \geq n$ for $\forall n$.

Proof. Note that $q_{n-1} \leq q_n$ for $\forall n \geq 1$; and that $q_{n-1} < q_n$ for $\forall n > 1$. ■

2. From the first observation, $x := [a_0; a_1, \ldots, a'_{n+1}]$ can be written as

$$x = \frac{p_{n+1}}{q_{n+1}} = \frac{a'_{n+1} p_n + p_{n-1}}{a'_{n+1} q_n + q_{n-1}}$$

Corollary 2.2 If $\frac{p_n}{q_n} (n \geq 0)$ is the nth convergent term of x, then

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}}$$

Proof. First note that for $k \geq 2$,

$$\begin{aligned}
p_{k-1} q_k - p_k q_{k-1} &= p_{k-1}(a_k q_{k-1} + q_{k-2}) - (a_k p_{k-1} + p_{k-2}) q_{k-1} \\
&= -(p_{k-2} q_{k-1} - p_{k-1} q_{k-2})
\end{aligned}$$

After computation,

$$
\begin{aligned}
\left|x - \frac{p_n}{q_n}\right| &= \left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right| \\
&= \left|\frac{a'_{n+1}p_n + p_{n-1}}{a'_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n}\right| = \left|\frac{p_{n-1}q_n - p_n q_{n-1}}{q_n(a'_{n+1}q_n + q_{n-1})}\right| \\
&= \left|\frac{(-1)^n p_1 q_0 - p_0 q_1}{q_n(a'_{n+1}q_n + q_{n-1})}\right| = \frac{1}{q_n(a'_{n+1}q_n + q_{n-1})} \\
&< \frac{1}{q_n q_{n+1}}
\end{aligned}
$$

■

Corollary 2.3 Furthermore, for the convergent term $\frac{p_n}{q_n}(n \geq 0)$ of x, we have

$$
\left|x - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}
$$

3. The sequence $\{[a_0, a_1, \ldots, a_n]\}$ is a Cauchy sequence. (Exercise)

btw, π is approximable by rational number of order 42.

2.2. Friday

2.2.1. Set Analysis

This lecture will discuss different kinds of sets. Now recall our common sense:

Definition 2.4 [Interval]

- Open interval:
$$(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$$
- Closed interval:
$$[a,b] = \{x \in \mathbb{R} \mid a \le x \le b\}$$
- Half open intervals:
$$[a,b) = \{x \in \mathbb{R} \mid a \le x < b\}$$
$$(a,b] = \{x \in \mathbb{R} \mid a < x \le b\}$$

■

Definition 2.5 [Open sets] A set $\boldsymbol{A}$ is open if $\forall x \in \boldsymbol{A}$, there exists $(a,b) \subseteq \boldsymbol{A}$ such that $x \in (a,b)$. ■

Theorem 2.2
1. An open set in $\mathbb{R}$ is a **disjoint** union of finitely many or countably many open intervals.
2. The union of any collection of open sets is open.
3. The intersection of **finitely** many open sets is open.

The proof is omitted, check Rudin's book for reference.

(R) Note that the intersection of **countably** many open sets may not be nessarily open.
$$\bigcup_{n=1}^{\infty}\left(-\frac{1}{n}, 1+\frac{1}{n}\right) = [0,1]$$

Definition 2.6 [Neighborhood] A **neighborhood** N of a point $a \in \mathbb{R}$ is an open interval containing a. ■

Definition 2.7 [Limit Point] x is a **limit point** of the set $\boldsymbol{A}$ if for any neighborhood N of x, N contains a point $a \in A$ such that $a \neq x$. ■

Definition 2.8 [Closed Set] A set $\boldsymbol{A}$ is **closed** if $\boldsymbol{A}$ contains all of its limit points. ■

Proposition 2.2 $\boldsymbol{A}$ is **closed** of and only if $\mathbb{R} \setminus \boldsymbol{A}$ is open.

2.2.2. Set Analysis Meets Sequence

Definition 2.9 [Limit Point of sequence] Given a sequence $\{a_n\}$, i.e.,

$$a_1, a_2, a_3, \dots,$$

a point x is said to be the **limit point** of $\{a_n\}$ if there exists a subsequence $\{x_{n_1}, x_{n_2}, \dots\}$ converging to x. ■

Does there exist a sequence of rational numbers such that every irrational number is a limit point? Yes, and we use an example as illustration.

■ **Example 2.3** $\{q_1, q_2, \dots\}$ is a sequence of all rational numbers. For example, to construct a subsequence with limit $\sqrt{2}$, we pick:

$$\begin{aligned}
&q_{m_1} \in (\sqrt{2}-1, \sqrt{2}+1) \setminus (\sqrt{2}-\frac{1}{2}, \sqrt{2}+\frac{1}{2}) \\
&q_{m_2} \in (\sqrt{2}-\frac{1}{2}, \sqrt{2}+\frac{1}{2}) \setminus (\sqrt{2}-\frac{1}{3}, \sqrt{2}+\frac{1}{3}) \\
&\dots \\
&q_{m_k} \in (\sqrt{2}-\frac{1}{k}, \sqrt{2}+\frac{1}{k}) \setminus (\sqrt{2}-\frac{1}{k+1}, \sqrt{2}+\frac{1}{k+1})
\end{aligned}$$

The same argument works for all irrational numbers, also for all rational numbers. ■

2.2.3. Completeness of Real Numbers

Now we use Cauchy sequence to construct the completeness of real numbers. First let's give a proof of three important theorems. Note that the proof and applications of these theorems are mandatory.

Theorem 2.3 — **Bolzano-Weierstrass.** Every bounded sequence has a convergent subsequence.

Theorem 2.4 — **Cantor's Nested Interval Lemma.** A sequence of nested closed bounded intervals $I_1 \supseteq I_2 \supseteq \cdots$ has a non-empty intersection, i.e., $\bigcap_{k=1}^{\infty} I_k \neq \emptyset$.

Theorem 2.5 — **Heine-Borel.** Any open cover $\{\mathcal{U}\}$ of a bounded closed set $\boldsymbol{E}$ consists of a finite sub-cover, i.e, $\boldsymbol{E} \subseteq$ the union of $\{\mathcal{U}\}$.

Proof for Bolzano-Weierstrass Theorem.

- Suppose $\{a_1, a_2, \dots\}$ is a bounded sequence, w.l.o.g., $\{a_1, a_2, \dots\} \subseteq [-M, M]$. We pick $a_{n_1} = a_1$.
- w.l.o.g., assume that $[0, M] \bigcap \{a_1, a_2, \dots\}$ is infinite (otherwise $[-M, 0] \bigcap \{a_1, a_2, \dots\}$ is infinite), then we pick $a_{n_2} \neq a_{n_1}$ such that $a_{n_2} \in [0, M]$.
- w.l.o.g., assume that $[0, \frac{M}{2}] \bigcap \{a_1, a_2, \dots\}$ is infinite, then we pick $a_{n_3} \neq a_{n_1}, a_{n_2}$ such that $a_{n_3} \in [0, \frac{M}{2}]$.

In this case, $\{a_{n_1}, a_{n_2}, \dots\}$ is Cauchy (by showing $|a_{n_k} - a_{n_l}| < \epsilon$ for large k, l), hence converges. ■

Proof for Cantor's Nested Interval Lemma.

1. Pick $a_k \in I_k$ for $k = 1, 2, \dots$, thus the sequence $\{a_1, \dots, a_k, \dots\}$ is bounded. By Theorem (2.3), there exists a convergent sub-sequence $\{a_{k_l}\}$ (with limit a). It suffices to show $a \in \bigcup_{m=1}^{\infty} I_k$.

2. For fiexed m, there exists index j such that $a_{k_l} \in I_m$ for all $l \geq j$. Since I_m is closed, it must contain a_{k_l}'s limit point, i.e., $a \in I_m$.
3. Our choice is arbitrary m and hence a belongs to the intersection of all nested closed intervals. The proof is complete.

■

Before the proof of third theorem, let's have a review for open cover definitions:

> **Definition 2.10** [Open Cover] Let $\boldsymbol{E}$ be a subset of a metric space X. An open cover $\{\mathcal{U}_\alpha\}_{\alpha \in A}$ of $\boldsymbol{E}$ is a collection of open sets in X whose union contains $\boldsymbol{E}$, i.e., $\boldsymbol{E} \subseteq \bigcup_{\alpha \in A} \mathcal{U}_\alpha$. A finite **subcover** of $\{\mathcal{U}_\alpha\}_{\alpha \in A}$ is a **finite** sub-collection of $\{\mathcal{U}_\alpha\}_{\alpha \in A}$ whose union still contains $\boldsymbol{E}$.

For example, consider $\boldsymbol{E} := [\frac{1}{2}, 1)$ in metric space $\mathbb{R}$. Then the collection

$$\{I_n\}_{n=3}^{\infty}, \qquad \text{where } I_n := (\frac{1}{n}, 1 - \frac{1}{n})$$

is a open cover of $\boldsymbol{E}$. Note that the finite subcover may not necessarily exist. In this example, the finite subcover of $\{I_n\}_{n=3}^{\infty}$ does not exist.

Proof for Heine-Borel Theorem.

Suppose $\boldsymbol{E} := [0, M]$ is a bounded closed interval with an open cover $\{\mathcal{U}\}$. The trick of this proof is to construct a sequence of nested closed bounded intervals.

- **Base case** We choose $I_1 = \boldsymbol{E} = [0, M]$
- **Inductive step** For example, Assume that $\boldsymbol{E}$ cannot be covered by finitely many open sets from $\{\mathcal{U}\}$, then at least one sub-interval $[0, \frac{M}{2}]$ or $[\frac{M}{2}, M]$ cannot be covered. Let I_2 be one of these sub-intervals that cannot be covered by finitely many elements of $\{\mathcal{U}\}$.

Repeating this process, we attain a nested bouned closed intervals $I_1 \supseteq I_2 \supseteq \cdots \supseteq$, which implies $\bigcap_{k=1}^{\infty} I_k \neq \emptyset$ (suppose $a \in \bigcap_{k=1}^{\infty}$), and $|I_k| = \frac{M}{2^k} \to 0$.

