0

QUANTIFIERS

In this lecture, we will learn about common sense logic and the use of quantifiers in mathematical statements. We have been using statements that involve the quantifier $\forall$ (called the universal quantifier), $\exists$ (called the existential quantifier) and logical connectives *and/or* since our school days. In the course of Linear Algebra we have seen how we prove a statement by contradiction and how to establish a result in contrapositive form.

Let us look at the following sentences. All the notations are standard and same as used in lectures.

- There exists a proper subspace of $\mathbb{R}$.
- There is a vector space which does not have any basis.

We know that both of the statements above are false. How do we prove that? What is the negation of these statements? If we want to prove that the first statement is false, we have to show that any proper subset of $\mathbb{R}$ is not a subspace. For the second one, we have to show that any vector space has a basis. Look at the difference carefully.

The above two statements can be abstracted as follows. There exists an element in a set X which has some property P. In terms of quantifiers, we write

$$\text{“}\exists\ x \in X\ (x \text{ has property } P)\text{”}$$

To prove that this sentence is false means that we negate the sentence. The negation of the above statement is

$$\text{“}\forall\ x \in X\ (x \text{ does not have property } P)\text{”}$$

Let us come back to the two statements mentioned in the beginning of this note.

- There exists a proper subspace of $\mathbb{R}$. In terms of quantifiers

  “$\exists\ A \subset \mathbb{R}$ (A is a subspace)”

  The negation of the above statement is

  “$\forall\ A \subset \mathbb{R}$ (A is not a subspace)”
- There is a vector space which does not have any basis. In terms of quantifiers, we have the following: The negation of

  “$\exists$ a vector space V (V does not have a basis)”

  is “$\forall$ vector spaces V (V has a basis)”

The following remark is important.

Remark 1. *Remember that when you have to prove statements like “$\forall\ A \subset \mathbb{R}$ (A is not a subspace)” or “$\forall$ vector space V (V has a basis)”, you cannot do it by taking* *a particular example. In the former case, you have take an arbitrary subset of $\mathbb{R}$ and prove show that it's not a subspace. In the later case, you have to take an arbitrary vector space (finite or infinite dimensional!) V and prove that it has a basis.*

The above sentences use the existential quantifier $\exists$. Now let us look at sentences which use the universal quantifier $\forall$:

- Every 2×2 real matrix has zero determinant.
- Every subset of $\mathbb{R}^2$ is linearly independent.

In abstract form, the above sentences follow the pattern:

$$\text{“}\forall\ x \in X\ (x \text{ has property } P)\text{”}$$

The negation of this sentence is

$$\text{“}\exists\ x \in X\ (x \text{ does not have property } P)\text{”}$$

In terms of quantifiers, the above two statements can be written as

- "$\forall\ A \in M_2(\mathbb{R})\ (|A| = 0)$". Its negation is
  "$\exists\ A \in M_2(\mathbb{R})\ (|A| \neq 0)$"
- "$\forall\ A \subset \mathbb{R}^2\ (A$ is L.I.)". Its negation is
  "$\exists\ A \subset \mathbb{R}^2\ (A$ is L.D.)"

Remark 2. *If you have to prove the statements "$\exists\ A \in M_2(\mathbb{R})\ (|A| \neq 0)$" or "$\exists\ A \subset \mathbb{R}^2$ (A is L.D.)", what you have to do is to give an example of a 2×2 matrix whose determinant is not zero in the first one, and in the second one, you should give an example of a subset of $\mathbb{R}^2$ which is L.D.*

We now consider sentences which are combined using the connectives *and* or *or*.

Let G be a nonempty set with a binary operation $*$. Consider the following sentences.