Note that $a \in \boldsymbol{E}$ implies that there exists an open set ξ in $\{\mathcal{U}\}$ such that $a \in \xi$. Thus $(a-\epsilon, a+\epsilon) \in \xi$ for small ϵ. Note that there exists sufficiently large k such that $\frac{M}{2^k} < 2\epsilon$, and $a \in I_k$, which implies $I_k \subseteq \xi$, which is a contradiction. ■

These theorems have simple applications:

Proposition 2.3 Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$ with the series convergent for $|x| < 1$. If for $\forall x \in [0,1)$, there exists $n := n(x)$ such that $\sum_{k=n}^{\infty} a_k x^k = 0$, then f is a polynomial (that is independent from x, i.e., n does not depend on x.)

In next lecture we will continue to study the completeness of real numbers and will speed up.

Chapter 3

Week3

3.1. Tuesday

3.1.1. Application of Heine-Borel Theorem

Theorem 3.1 Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$ which converges in $|x| < 1$. If for every $x \in [0,1)$, there exists $n(= n(x))$ such that $\sum_{n+1}^{\infty} a_k x^k = 0$, then f is a polynomial, i.e., n does not depend on x.

The idea is to construct a sequence of points $\{x_n\}$ satisfying $f(x_k) = a_0 + \cdots + a_m x_k^m$, i.e., infinite points coincide $f(x)$ with a polynomial, which implies f is a polynomial.

Proof. Construct $E_N := \{x \in [0, \frac{1}{2}] \mid \sum_{k=N+1}^{\infty} a_k x^k = 0\}$. It follows that

$$[0, \frac{1}{2}] = \bigcup_{N=1}^{\infty} E_N,$$

which implies that at least one E_N is uncountable, say, E_m is uncountable. In particular, E_m is infinite

By Bolzano-Weierstrass Theorem, there exists a sequence $\{x_k\} \subset E_m$ with limit x_0 in E_m as E_m is closed. Hence, $f(x) = a_0 + a_1 x + \cdots + a_m x^m$ holds for the sequence $\{x_m\}$. Intuitively we conclude the power series and the analytics function coincide each other for every point $x \in (-1,1)$.

$$f(x) \equiv a_0 + a_1 x + \cdots + a_m x^m$$

■

However, the proof above does not show why a sequence coincide $f(x)$ with a polynomial could imply f is a polynomial for every point. We summarize this induction as the proposition(3.1) and give a proof below. Before that we formulate what we want to prove precisely:

> Let f be analytic, i.e., $f(x) = a_0 + a_1 x + \cdots + a_n x^n + \cdots$ on $(-1,1)$; and $f(x_k) = \sum_{i=1}^{m} a_i x_k^i$ for all $k \geq 1$, where $\{x_k\}$ is a sequence with limit x_0. Then $f(x) = \sum_{i=1}^{m} a_i x^i$ on $(-1,1)$.

To show this statement, we construct

$$g(x) = f(x) - \sum_{i=1}^{m} a_i x^i \implies g(x_k) = 0, \forall k \geq 1$$

It suffices to show $g \equiv 0$ on $(-1,1)$. Moreover, if we construct $y_k := x_k - x_0$, and set $f(x) = a_0 + a_1(x - x_0) + \cdots$, then it suffices to prove the proposition given below:

Proposition 3.1 Let g be analytic, i.e., $g(x) = b_0 + b_1 x + \cdots + b_n x^n + \cdots$ on $(-1,1)$; and $g(x_k) = 0$ for all $k \geq 1$, where $\{x_k\} \to 0$. Then $g \equiv 0$ on $(-1,1)$ (i.e., $b_0 = b_1 = \cdots = 0$)

Proof.

- Note that $g(0) = 0$ due to continuity property. Also, $g(0) = b_0 = 0$, which follows that

$$g(x) = x(b_1 + b_2 x + \cdots + b_n x^{n-1} + \cdots \tag{3.1}$$

- Substituting x with x_k in Eq.(3.1), we derive

$$0 = g(x_k) = x_k(b_1 + b_2 x_k + \cdots + b_n x_k^{n-1} + \cdots \tag{3.2}$$

  Taking limit both sides for (3.2), we derive $b_1 = 0$.

- By applying the same trick, we conclude $b_0 = b_1 = \cdots = 0$ (the rigorous proof requires induction).

■

Now we talk about some advanced topics in Analysis.

3.1.2. Set Structure Analysis

Definition 3.1 [Nowhere Dense] A set $\boldsymbol{B}$ is said to be **nowhere dense** if its closure $\overline{B}$ contains no non-empty open set. ▪

For example,

$$B = \{1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, ..\} \implies \overline{B} = B \bigcup \{0\},$$

which contains no non-empty open set.

Definition 3.2 [1st category] A set of $\boldsymbol{B}$ is said to be of 1st category if it can be written as the **union** of **finitely** many or **countably** many **nowhere** dense sets. ▪

Definition 3.3 [2rd category] A set is said to be of 2rd category if it is **not** of 1st category ▪

Theorem 3.2 — **Baire-Category Theorem.**
- $\mathbb{R}$ is of 2rd category, i.e.,
- $\mathbb{R}$ cannot be written as the union of countably many nowhere dense sets, i.e.,
- if $\mathbb{R} = \bigcup_{n=1}^{\infty} A_n$, then at least one A_n whose closure contains a non-empty open set.

Proof.
- Assume $\mathbb{R} = \bigcup_{n=1}^{\infty} A_n$ such that all A_n's are nowhere dense. It follows that

$$\mathbb{R} \setminus \overline{A_1} \text{ is open,}$$

since $\overline{A_1}$ is closed and its complement is open.

- We construct an open set N_1 such that $\overline{N_1} \subseteq \mathbb{R} \setminus \overline{A_1}$. (e.g., there exists ε and $x \in \mathbb{R} \setminus \overline{A_1}$ such that $N_1 := B(x, \varepsilon) \subseteq \overline{N_1} \subseteq \mathbb{R} \setminus \overline{A_1}$.)

- Since A_2 is nowhere dense, we imply $\overline{A_2}$ does not contain N_1, i.e., $N_1 \setminus \overline{A_2}$ is open.

- By applying similar trick, we obtain a sequence of nested sets

$$\overline{N_1} \supseteq N_1 \supset \overline{N_2} \supset N_2 \cdots$$

  The cantor's theorem implies that $\bigcap_{k=1}^{\infty} \overline{N_k} \neq \emptyset$.

- On the other hand, $\bigcap_{k=1}^{\infty} \overline{N_k} \subseteq \mathbb{R} \setminus \bigcup_{n=1}^{m} A_n$ for any finite m.

- Therefore, $\emptyset \neq \bigcup_{k=1}^{\infty} \overline{N_k} \subseteq \mathbb{R} \setminus \bigcup_{n=1}^{\infty} A_n = \emptyset$, which is a contradiction.

■

R $\mathbb{R}$ is of 2nd category, i.e., if $\mathbb{R} = \bigcup_{n=1}^{\infty} A_m$, then at least A_n whose closure contains a **non-empty** open sets; The theorem also holds if we replace $\mathbb{R}$ by a **complete** metric space (essentially the same proof).

Most proof for $\mathbb{R}$ can be generalized into metric space, the proof for which is essentially the same. Now let's introduce the metric space informally.

Metric Space. A metric space is an ordered pair (M, d), where M is a set and d is a metric on M, i.e., d is a distance function defined for two points on M. Here we list several examples:

The Real Line. For $\mathbb{R}$, $d(x, y) = |x - y|$. Note that $(\mathbb{Q}, d)$ and $(\mathbb{R} \setminus \mathbb{Q}, d)$ are also metric spaces, but not complete.

n-Cell Real Space. $\mathbb{R}^n$, with $d(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|$ is a metric space.

Bounded Sequences. The set of all bounded sequences on $\mathbb{R}$ is a metric space, with d defined as:

$$d(\{x_n\}, \{y_n\}) = \sup\{|x_i - y_i| \mid i = 1, 2, \ldots\}$$

Bounded Functions. Similarly, the set of all bounded continuous functions on $\mathbb{R}$ (different domains), with

$$d_1(f, g) = \sup\{|f(x) - g(x)| \mid x \in \mathbb{R}\},$$

or

$$d_2(f,g) = (\int_0^1 |f(x) - g(x)|^2 \, dx)^{1.2}$$

is a metric space. Note that $(\mathcal{C}[0,1], d_1)$ is complete, and $(\mathcal{C}[0,1], d_2)$ is not complete. (exercise)

R

Different distance definition corresponds to different metric spaces.

Recall that a metric space is complete if all Cauchy sequence of which converge.

3.1.3. Reviewing

Definition 3.4 [Sequence] A sequence is defined as a kind of function $f : \mathbb{N} \to \mathbb{R}$, denoted as $\{f(0), f(1), \dots\}$. Conventionally we denote it as $x_1, x_2, \dots$ ■

Definition 3.5 [Limit] A number α is the limit of $\{x_1, x_2, \dots\}$ if $\forall \epsilon > 0$, there $\exists N = N(\epsilon)$ such that $|x_k - \alpha| < \epsilon$ for $\forall k \geq N$, denoted by $\alpha_n \to \alpha$ ■

Definition 3.6 [liminf & limsup]

$$\liminf_{k\to\infty} x_k := \lim_{n\to\infty} \inf_{k\geq n} x_k$$

which is the smallest limit point of the sequence

$$\limsup_{k\to\infty} x_k := \lim_{n\to\infty} \sup_{k\geq n} x_k$$

which is the largest limit point of the sequence. ■

A sequence always has liminf and limsup.

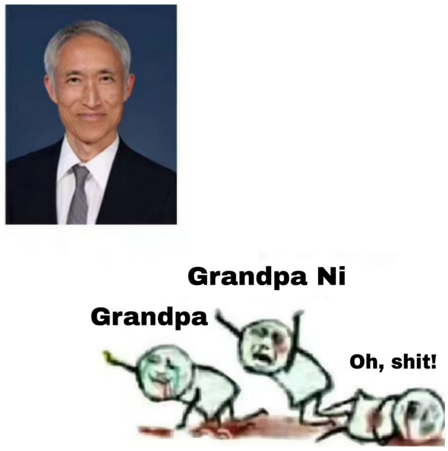
Definition 3.7 [Partial Sum] Given the sequence $\{a_n\}$, its n-th partial sum are defined as:

$$s_n = a_1 + \cdots + a_n,$$

the series $\sum_i a_i$ is defined as the limit of the partial sum, ■

Next lecture we will show that most continuous function is nowhere differentiable, by applying the Baire Category Theorem on $(\mathcal{C}[0,1], d_1)$

3.2. Friday



"Our first quiz is at 1:30-2:20pm on September 30th. That is next Sunday. There will be around 5 questions." the Grandpa said breezily,

3.2.1. Review

This lecture will review the continuty of function. Let's start with some easy examples:

$$f(x) = \begin{cases} x, & x \in \mathbb{Q} \\ -x, & x \notin \mathbb{Q} \end{cases}$$

This function is continuous nowhere expect for $x = 0$.

Definition 3.8 [Continuous] Given a function $f : D \mapsto \mathbb{R}$,

- we say f is **continuous** at $x_0 \in D$ if for $\forall \varepsilon > 0$, $\exists \delta := \delta(\varepsilon, x_0) > 0$ s.t. $|f(y) - f(x_0)| < \varepsilon$ for $\forall |y - x_0| < \delta$

- f is continuous on D if it is continuous at every point in D.
- If $\delta := \delta(\varepsilon)$, i.e., δ is independent of $x_0 \in D$, then f is said to be **uniformly continuous** on D.

R The following statements are equivalent, you should show by yourself.

1. f is continuous
2. If $\{x_n\} \to x_0$ as $n \to \infty$, then $\{f(x_n)\} \to f(x_0)$ as $n \to \infty$.
3. $f^{-1}(\boldsymbol{A})$ is open/closed if the set $\boldsymbol{A}$ is open/closed.

Definition 3.9 [Compact] A set K is **compact** (cpt) if for evey open cover of K, there exists a **finite** sub-cover.

R The compactness has an important connection with continuity, e.g., the continuous function f maps compact sets to compact sets.

There is a useful way to determine whether a point is continuous at f, which will be discussed in this lecture.

3.2.2. Continuity Analysis

Let's raise some examples first. From these examples we can see that the proof of continuousness is non-trival.

Example 3.1 1. Given a funciton

$$f(x) = \begin{cases} \sin\frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

From the graph we can see that f oscillates heavily near zero point.

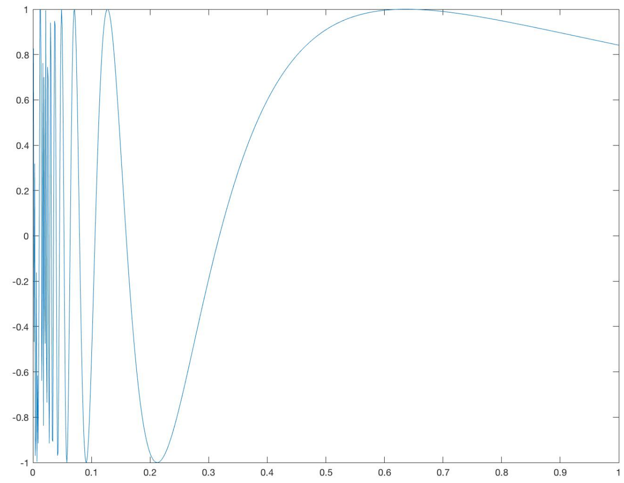


Figure 3.1: Graph for f

It is easy to show that $\sup_{x,y\in N_\delta(0)} |f(x) - f(y)| = 2$ however small δ is.

2. For another function

$$g(x) = \begin{cases} x\sin\frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Conversely, it oscillates weakly near the zero point.

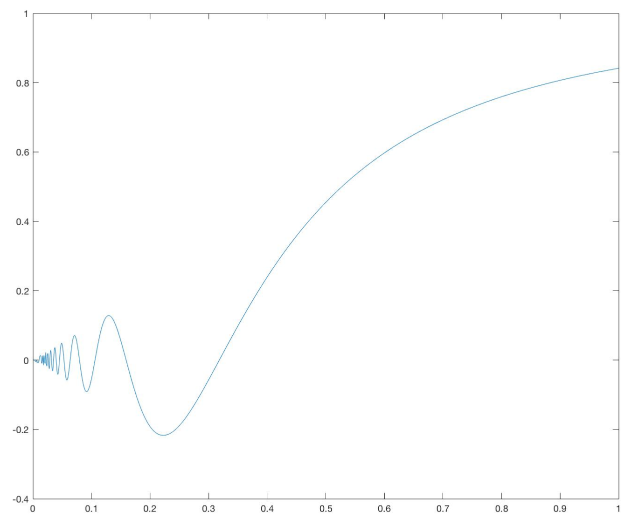


Figure 3.2: Graph for g

It is easy to show that $\sup_{x,y \in N_\delta(0)} |g(x) - g(y)| = 0$ as $\delta \to 0$.

Definition 3.10 [oscillation]

- The oscillation of a function f on E is defined as

$$\omega(f;E) := \sup_{x,y \in E} |f(x) - f(y)|$$

- The oscillation of f at a single point x_0 is defined as

$$\lim_{\delta \to 0} \omega(f, N_\delta(x_0)) := w(f;x_0)$$

R

- Here we abuse the notation to denote the oscillation at x_0 with $\omega(f;x_0)$, but note that $\omega(f;x_0) \neq \omega(f;\{x_0\})$.
- The well-definedness of $w(f;x_0)$ is because $\omega(f, N_\delta(x_0))$ is non-incresing as δ decreases and has a lower bound 0.
- A function f is continuous at x_0 iff $w(f, x_0) = 0$. (verify by yourself)

An classical example that illustrates a function can have continuous points in $\mathbb{R} \setminus \mathbb{Q}$ is shown below. We have faced this example in the diagnostic quiz:

Example 3.2 The Dirichlet function is defined for $\mathbb{R} \setminus \{0\}$:

$$f(x) = \begin{cases} 0, & x \notin \mathbb{Q} \\ \frac{1}{q}, & x = \frac{p}{q}, q > 0, (p,q) = 1 \end{cases}$$

The function f is continuous at x iff $x \notin \mathbb{Q}$. The set of all discontinuous points of f is $\mathbb{Q}$.

Now the question turns out:

> *Does there exists a function g of which the set of all discontinuous points of g is $\mathbb{R} \setminus \mathbb{Q}$?*

Applying Baire-Category Theorem, we will show the answer to this question is no.

Proposition 3.2 Suppose f is continuous on a dense set in $\mathbb{R}$. Then the set of all discontinous points of f, denoted as T, must form a set of first category, i.e., (a countably union of nowhere dense sets)

R $\mathbb{R} \setminus \mathbb{Q}$ is of second category, otherwise assume

$$\mathbb{R} \setminus \mathbb{Q} = \bigcup_{i=1}^{\infty} I_i$$

for nowhere dense sets I_i, which implies $\mathbb{R} = [\cup_{i=1}^{\infty} I_i] \bigcup [\cup_{q \in \mathbb{Q}} \{q\}]$ is a countably union of nowhere dense sets. Thus by applying proposition(3.2), the irrational number space cannot be the set of discontinuities.

The idea of the proof is to express T as countably union of sets, and argue that at least one of which must be nowhere dense.

Proof. We construct $D_n = \{x \in \mathbb{R} \mid w(f;x) \geq \frac{1}{n}\}$, which follows that

$$T = \bigcup_{n=1}^{\infty} D_n.$$

It suffices to show that D_n is nowhere dense for every n by contradiction.

Assume for some fixed n, D_n is not nowhere dense, i.e., $\overline{D_n}$ contains an **open** interval $\boldsymbol{I}$. Note that the set of continuous points is dense, we conclude that there exists a point a inside the interval $\boldsymbol{I}$ such that f is continuous at a. (why?) Also, there exists a sequence $\{b_k\} \subseteq D_n$ with limit a. (since you can verify D_n is closed)

Since f is continuous at a, there exists $\delta > 0$ such that

$$|f(x) - f(a)| < \frac{1}{4n} \text{for } |x - a| < \delta. \tag{3.3a}$$

At the same time $\{b_k\} \subseteq (a-\delta, a+\delta)$ for all large k, i.e., $\omega(f; b_k) \geq \frac{1}{n}$. Hence, there exists a sequence $\{c_{kl}\}$with limit b_k, such that the difference $|f(c_{kl}) - f(b_k)|$ is at least greater than $\frac{1}{2n}$ (why not $\frac{1}{n}$?), i.e.,

$$|f(c_{kl} - f(b_k)| \geq \frac{1}{2n} \tag{3.3b}$$

Meanwhile, for large l, note that c_{kl} is close to a, i.e., from (3.3a) we have

$$|f(c_{kl}) - f(a)| < \frac{1}{4n} \tag{3.3c}$$

Also, note that b_k is close to a for large k, i.e.,from (3.3a) we have

$$|f(b_k) - f(a)| < \frac{1}{4n} \tag{3.3d}$$

Three inequalities (3.3b) to (3.3d) show a contradiction:

$$|f(c_{kl}) - f(b_k)| \leq |f(c_{kl}) - f(a)| + |f(b_k) - f(a)| < \frac{1}{4n} + \frac{1}{4n} = \frac{1}{2n}$$

■

Theorem 3.3 Let f be the **pointwise** limit of a sequence of continuous functions $\{f_n\}$, i.e., $f(x) = \lim_{n\to\infty} f_n(x)$. Then the set if all discontinuous points of f must be a set of first category.

R Review the uniform limit version of this theorem;

Proof. We claim $D_\varepsilon = \{x \in \mathbb{R} \mid \omega(f,x) \geq \varepsilon\}$ is nowhere dense for any $\varepsilon > 0$. (fixed ε).

Assume $\overline{D_\varepsilon}$ contains an open set, or equivalently, D_ε contains an open set $\mathcal{U}$(since D_ε is closed, i.e., $\overline{D_\varepsilon} = D_\varepsilon$). Define

$$A_{mn} = \left\{x \in U \mid |f_m(x) - f_n(x)| \leq \frac{\varepsilon}{4}\right\}.$$

Proposition 3.3 The set A_{mn} is closed.

The proof of proposition is moved in the end.