- (Existence of identity) There exists an element $e \in G$ such that $a * e = e * a = a$ for all $a \in G$. In terms of quantifiers we write
  "$\exists\ e \in G\ (\forall\ a \in G\ (a * e = e * a = a))$". Note that if we want to claim that this nested sentence is false we have to negate it. Its negation reads as
  "$\forall\ e \in G\ (\exists\ a \in G\ (a * e \neq e * a$ or $a * e \neq a$ or $e * a \neq a)))$".
- (Existence of inverse) For each element $a \in G$, there exists $b \in G$ such that $a * b = b * a = e$, where the e is the identity element. In terms of quantifiers we have
  "$\forall\ a \in G\ (\exists\ b \in G\ (a * b = b * a = e)))$". Its negation is
  "$\exists\ a \in G\ (\forall\ a \in G\ (a * b \neq b * a$ or $a * b \neq e$ or $b * a \neq e)))$".

Before proceeding further we just add a remark regarding the previous two examples. Let $x, y, z \in X$, where X in any set. Consider the statement

"$x = y = z$".

This is actually a combination of three statements,

$x = y$ and $y = z$ and $x = z$.

When we say that these combined statements are not true, it means that atleast one of them should be false. That is, $x \neq y$ or $y \neq z$ or $x \neq z$. Thus we arrive at the following: If a statement is of the form P and Q and R, then its negation is P is false or Q is false or R is false, that is, Not P or Not Q or Not R.

Exercise 3. *Let A, B, C be subsets of a set X. Try to write the negation of the following sentences.*

(1) $x \in A \cap B$.
(2) $x \in A \cup B$.
(3) $A \neq B$.

Exercise 4. *Negate the following sentence using quantifiers:*

"every road in the city has at least one home in which we can find a man who is wealthy and handsome or educated and generous."

Let us now deal with a more complicated sentence. Suppose there is a garden which is full of trees. We make the following statement.

"In each tree in the garden, there is a branch on which all of the leaves are red.

Try to negate this sentence using your common sense. You'll be in lot of trouble.

Let us turn this statement into a mathematical sentence? We first fix some notations. Let T denotes the set of all trees in the garden. For $t \in T$, let B_t denotes the set of all branches on the tree t. For $b \in B_t$, we let L_b denotes the set of all leaves on the branch b. Using quantifiers, the above sentence reads as

$$\text{“}\forall\ t \in T\ (\exists\ b \in B_t\ (\forall\ l \in L_b\ (l \text{ is red})))\text{”}$$

How do we negate it? We do it layer by layer. So,

$$\text{“}\exists\ t \in T\ (\forall\ b \in B_t\ (\exists\ l \in L_b\ (l \text{ is not red})))\text{”}$$

Now the question is that do such complicated sentence appear in Mathematics. Yes, when will study convergence of a sequence of real numbers.

Exercise 5. *Let A, B, C be three subsets of $\mathbb{R}$. What is the negation of the following statement:*

For every $\varepsilon > 1$, there exists $a \in A$ and $b \in B$ such that for all $c \in C$, $|a - c| < \varepsilon$ and $|b - c| > \varepsilon$.

1

THE REAL NUMBER SYSTEM

In this lecture we shall study some properties of real numbers. We are already familiar with the set of natural numbers (or positive integers) $\mathbb{N} = \{1, 2, 3, \ldots\}$[1]. The set of integers $\mathbb{Z}$[2] consists of positive integers, 0, and negative integers, i.e., $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Quotients of integers are called rational numbers denoted by $\mathbb{Q}$, that is, $\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$.

We can represent integers by points on a straight line (called the number line) by fixing the number 0 and a unit distance. By subdividing the segment between 0 and 1, we can represent the rational numbers $\frac{1}{n}$, where $n \in \mathbb{N}$. Thus, we can use this to represent any rational number by a unique point on the number line. Though rational numbers seems to fill the number line, there are some gaps. For example, the number $\sqrt{2}$ is not a rational number (can be easily proved). Similarly, numbers like π and e are also irrationals (not so easy to prove!). Such numbers are called irrational numbers. The rational numbers and the irrational numbers together constitute the set $\mathbb{R}$.