Set $A_m = \bigcap_{n\geq m} A_{mn}$, which is also closed, and

$$A_m \subseteq \left\{ x \in U \mid |f_m(x) - f(x)| \leq \frac{\varepsilon}{4} \right\}$$

For every $x \in U$, as $\lim_{n\to\infty} f_n(x) = f(x)$, we have $x \in \bigcup_{m=1}^{\infty} A_m$, which implies $U \subseteq \bigcup_{m=11}^{\infty} A_m$. Applying the Baire Category Theorem, there exists one A_m containing an open set W.

For $x_0 \in W \subseteq \mathcal{U} \subseteq D_\varepsilon$, pick $\{x_n\} \subseteq W$ with limit x_0 such that

$$|f(x_n) - f(x_0)| \geq \frac{3}{4}\varepsilon \tag{3.4a}$$

At the same time, since $x_n, x_0 \in W$, it follows that

$$|f_m(x_n) - f(x_n)| \leq \frac{\varepsilon}{4} \tag{3.4b}$$

$$|f_m(x_0) - f(x_0)| \leq \frac{\varepsilon}{4} \tag{3.4c}$$

From (3.4a) to (3.4c), we conclude that

$$\begin{aligned} \frac{3}{4}\varepsilon &\leq |f(x_n) - f(x_0)| \\ &\leq |f_m(x_n) - f(x_n)| + |f_m(x_n) - f_m(x_0)| + |f_m(x_0) - f(x_0)| \\ &\leq \frac{1}{2}\varepsilon + |f_m(x_n) - f_m(x_0)| \end{aligned}$$

Or equivalently,

$$|f_m(x_n) - f_m(x_0)| \geq \frac{\varepsilon}{4}, \forall n \tag{3.4d}$$

which implies $\omega(f_m; x_0) \geq \frac{\varepsilon}{4}$, which implies f_m is discontinuous at x_0, which is a contradiction. ■

R

- The proof for A_{mn} is closed is easy: pick any sequence $\{x_k\} \subseteq A_{mn}$ with limit x, it suffices to show $x \in A_{mn}$, i.e.,

$$\lim_{k \to \infty} |f_m(x_k) - f_m(x_k)| \le \frac{\varepsilon}{4}$$

- The applications of Baire Category Theorem give us an estimation of how large and how small a set is. In next lecture we will see how large is the set of continuous but nowhere differential functions.

Chapter 4

Week4

4.1. Wednesday

This lecture will talk about the applications of Baire-Category Theorem and continuity analysis.

4.1.1. Function Analysis

In last lecture we have studied that given a analytic function f, if f can be expressed as a partial sum of its series for each x, then f is a polynomial:

Proposition 4.1 Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$ in $|x| < 1$. If for every $x \in (-1,1)$, there exists $n(= n(x))$ such that $\sum_{k=n+1}^{\infty} a_k x^k = 0$, then f is a polynomial, i.e., n is independent of x.

The idea of the proof is to construct a sequence of points such that f coincide with a polynomial over these points, which implies f is indeed a polynomial.

Now we study its stronger version, i.e., f may not be analytic, it only needs to be infinitely differentiable:

Proposition 4.2 Suppose $f \in \mathcal{C}^{\infty}[-1,1]$. If for every $|x| \le 1$, there exists $n(= n(x))$ such that $f^{(n)}(x) = 0$, then f is a polynomial.

R Note that an analytic function (i.e., can be expressed as power series) is always infinitely differentiable, but the reverse direction is necessarily not.

For example, recall we have learnt a function

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & x \neq 0 \\ 0, & x = 0 \end{cases},$$

such that it is infinitely differentiablem but $f^{(n)} = 0$ for $n = 1, 2, 3, \ldots$. Hence, this function is not analytic at $x = 0$.

Proof. Construct a sequence of set

$$E_n = \{x \in [-1,1] \mid f^{(n)}(x) = 0\} \implies [-1,1] = \bigcup_{n=1}^{\infty} E_n,$$

with E_n closed (Exercise #1). Applying Baire-Category Theorem to $[-1,1]$, at least one E_{N_1} contains a non-empty open interval, say I_1 (Exercise #2).

1. On I_1, $f^{(N_1)} \equiv 0$, which implies f is a polynomial of degree $N_1 - 1$(Exercise #3).
2. If $I_1 = (-1,1)$, the proof is complete.
3. Otherwise, $[-1,1] \setminus I_1 \neq \emptyset$. Applying Baire-Category Theorem on the set $[-1,1] \setminus I_1 := \bigcup_{n=1}^{\infty} E_n \setminus I_1$, we conclude that at least one $E_{N_2} \setminus I_1$ contains a non-empty open interval, say I_2, on which f is a polymial of degree $N_2 - 1$.
4. Each time applying the same trick to construct $I_1, I_2, \ldots$, and make sure these are the **maximal** intervals with the desired properties. Finally, we reach the stage that:

   $f|_{x \in I_j}$ is a polynomial of order $N_j - 1$ for $j = 1, \ldots, \infty$ and $\bigcup_{j=1}^{\infty} I_j$ is dense on $[-1,1]$ (Exercise #4).

5. Construct and claim that

$$H = [-1,1] \setminus \bigcup_{j=1}^{\infty} I_j = \{-1,1\}. \text{(Exercise \#5)}$$

6. Combining (4) and (5), we derive f satisfies the condition in Proposition(4.2), and therefore is a polynomial. (Exercise #6)

■

Verification. Here we give some hints for the exercises above:

1. Since the inverse image of $\{0\}$ is closed for continuous functions, and $f^{(n)}(\cdot)$ is continuous, we derive E_n's are closed.
2. The Baire-Category Theorem asserts that for a non-empty complete metric space X, or any subsets of X with **non-empty** interior, if it is the countably union of **closed** sets, then one of these **closed** sets has non-empty interior.
3. By integrating $f^{(N_1)}$ for N_1 times, e.g.,

$$f^{(N_1)} = 0 \implies f^{(N_1-1)} = \int f^{(N_1)}\,\mathrm{d}x = a_0 \implies \cdots \implies f = a_{N_1-1}x^{N_1-1} + \cdots + a_0$$

4. Let $I \subseteq [-1,1]$ be any open interval. Its clousure can be expressed as:

$$\overline{I} = \bigcup_{n=1}^{\infty} \overline{I} \bigcap E_n$$

   Applying Baire category theorem to $\overline{I}$, at least one $\overline{I} \bigcap E_{n'}$ contains an open interval I'. Thus, $I' \subseteq I$ and $I' \subseteq E_{n'}$, which implies $I' \in \bigcup_{j=1}^{\infty} I_j$ (recall that I_j's are picked maximally). This means that $\bigcup_{j=1}^{\infty} I_j \bigcap I$ is non-empty for arbitrary open interval I, which implies $\bigcup_{j=1}^{\infty} I_j$ is dense.
5. We have seen that $\bigcup_{j=1}^{\infty} I_j$ is an open, dense **proper** subset of $[0,1]$, which means H is **non-empty**, closed, and nowhere dense in $[0,1]$. In order to show $H = \{-1,1\}$, it suffices to show H does not contain open intervals. Otherwise applying Baire-Category Theorem to $H = \bigcup_{n=1}^{\infty} E_n \bigcap H$ again, for some fixed n^*, $E_{n^*} \cap H$ contains an open interval I^*. Thus, $I^* \subseteq E_{n^*}$ and $I^* \subseteq H \implies I^* \subseteq (\bigcup_{j=1}^{\infty} I_j)^c$, which leads to a contradiction as I_j's are picked maximally.
6. Hence, $\bigcup_{j=1}^{\infty} I_j = (-1,1)$, i.e., $f(x) = \sum_{k=0}^{\infty} a_k x^k$ in $|x| < 1$.

A simpler and more clear proof is presented in the website

`https://mathoverflow.net/questions/34059/`

`if-f-is-infinitely-differentiable-then-f-coincides-with-a-polynomial`

We have seen some examples of nowhere differentiable functions. Now we show that almost functions are nowhere differentiable.

Notations. We denote $\mathcal{C}[0,1]$ as the set of all continuous functions on $[0,1]$. One corresponding metric is defined as:

$$d(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|, \quad \forall f,g \in \mathcal{C}[0,1].$$

Remember that $(\mathcal{C}[0,1], d)$ is complete.

Theorem 4.1 The set of all nowhere differentiable functions in $(\mathcal{C}[0,1], d)$ is dense, i.e., forms a 2nd Category.

The trick is to show the complement of the set of nowhere differentiable functions, i.e., the set of functions that have a **finite** derivative at some point, forms a 1st Category.

Proof. Construct

$$E_n = \left\{ f \in \mathcal{C}[0,1] \,\middle|\, \begin{array}{r} \forall 0 < h < 1 - x, \left| \dfrac{f(x+h) - f(x)}{h} \right| \le n, \\ \text{for some } 0 \le x \le 1 - \frac{1}{n} \end{array} \right\}$$

Thus the union of all E_n will contain all functions having a finite **right hand derivative** at some point in $[0,1)$.

Proposition 4.3 E_n is closed, i.e., for a sequence of function $\{f_m\} \subseteq E_n$ such that $f_m \to f$, we have $f \in E_n$.

Proposition 4.4 E_n is nowhere dense, i.e., ($\mathcal{C} \setminus E_n$ is dense):

After showing these two propositions, we conclude that the set of functions, with a right derivatives at some point, is a set of the first category. Similarly, we can repeat these steps for left derivatives. In summary, the set of functions with a well-defined derivatives forms a 1st Category. The proof is complete. ∎

Proof of Proposition(4.3). Since $\{f_m\} \subseteq E_n$, there exists a sequence of $\{x_m\}$ such that for each m,

$$0 \leq x_m \leq 1 - \frac{1}{n}$$
$$|f_m(x_m + h) - f_m(x_m)| \leq hn,$$

for $\forall 0 < h < 1 - x_m$. As $\{x_m\}$ is bounded, there exists a subsequence $\{x_{m,k}\}$ of $\{x_m\}$ with limit $x \in [0, 1 - \frac{1}{n}]$.

For $\forall 0 < h < 1 - x$, we have that $0 < h < 1 - x_{m,k}$ for large k. Applying triangle inequality, we obtain:

$$\begin{aligned} |f(x+h) - f(x)| &\leq |f(x+h) - f(x_{m,k}+h)| + |f(x_{m,k}+h) - f_m(x_{m,k}+h)| \\ &\quad + |f_m(x_{m,k}+h) - f_m(x_{m,k})| + |f_m(x_{m,k}) - f(x_{m,k})| + |f(x_{m,k}) - f(x)| \\ &\leq |f(x+h) - f(x_{m,k}+h)| + d(f, f_k) + nh + d(f_k, f) + |f(x_k) - f(x)|. \end{aligned}$$

Taking $k \to \infty$, we find all terms in RHS goes to zero except nh:

$$|f(x+h) - f(x)| \leq nh \implies f \in E_n.$$

∎

Proof of Proposition(4.4). In order to show E_n is nowhere dense, by using the fact that E_n is closed, it suffices to show that an arbitrary open neighborhood $B(f, \varepsilon)$ will contain elements from the set $\mathcal{C}[0,1] \setminus E_n$, i.e., it suffices to create a function in $B(f, \varepsilon)$ that cannot be in E_n for fixed ε.

- Construct a piecewise linear function $\phi_N(x)$ on $[0,1]$ first:

$$\phi_N(x) = \begin{cases} N(x - \frac{k}{N}), & \frac{k}{N} \leq x \leq \frac{k+1}{N}, k = 0, 2, \ldots, N \\ -N(x + \frac{k+1}{N}), & \frac{k}{N} \leq x \leq \frac{k+1}{N}, k = 1, 3, \ldots, N-1 \end{cases}$$

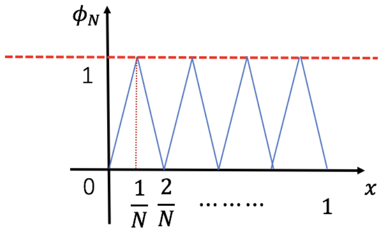


Figure 4.1: Plot of function $\phi_N(x)$

As we can see, N is the maximum slope of the piecewise linear function ϕ_N.

- Let M be the maximum slope of the piecewise linear function f, and pick a positive even integer m such that

$$\frac{1}{2}mN\varepsilon > M + n.$$

Then we construct function

$$g(x) = f(x) + \frac{1}{2}\varepsilon\phi_{mN}(x)$$

As we can see, $d(f,g) = \frac{1}{2}\varepsilon < \varepsilon$, thus $g \in B(f,\varepsilon)$. Also note that

$$\begin{aligned}
\left|\frac{g(x+h)-g(x)}{h}\right| &\geq \left|\frac{1/2\varepsilon(\phi_{mN}(x+h)-\phi_{mN}(x))}{h}\right| - \left|\frac{f(x+h)-f(x)}{h}\right| \\
&\geq \frac{1}{2}\varepsilon\left|\frac{(\phi_{mN}(x+h)-\phi_{mN}(x))}{h}\right| - M \\
&= \frac{1}{2}mN\varepsilon - M > n
\end{aligned}$$

for x in $(0, 1-\frac{1}{mN})$ and some $h \in (0, 1-x)$. Hence, $g \notin E_n$. The proof is complete.

■

4.1.2. Continuity Analysis

Recall the definition for continuity:

- A function f is said to be continuous at $x_0 \in I$ if $\forall \varepsilon > 0$, there exists $\delta > 0$ (δ depends on x_0 and ε) such that

$$|f(x) - f(x_0)| < \varepsilon, \quad \forall |x - x_0| < \delta$$

- A function f is continuous on I if it is continuous at every point in I.

Definition 4.1 [Uniform] We say f is **uniformly continuous** on I if $\forall \varepsilon > 0$, there exists δ (depend only on ε, but independent of $x \in I$) such that

$$|f(y) - f(x)| < \varepsilon, \text{ if } |x - y| < \delta$$

R It is useful to note that the uniform continuity places a upper bound on the growth of the function at every point, i.e., the function cannot grow too fast.

Example 4.1 Given a function $f(x) = x^2$,

1. Is it uniformly continuous on $[0,1]$?

   Yes, intuitively the growth of x^2 is limited within bounded interval.

2. Is it uniformly continuous on $\mathbb{R}$?

   No, intuitively the growth of x^2 tends to infinite as $x \to \infty$.

   Proof: For fixed x, if $|y - x| < \delta$, if we choose $|x| \ge \frac{\varepsilon}{2\delta} + \frac{\delta}{2}$, then

$$\begin{aligned}
\underbrace{|f(y) - f(x)|}_{\varepsilon} &= |y^2 - x^2| = |y + x| \underbrace{|y - x|}_{\delta} \\
&\ge (|2x| - |x - y|)|y - x| \ge (\frac{\varepsilon}{\delta} + \delta - \delta)\delta = \varepsilon
\end{aligned}$$

   which is a contradiction.

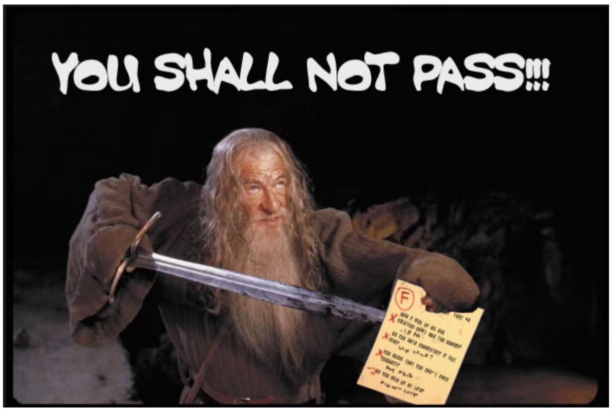


Figure 4.2: The proof and application for the Theorem(4.2) is Mandatory. If you don't know how to do it in the exam, Prof.Ni will fail you without hesitation.

Theorem 4.2 Suppose that f is continuous on a compact set D. Then f is uniformly continuous on D.

Proof. For given $\varepsilon > 0$, since f is continuous at x, there exists $\delta_x > 0$ s.t.

$$|f(y) - f(x)| < \tfrac{\varepsilon}{2}, \quad \text{if } |y - x| < \delta_x.$$

Construct an open cover $\{B_{\delta_x}(x) \mid x \in D\}$ of D with

$$B_{\delta_x}(x) = \{y \in D \mid |y - x| < \frac{1}{2}\delta_x\}.$$

The set D is compact implies there exists a finite subcover:

$$D \subseteq B_{\delta_{x_1}}(x_1) \bigcup B_{\delta_{x_2}}(x_2) \bigcup \cdots \bigcup B_{\delta_{x_k}}(x_k). \tag{4.1}$$

Construct $\delta > 0$ such that $B_\delta(x)$ must be contained entirely in one of the ball, say $B_{\delta_{x_j}}(x_j)$ (Exercise #7)

Therefore given $|y - x| < \delta$ we imply $x, y \in B_{\delta_{x_j}}(x_j)$ for some j, which follows that

$$|f(y) - f(x)| \le |f(y) - f(x_j)| + |f(x_j) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

■

Verification of Exercise. Such a δ is constructed as

$$\delta = \frac{1}{2}\min\{\delta_{x_1}, \delta_{x_2}, \ldots, \delta_{x_k}\}.$$

Thus for any x, y with $|y - x| < \delta$, by(4.1), there exists j such that $x \in B_{\delta_{x_j}}(x_j)$, and hence

$$|x - x_j| < \frac{1}{2}\delta_{x_j} \tag{4.2}$$

Also, we have

$$|y - x_j| \le |y - x| + |x - x_j| \le \delta + \frac{1}{2}\delta_{x_j} \le \delta, \tag{4.3}$$

i.e., y is also in $B_{\delta_{x_j}}(x_j)$.

Definition 4.2 [Convex] A real-valued function f defined in (a, b) is said to be convex if

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y)$$

whenever $a < x < b, a < y < b, 0 < t < 1$.

Check Rudin's book for the proof that a convex function is always continuous.

4.2. Friday

This lecture will finish the topic for continuity, and we will have a very simple, easy quiz on Sunday.

4.2.1. Continuity Analysis

Definition 4.3 [Lipschitz Continuity] A function f is **Lipschitz continuous** at x_0 if there exists a constant M (depend on x_0) such that

$$|f(x) - f(x_0)| \le M|x - x_0|,$$

for $\forall |x - x_0|$ small.

R

- Note that Lipschitzness places a upper bound on the growth of the funciton that is linear in the perturbation, i.e., $|x - x_0|$.
- Also notice that Lipschitz functions need not be differentiable, e.g., $f(x) = |x|$ is Lipschitz continuous at $x = 0$.
- However, differentiable functions with bounded derivative are always Lipschitz.
- Lipschitz continuous functions are always continuous. (choose $\delta = \varepsilon / M$)

A property similar to Lipschitzness is that of Holder continuity.

Definition 4.4 [Holder] A function f is said to be **Holder continuous** of order α at x_0 if there exists a constant M such that

$$|f(x) - f(x_0)| \le M|x - x_0|^{\alpha}$$

for $\forall |x - x_0|$ small, where $0 < \alpha < 1$.

R

- $f(x) = \sqrt{|x|}$ is Holder continuous of order $1/2$ at $x = 0$.
- Holder continuous functions are always continuous (choose $\delta = (\varepsilon/M)^{1/\alpha}$)

Holder Continuity for Differentiable Equation. Solving differentiable equations is a core topic in pure math. When solving the ODE $u'' = f(x)$ with f continuous, we can say $u \in \mathcal{C}^2$. However, when talking about the PDE $u_{xx} + u_{yy} = f(x,y)$ with continuous f, u is not necessarily twice continuously differentiable. Instead, u is almost $\mathcal{C}^2$. However, if given the extra condition $f \in \mathcal{C}^{\alpha}_{\text{holder}}$, then we imply $u \in \mathcal{C}^{2+\alpha}$. Holder continuous is vital important for future mathematics study.

Definition 4.5 [Convex] A funciton f is said to be **convex** in (a,b) if

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

holds for $\forall x,y \in (a,b)$ and $\forall t \in [0,1]$

R The geometrically meaning for convexity is that the function evaluated in the line segment is lower than secant line between x and y, i.e., a convex function f lies below secant line.