But what exactly are irrational numbers? Or for that matter real numbers? We haven't actually given the precise (mathematical) definition of real numbers. To define the real numbers, one begins with the set $\mathbb{Q}$ and construct the set $\mathbb{R}$. There are two standard approaches for the construction of $\mathbb{R}$ from $\mathbb{Q}$, one due to Dedekind (through Dedekind cuts) [3], and the other due to Cantor (through Cauchy sequences) [1]. Both Cantor and Dedekind published their construction in 1872. We will not study these constructions in this course because they are little complicated, rather we will study some other useful and important properties of real numbers.

We already know that the set of real numbers $\mathbb{R}$ is a field under usual addition and multiplication. Moreover, under usual addition and scalar multiplication, $\mathbb{R}(\mathbb{R})$ is a vector space. We also know that the set of rational numbers $\mathbb{Q}$ is a subfield of $\mathbb{R}$ and it is also a subspace of $\mathbb{R}(\mathbb{Q})$.

Order Properties

We refer to the vector space and field properties of real numbers as algebraic properties. The set $\mathbb{R}$ contains a subset $\mathbb{R}^+$, called the set of all positive real numbers, satisfying the following properties:

[1] In some texts the natural numbers start at 0 instead of 1. This is just a matter of notational convention. The first evidence we have of zero is from the Sumerian culture in Mesopotamia, some 5000 years ago. For more details see [2].

[2] The letter $\mathbb{Z}$ stands for the German word *Zahlen* for numbers.

(1) Given any $x \in \mathbb{R}$, exactly one of the following statements is true:

$$x \in \mathbb{R}^+; \; x = 0; \; -x \in \mathbb{R}^+.$$

(2) If $x, y \in \mathbb{R}^+$, then $x + y \in \mathbb{R}^+$ and $xy \in \mathbb{R}^+$.

We define an order relation on $\mathbb{R}$ as follows:

For $x, y \in \mathbb{R}$, we define x to be less than y, we write $x < y$, if $y - x \in \mathbb{R}^+$. We also write $y > x$ and say that y is greater than x. It follows that $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$. Using the algebraic properties of $\mathbb{R}$ and the properties (1) and (2) above we can easily prove the following:

(i) Given any $x, y \in \mathbb{R}$, one of the following is true.

$$x < y; \; x = y, \; y < x.$$

This is called the **Law of Trichotomy**.

(ii) If $x, y, z \in \mathbb{R}$ such that $x < y$ and $y < z$, then $x < z$.

(iii) If $x, y, z \in \mathbb{R}$ such that $x < y$, then $x + z < y + z$. Moreover, if $z > 0$, then $xz < yz$, whereas if if $z < 0$, then $xz > yz$.

The notation $x \leq y$ means that either $x < y$ or $x = y$. Likewise, $x \geq y$ means $x > y$ or $x = y$.

Completeness Property

We begin with the following definitions.

Definition 1 (Bounded above and below). *Let S be a nonempty subset of $\mathbb{R}$.*

- *We say that S is* **bounded above** *if there exists $\alpha \in \mathbb{R}$ such that $x \leq \alpha$ for all $x \in S$. Any such α is called an* **upper bound** *of S. Express this definition in terms of quantifiers.*
- *We say that S is* **bounded below** *if there exists $\beta \in \mathbb{R}$ such that $x \geq \beta$ for all $x \in S$. Any such β is called a* **lower bound** *of S.*
- *The set S is said to be* **bounded** *if it is bounded above as well as bounded below; otherwise, S is said to be unbounded.*

Examples.

(1) $\mathbb{N}$ is bounded below, and any real number $\beta \leq 1$ is a lower bound of $\mathbb{N}$.

(2) The set $S = \{\frac{1}{n} : n \in \mathbb{N}\}$ is bounded. Indeed, any real number $\alpha \geq 1$ is an upper bound of S, whereas any real number $\beta \leq 0$ is a lower bound of S.