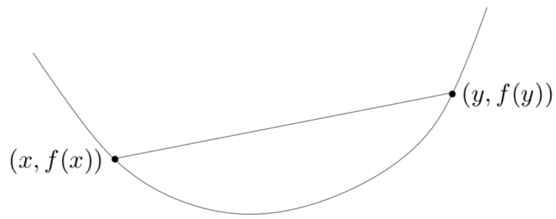
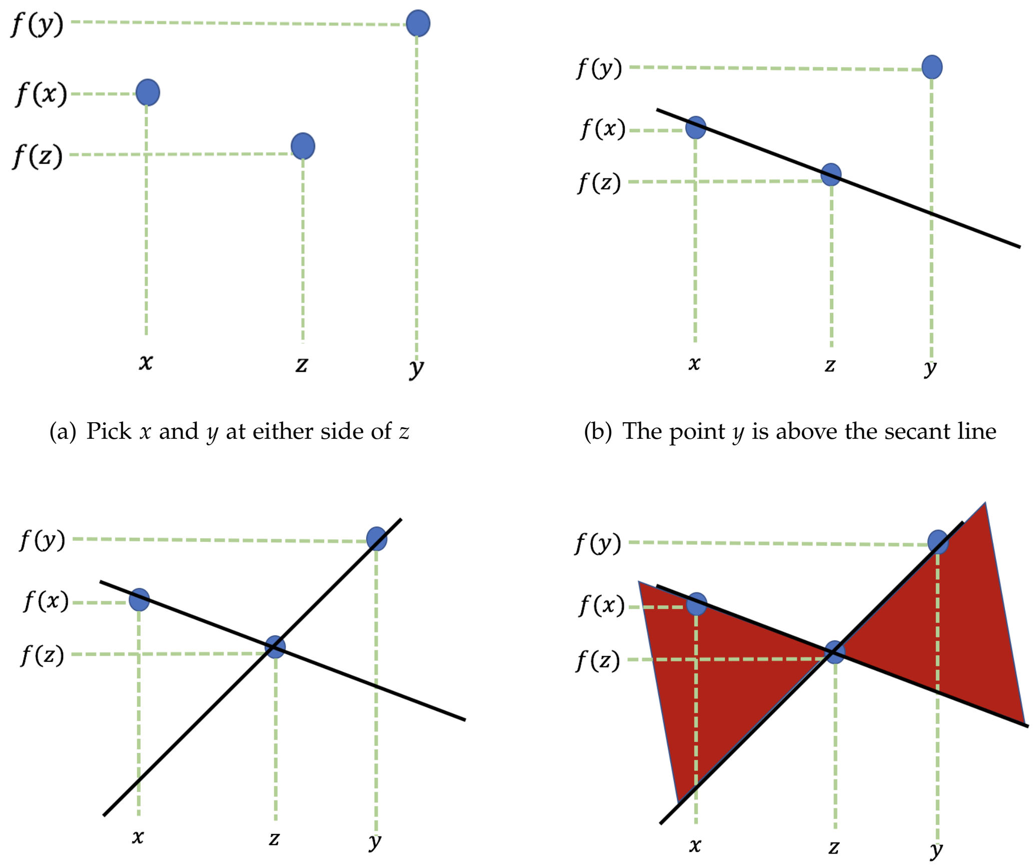


Figure 4.3: Graph of a convex function. The line segment between any two points on the graph lies above the graph.

An intuitive statement is that a convex function is always continuous. This statement can be shown pictorially.

Proposition 4.5 A convex function must be continuous.



(a) Pick x and y at either side of z

(b) The point y is above the secant line

(c) The point x is below the secant line

(d) The function is inside the red region

Figure 4.4: Graphic Proof for Proposition(4.5)

Proof. The proof is given in Fig:(4.4):

- To show the continuity for z, pick x and y at either side.
- By convexity, the y lies above the secant line between x and z, otherwise draw a secant line between x and y, then z lines above the secant line (contradiction).
- Again, x lines below the secant line between z and y.
- Hence, the function lies inside the red region shown in Fig:(4.4(d)).
- Pick a sequence $\{z_n\} \to z$, the function $f(z_n)$ must converge to $f(z)$.

The proof is complete. ■

The following two statements are left as exercise:

Proposition 4.6 Assume that f is continuous real function defined in (a,b) such that

$$f\left(\frac{x+y}{2}\right) \le \frac{f(x)+f(y)}{2}$$

for all $x, y \in (a,b)$. Then f is convex.

R We can pick an example to show that dropping out the continuity condition will make this statement false.

4.2.2. Monotone Analysis

Definition 4.6 [Monotone] The function f is said to be **monotone increasing** (**decreasing**) if $f(x) \le f(y)$ $(f(x) \ge f(y))$ whenever $x < y$.

R

- For monotone functions,

$$\lim_{x\to x_0+} f(x), \quad \lim_{x\to x_0-} f(x),$$

always exist.

- The difference

$$\lim_{x \to x_0+} f(x) - \lim_{x \to x_0-} f(x)$$

is called the **jump discontinuity** at $x = x_0$.

Proposition 4.7 A monotone function f on (a, b) is continuous except at possibly a countable number of points.

We show the statement is true for interval $[a + \frac{1}{n}, b - \frac{1}{n}]$ first. That's because the function inside a closed intercal is finite.

Proof.

- Let S_k denote the set of all points in $[a + \frac{1}{n}, b - \frac{1}{n}]$ on which f has a jump discontinuity no less than $\frac{1}{k}$.
- Then the set S_k must be finite for every k, otherwise the total jump of S_k will be infinite, which contradicts the range of f.
- The set of all discontinuous points of f on $[a + \frac{1}{n}, b - \frac{1}{n}]$ is the union $\bigcup_{k=1}^{\infty} S_k$, therefore is at most countable.
- Note that

$$(a, b) = \bigcup_{n=1}^{\infty} [a + \frac{1}{n}, b - \frac{1}{n}],$$

the countably union of at most countable sets is at most countable. Tshe proof is complete.

∎

Exercise. Given the Riemann function

$$f(x) = \begin{cases} 0, & x \in \mathbb{Q} \\ \frac{1}{q}, & x = \frac{p}{q}, (p, q) = 1. \end{cases}$$

Is f Lipschitz continuous, Holder continuous, or whatever at the given point $x_0 \notin \mathbb{Q}$?

4.2.3. Cantor Set

Now we describe a complicated subset of $\mathbb{R}$ which is uncountable. This set is called the Cantor set.

Geometric description. We start with the unit interval

$$F_0 = [0,1]$$

Now define a new set

$$F_1 = F_0 \setminus (\frac{1}{3},\frac{2}{3}) = [0,1/3]\bigcup[2/3,1],$$

i.e., we obtain F_1 by deleting the open middle third of F_0.

Next we obtain a new set F_2 by deleting the open middle thirds of each of the intervals making up F_1:

$$F_2 = [0,1/9]\bigcup[2/9,1/3]\bigcup[2/3,7/9]\bigcup[8/9,1]$$

Continue in this way to obtain sets F_n, $n \geq 0$, where F_n consists of 2^n disjoint closed intervals of length 3^{-n}, formed by deleting the middle thirds of the intervals making up F_{n-1}.

The **Cantor set** is defined to be the intersection of these sets:

$$F = \bigcup_{n=1}^{\infty} F_n$$

Arithmetic description. We also have

$$F = \left\{ x \in [0,1] : x = \sum_{n=1}^{\infty} a_n 3^{-n}, a_n \in \{0,2\}, n \geq 1 \right\}$$

Here we can also describe F as the set of reals with a tenary expansion

$$0.a_1a_2\dots a_n\dots \qquad a_n \in \{0,2\}.$$

For example, given $x = 3/4$, how to find the tenary expansion $x \sim 0.a_1a_2\ldots$? Clearly $3/4 \in [\frac{2}{3}, 1]$, i.e., $a_1 = 1$. The next interval is determined by which interval among $[\frac{2}{3}, \frac{7}{9}]$ and $[\frac{8}{9}, 1]$ that x belongs. It is clearly that it belongs to the first interval, and so $a_2 = 0$. Thus we can find a_n recursively via this way to get the tenary expansion.

Proposition 4.8 There is a one-to-one correspondence between F and F_0.

Proof. For any $x \in F$, we write $x = 0.a_1a_2\cdots a_j$ with $a_j = 0$ or 2. Also, we can construct our $y = 0.b_1b_2\cdots$ with $b_j = \frac{a_j}{2}$, i.e., y is any number with binary expansion. Therefore, we establishes a one to one correspondence between F and the set of points in $(0,1)$. ∎

Also, there is a one-to-one correspondence between F and F_0 implies F is **uncountable**.

Proposition 4.9 The measure of F is 0.

Proof. The F_1 is constructed by taking away an interval with length $\frac{1}{3}$. The F_2 is constructed by taking away intervals with total length $\frac{2}{9}$, etc. Therefore, the toal length of taking-away intervals is given by:

$$\begin{aligned}
\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots \\
&= \frac{1}{3} + \frac{2}{3^2} + \frac{4}{3^3} + \cdots \\
&= \frac{1}{3}(1 + \frac{2}{3} + (\frac{2}{3})^2 + \cdots) \\
&= 1
\end{aligned}$$

∎

Proposition 4.10 F is closed and nowhere dense.

Proof. F is closed since it is the complement of the union of open intervals. The Cantor set is nowhere dense as its closure has empty interior. ∎

The proof for proposition(4.11) is left as exercise.

Proposition 4.11 Every point in F is a limit point of F

Definition 4.7 [Perfect] A closed set S is **perfect** if every point in S is a limit point of S. ■

The proof for proposition(4.12) is left as exercise.

Proposition 4.12 Any perfect set is uncountable.

Chapter 5

Week5

5.1. Wednesdays

Today we will discuss topics about differentiation. Note that the topics in Friday will be more difficult.

5.1.1. Differentiation

Notations and Conventions. Given two functions ϕ and ξ, then we denote $\phi(x) = O(\xi(x))$ near x_0 if there exists a constant C such that

$$\left|\frac{\phi(x)}{\xi(x)}\right| \le C \text{ for } x \text{ near } x_0,$$

i.e., $|\phi(x)| \le C|\xi(x)|$; also, $\phi(x) = o(\xi(x))$ near x_0 if

$$\left|\frac{\phi(x)}{\xi(x)}\right| \to 0 \text{ as } x \to x_0,$$

i.e., $\forall \varepsilon > 0$, there exists $\delta > 0$ s.t. $|\phi(x)| \le \varepsilon|\xi(x)|$ if $|x - x_0| \le \delta$.

R

1. In particular, if $f(x) \to 0$ as $x \to x_0$, we write $f(x) = o(1)$ near x_0.
2. $|\frac{\phi(x)}{h}| = o(1) \iff |\phi(x)| = o(h)$.

Definition 5.1 [Derivative] Given a function f, if the limit

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} \text{ exists,}$$

we say that f is **differentiable** at x_0 and the limit is called the **derivative** of f at x_0, denoted as $f'(x_0)$.

Geometric Interpretation. Derivative has its geometric meaning:

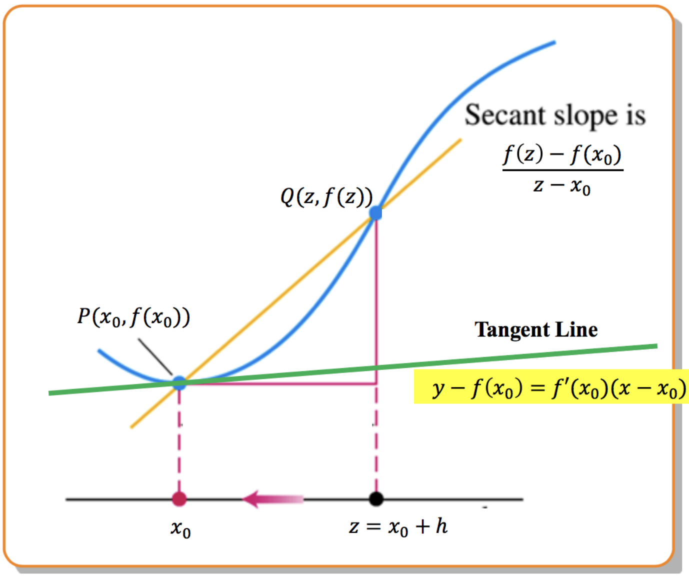


Figure 5.1: Interpretations of Derivative

Here the derivative is essentially the **slope** of the tangent line to the curve $y=f(x)$ at $x=x_0$, where the tangent line is

$$y-f(x_0)=f'(x_0)(x-x_0)$$

Here the tangent line is defined as the limit of secant line, i.e., taking limit $z\to x_0$, the secant line between x_0 and z becomes the tangent line.

Arithmetic Insights of derivative. The limit

$$f'(x_0):=\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h},$$

can be equivalently rewritten as:

- $\left|\frac{f(x_0+h)-f(x_0)}{h} - f'(x_0)\right| \to 0(=o(1))$ as $h \to 0$,
- $|f(x_0+h) - f(x_0) - f'(x_0)h| = o(h) \qquad$ as $h \to 0$,
- $f(x_0+h) = f(x_0) + f'(x_0)h + o(h)$ as $h \to 0$. (useful equivalent definition for derivative)

Substituting $h = x - x_0$ into the above equation, we derive

$$|f(x) - [f(x_0) + f'(x_0)(x-x_0)]| = o(x-x_0), \quad \textbf{Important Formula}$$

which essentially means that *the differential provides the best linear approximation to the function in a neighborhood of a point*.

5.1.2. Basic Rules of Differentiation

Given two functions g, f, we study the derivative evaluated for the composite function $g \circ f$:

Proposition 5.1 $(g \circ f)' = g'(f(x)) \cdot f'(x)$

Let's see how the engineer gives a proof: (recall MAT1001 slides, they exactly done this)

Engineers' proof. Recall the definition

$$(g \circ f)'(x) := \lim_{h \to 0} \frac{(g \circ f)(x_0+h) - (g \circ f)(x)}{h}.$$

Note that

$$\frac{(g \circ f)(x_0+h) - (g \circ f)(x)}{h} = \frac{(g \circ f)(x_0+h) - (g \circ f)(x)}{f(x+h) - f(x)} \frac{f(x+h) - f(x)}{h}.$$

Taking the limit $h \to 0$, the first term approaches $g'(f(x))$, and the second approaches $f'(x)$. ■

Comments. Note that $f(x+h)-f(x)$ may not be necessarily non-zero, and there is meaningless to pick a zero term in denominator.

Mathematicians' Proof. Recall the definition

$$(g\circ f)(x+h)-(g\circ f)(x):=g(f(x)+f'(x)h+o(h))-g(f(x))$$

The differentiablity of g gives us $g(y)=g(y_0)+g'(y_0)(y-y_0)+o(y-y_0)$, and therefore

$$\begin{aligned}(g\circ f)(x+h)-(g\circ f)(x)&=g(f(x)+f'(x)h+o(h))-g(f(x))\\&=\{g(f(x))+g'(f(x))\left[(f'(x)h)+o(h)\right]\}-g(f(x))\\&=g'(f(x))f'(x)h+\underbrace{g'(f(x))}_{\text{fixed as }x\text{ is fixed}}o(h)=g'(f(x))f'(x)h+o(h),\end{aligned}$$

■

and therefore $(g\circ f)'(x)=g'(f(x))f'(x)$.

5.1.3. Analysis on Differential Calculus

Differentiablity implies continuity.

Proposition 5.2 If f is differentiable at x_0, then it is continuous at x_0.

Proof. By definition of differentiablity,

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

Taking the limit $x\to x_0$, the RHS approaches 0, i.e., $\lim_{x\to x_0}f(x)=f(x_0)$. ■

Theorem 5.1 — Rolle's Theorem. Suppose f is differentiable on $[a,b]$, then $f'(x_0)=0$ if $x_0\in(a,b)$ is a local maximum or local minimum.

Proof. Note that

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Suppose that x_0 is a local maximum, then

- $x > x_0$ and x close to x_0 implies $f(x) - f(x_0) \le 0$, i.e., $\lim_{x \to x_0 +} \frac{f(x)-f(x_0)}{x-x_0} \le 0$
- Similarly, $x < x_0$ and x close to x_0 implies $\lim_{x \to x-} \frac{f(x)-f(x_0)}{x-x_0} \ge 0$.

Therefore

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0 +} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x-} \frac{f(x) - f(x_0)}{x - x_0} = 0.$$

■

R Geometrically this theorem is obvious. It asserts that at an extreme point of a differentiable function, the tangent line to its graph is horizontal.

Estimation on finite increments. The following two proposition are the most frequently used and important methods of studying numerical-valued functions.

> **Theorem 5.2 — Mean-Value Theorem.** Suppose f is differentiable on $[a, b]$, then there exists a point $c \in (a, b)$ such that
>
> $$\frac{f(b) - f(a)}{b - a} = f'(c)$$

The proof relies on Rolle's theorem, i.e., we need to construct a function h on $[a, b]$ that has a interior local maximum or local minimum. This is one useful trick.

Proof. We set $g(x) = \frac{f(b)-f(a)}{b-a}(x - a) + f(a)$, i.e., g is a secant line between $(a, f(a))$ and $(b, f(b))$. Then we consider the auxiliary function $h(x) = f(x) - g(x)$, which implies $h(a) = h(b) = 0$.

- If $h \equiv 0$, then $g \equiv f$, and therefore $g' \equiv f'$, i.e.,

$$g'(x) = \frac{f(b) - f(a)}{b - a} = f'(x), \qquad \forall x \in (a, b)$$

- Otherwise, h is positive or negative somewhere in (a, b). w.l.o.g., h is positive somewhere in (a, b). Thus h assumes its maximum in (a, b), say c (Exercise # 1). By rolle's theorem $h'(c) = 0$, which impleis $f'(c) = g'(c) = (f(b) - f(a))/(b - a)$.

■

Applying Mean-Value Theorem, we give a more useful verison of Rolle's Theorem:

Corollary 5.1 [Rolle's Theorem verison 2] Suppose f is differentiable on $[a, b]$ with $f(a) = f(b)$, then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.

Proof. By Mean-Value Theorem,

$$f(b) - f(a) = 0 = f'(c)(b - a) \implies f'(c) = 0.$$

■

Exercise Verification. Recall the proposition we have learnt:

Proposition 5.3 The range of a continuous function over a compact set is also compact.

As a result, $\sup_{x \in [a,b]} h = \max_{x \in [a,b]} h$, i.e., h assumes its maximum in $[a, b]$.

R Mean Value Theorem is correct, but not precise, i.e., we obtain an estimate of a function using affine. Now we are interested in approximations of a function by a polynomial $P_n(b) = P_n(b; a) = c_0 + c_1(b - a) + \cdots + c_n(b - a)^n$. The Taylor's theorem gives the answer using integrals.

Theorem 5.3 — Taylor's Theorem. Let f be n times differentiable on the open interval with $f^{(n-1)}$ continuous on the closed interval between a and x, then

$$f(x)=f(a)+f'(a)(x-a)+f''(a)\frac{(x-a)^2}{2!}+\cdots+f^{(n-1)}(a)\frac{(x-a)^{n-1}}{(n-1)!}+R_n(x),$$

where the remainder is given by:

$$R_n(x)=\frac{1}{(n-1)!}\int_a^x f^{(n)}(t)(x-t)^{n-1}\,\mathrm{d}t$$

Proof. All techniques in this proof is **integration by parts**:

$$\begin{aligned}f(x)&=f(a)+\int_a^x f'(t)\,\mathrm{d}t=f(a)-\int_a^x f'(t)\,\mathrm{d}(x-t)\\&=f(a)-f'(t)(x-t)|_b^a+\int_a^x f''(t)\,\mathrm{d}t=f(a)+f'(a)(x-a)+\int_a^x (x-t)f''(t)\,\mathrm{d}t\end{aligned}$$

Applying the similar trick gives the result. ■

Motivation of Continuosly differentiable.

R If $f'(x_0)>0$, then f is increasing at x_0. However, it does not mean that f is increasing in a neighborhoold of x_0. That's because f' may not be continuous.

Some mathematicians give an example of a function that is everywhere differentiable but nowhere monotone:

■ **Example 5.1** There exists a function haing the following properties:

1. f is differentiable with $|f'|\le 1$ on $\mathbb{R}$
2. Both $\{x\in\mathbb{R}\mid f'(x)>0\}$ and $\{x\in\mathbb{R}\mid f'(x)<0\}$ are dense in $\mathbb{R}$
3. $\{x\in\mathbb{R}\mid f'(x)=0\}$ is also dense in $\mathbb{R}$
4. f is the difference of 2 monotone functions
5. f' is **not** Riemann integrable.

`Reference: http://citeseerx.ist.psu.edu/viewdoc/download?`

```
doi=10.1.1.843.5011&rep=rep1&type=pdf
```

In the future when handling such topic, we will assume f is **continuously differentiable**, denoted as $f \in \mathcal{C}^1$. Then the example above will not appear.

5.2. Friday

We plan to have a make-up lecture tomorrow, which will cover problems in Quiz 1.

5.2.1. Analysis on Derivative

Not all functions could be the derivative of some functions. Let's give sufficient or necessary conditions for that.

Criteria 1: The image of the derivative of any function should be continuous.

> **Theorem 5.4 — Intermediate-Value Property for derivative.** Let f be differentiable on $[a,b]$, and let y_0 be a number between $f'(a)$ and $f'(b)$. Then there exists $c \in (a,b)$ such that $f'(c) = y_0$.

Proof. Consider an auxiliary function

$$h(x) = f(x) - y_0 x,$$

which follows that $h'(a) = f'(a) - y_0$ and $h'(b) = f'(b) - y_0$. w.l.o.g., $f'(a) < f'(b)$, and therefore $h'(b) < 0 < h'(a)$, i.e.,

$$\begin{cases} \lim_{x \to b+} \frac{h(x)-h(b)}{x-b} < 0 \\ \lim_{x \to a-} \frac{h(x)-h(a)}{x-a} > 0 \end{cases} \implies \begin{cases} h(x) < h(a) \text{ for some } x > a \\ h(x) < h(b) \text{ for some } x > b \end{cases},$$

i.e., h has a minimum on $[a,b]$, say at c. By Rolle's theorem, $h'(c) = 0$, i.e., $f'(c) = y_0$. ∎

R This theorem tells us that the image for the derivative of any function (on $[a,b]$) contains the whole interval between $f'(a)$ and $f'(b)$. Hence, the step function cannot be the derivative of other functions, since its image does not contain the interval $[0,1]$ but only the endpoints.

However, it does not asset that the derivative of any function should be continuous (the image should be). A simple example of this is the function

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

the derivative $f'(x)$ is not continuous at $x = 0$.

The question turns out how continuity should be for the derivative of any function?

Criteria 2: The derivative of any function should be continuous on a dense set. For the derivative of any function f, if exists, can be expressed as the pointwise limit:

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{n \to \infty} \underbrace{n \left[f(x + \frac{1}{n}) - f(x) \right]}_{f_n(x)} \\ &:= \lim_{n \to \infty} f_n(x), \end{aligned}$$

with $f_n(x) = n\left[f(x + \frac{1}{n} - f(x))\right]$. Thus the set of all discontinuous points of f' is a set of first categroy (Recall Theorem(3.3)). In particular, f' must be continuous on a dense set.

5.2.2. Analysis on Mean-Value Theorem

The following proposition is a useful generalization of the standrad mean-value theorem, and is also based on the Rolle's theorem. From this theorem we can also imply the L-Hopital's Rule:

Theorem 5.5 — Cauchy's Mean-Value Theorem. Let f and g be two differentiable function on $[a,b]$. Then there exists a point $c \in (a,b)$ such that

$$g'(c)[f(b)-f(a)] = f'(c)[g(b)-g(a)] \tag{5.1}$$

R

1. When $g(x) := x$, Cauchy's Mean-Value Theorem becomes the Mean-Value Theorem.
2. If in addiction $g'(x) \neq 0$ for each $x \in (a,b)$, then $g(b) \neq g(a)$ and we have the equality version of (5.1):

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

The idea of proof is to construct an auxiliary function satisfying the hypotheses of Rolle's theorem version 2 (Corollary (5.1)), which is the same trick used in Theorem(5.2) and (5.4).

Proof. We construct a function

$$h(x) = g(x)[f(b)-f(a)] - f(x)[g(b)-g(a)],$$

which follows that $h(a) = h(b) = g(a)f(b) - f(a)g(b)$. By Rolle's theorem, there exists $c \in (a,b)$ s.t. $h'(c) = 0$, i.e., $g'(c)[f(b)-f(a)] = f'(c)[g(b)-g(a)]$. ■

Now we pause to discuss a special but very useful technique for findind the limit of a ratio of functions, known as L-Hopital's Rule[1].

Theorem 5.6 — L-hospital's Rule. Suppose f and g are both differentiable on (a,b), and $g'(x) \neq 0$ in (a,b). Suppose that

$$\lim_{x\to a+} \frac{f'(x)}{g'(x)} = l,$$

[1] G.F.de l'Hopital, a French mathematician, a capable student of Johann Bernoulli. The L-Hopital's Rule is really due to Johann Bernoulli, but l'Hopital is so rich so that there is a deal on the table, and thus the rule was published in slightly altered altered form by "l'Hopital".

then

1. $\lim_{x \to a+} f(x) = 0 = \lim_{x \to a+} g(x)$ implies $\lim_{x \to a+} \frac{f(x)}{g(x)} = l$.
2. $\lim_{x \to a+} f(x) = +\infty, \lim_{x \to a+} g(x) = \infty$ implies $\lim_{x \to a+} \frac{f(x)}{g(x)} = l$.

We discuss the proof for the first case, and the second case is left as exercise.

Proof. Note that $g(x) \neq g(y)$ for $\forall x, y \in (a, b)$, otherwise there will be some $c \in (a, b)$ such that $g'(c) = 0$. By Cauchy's Mean-Value Theorem, for $x, y \in (a, b)$, there exists $z \in (y, x)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(z)}{g'(z)}$$

Or equivalently,

$$\frac{f(x)}{g(x)} = \frac{f(y)}{g(x)} + \frac{f'(z)}{g'(z)}\left[1 - \frac{g(y)}{g(x)}\right], \forall x, y \in (a, b). \tag{5.2}$$

For fixed x, pick $y < x$ which is close to $a+$ such that $\frac{f(y)}{g(x)}$ and $\frac{g(y)}{g(x)}$ both small (the reason we can do that is because $\lim_{y \to a+} f(y) = 0 = \lim_{y \to a+} g(y)$). Thus taking the limit $x \to a+$, by (5.2), we derive

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = \lim_{x \to a+} \frac{f'(z)}{g'(z)} = l.$$

■

We can also derive the L-hopital's rule by Taylor expansion, but keep note that we should add one moer condition that $f, g \in \mathcal{C}^1$ (so that we can apply taylor expansion)

Proof. We expand $f(x)$ and $g(x)$ in a small neighborhood of a with $x \neq a$:

$$\begin{aligned}
\frac{f(x)}{g(x)} &= \frac{f(a) + f'(a)(x-a) + o(x-a)}{g(a) + g'(a)(x-a) + o(x-a)} \\
&= \frac{(x-a)[f'(a) + o(1)]}{(x-a)[g'(a) + o(1)]} \\
&= \frac{f'(a) + o(1)}{g'(a) + o(1)} \to \frac{f'(a)}{g'(a)}, \qquad x \to a.
\end{aligned}$$

■

Note that L-hopital's rule is a technique that should be cleverly used, otherwise the limit can be messy to handle. Prof.Yeye Ni does not like this technique. Now we give an example:

■ **Example 5.2**

$$
\begin{aligned}
\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^{1/(1-\cos x)} &= \lim_{x\to 0}\exp\left[\frac{\ln(\frac{\sin x}{x})}{1-\cos x}\right] \\
&= \exp\left[\lim_{x\to 0}\frac{\ln(\frac{\sin x}{x})}{1-\cos x}\right] \\
&= \exp\left[\lim_{x\to 0}\frac{\cos x-\sin x/x}{\sin^2 x}\right] \leftarrow \text{L-hopital's Rule} \\
&= \exp\left[\lim_{x\to 0}\frac{x\cos x-\sin x}{x\sin^2 x}\right] \\
&= \exp\left[\lim_{x\to 0}\frac{\cos x-x\sin x-\cos x}{\sin^2 x+2x\sin x\cos x}\right] \leftarrow \text{L-hopital's Rule} \\
&= \exp\left[\lim_{x\to 0}\frac{-x}{\sin x+2x\cos x}\right] \\
&= \exp\left[\lim_{x\to 0}\frac{-1}{\cos x+2\cos x-2x\sin x}\right] \leftarrow \text{L-hopital's Rule} \\
&= e^{-1/3}
\end{aligned}
$$

An alternative way is to apply Taylor expansion. Recall that

$$
\begin{aligned}
\sin x &= x-\frac{x^3}{6}+o(x^3) \\
\cos x &= 1-\frac{1}{2}x^2+o(x^2) \\
\ln(1+z) &= z-z^2+o(z^2).
\end{aligned}
$$

Therefore we have

$$\ln(\frac{\sin x}{x}) = \ln(1-\frac{1}{6}x^2+o(x^2)) = \ln(1+x^2(-\frac{1}{6}+o(1)))$$

Finally as $x\to 0$,

$$\frac{\ln(\frac{\sin x}{x})}{1-\cos x} = \frac{x^2(-\frac{1}{6}+o(1))}{x^2(\frac{1}{2}+o(1))} = -\frac{1}{3}$$

■

Plan for next week. Next week, with Taylor theorem, we will discuss taylor's polynomial, and finally taylor series (let $n \to \infty$).

Also, given a function $f \in \mathcal{C}^\infty$, we will discuss topics about:

1. do we always have the taylor series converge in some neighborhood?
2. Suppose it does, does it necessarily converge to $f(x)$?

One counter-example is the typical function

$$f(x) = \begin{cases} \exp\left[-\frac{1}{x^2}\right], & x \neq 0 \\ 0, & x = 0 \end{cases}.$$

The taylor expansion at $x = 0$ is always zero, and therefore does not converge to $f(x)$.

Hence what conditon could guarantee the correctness of convergence?