(3) The set $S = \{x \in \mathbb{Q} : x^2 < 2\}$ is bounded. Here, 2 is an upper bound, while -2 is a lower bound.

Exercise 2. *When do you say a real number is not a lower bound of S? When do you say S is not bounded below? Express these in terms of quantifiers.*

Definition 3 (Supremum and Infimum). *Let S be a nonempty subset of $\mathbb{R}$.*

- *An element $M \in \mathbb{R}$ is a* **supremum** *or a* **least upper bound** *of S if*
  (i) M is an upper bound of S, and
  (ii) If α is an upper bound of S, then $M \leq \alpha$.
  The symbol LUB is a shorthand notation for least upper bound. We denote the supremum of S by $\sup S$ or $\operatorname{lub} S$.
- *An element $m \in \mathbb{R}$ is a* **infimum** *or a* **greatest lower bound** *of S if*
  (i) m is a lower bound of S, and
  (ii) If β is a lower bound of S, then $m \geq \beta$.
  The symbol GLB is a shorthand notation for greatest lower bound. We denote the infimum of S by $\inf S$ or $\operatorname{glb} S$.
- *$\sup S$ and $\inf S$ may belong to the set S. If $\sup S \in S$, it is called the maximum of S, and denoted by $\max S$; likewise, if $\inf S \in S$, it is called the minimum of S, and denoted by $\min S$.*

Exercise 4. *Prove that if S has a supremum, it must be unique. The same assertion holds for infimum as well.*

Problem 5. *Find the supremum and infimum of the set $S = \{x + x^{-1} : x > 0\}$.*

Solution. First observe that if $x > 0$, $x + \frac{1}{x} > x$. Hence, S is not bounded above.

On the other hand, if $x > 0$, $(x-1)^2 \geq 0 \implies x^2 - 2x + 1 \geq 0$. Dividing by x we get, $x - 2 + \frac{1}{x} \geq 0$ or $2 \leq x + \frac{1}{x}$. Thus, 2 is a lower bound of S. Further, note that $2 = 1 + \frac{1}{1} \in S$. This implies that $\inf S = 2$.

The following characterization of the LUB will be very useful.

Proposition 6. *Let S be a nonempty subset of $\mathbb{R}$. Then $M = \sup S$ if and only if*

(1) M is an upper bound of S,

(2) If $\beta < M$, then β is not an upper bound of S.

Proof. The direct part is clear. For the converse part, let α be an upper bound of S. We claim that $M \leq \alpha$. If not, then $M > \alpha$. So, by the hypothesis (2), α is not an upper bound, contradicting our assumption. Therefore, $M \leq \alpha$, and $M = \sup S$. $\square$

Exercise 7. *State and prove an analogous to the above Proposition for GLB.*

It is time to state the most important property of $\mathbb{R}$ which is called the **completeness property** or the **LUB property**. It says that

"Every nonempty subset of $\mathbb{R}$ that is bounded above has a LUB".

Theorem 8. *Let S be a nonempty subset of $\mathbb{R}$ that is bounded below. Then S has an infimum.*

Proof. Let $A = \{\alpha \in \mathbb{R} : \alpha \le x, \ \forall \ x \in S\}$. Since S is bounded below, $A \neq \emptyset$. Let $x \in S$. By definition of the set A, $\beta \le x$ for all $\beta \in A$. Then, x is an upper bound of A. This implies that $\sup A \le x$ (sup A exists because A is bounded above). Thus, $\sup A$ is a lower bound of S. We claim that $\sup A = \inf S$. Indeed, if α is any lower bound of S, then $\alpha \in A$. Therefore, $\alpha \le \sup A$, which completes the proof. $\square$

The following theorem is an important consequence of the completeness property of $\mathbb{R}$.

Theorem 9 (Archimedean property (AP))**.** *Given $x, y \in \mathbb{R}$ with $x > 0$, there exists $n \in \mathbb{N}$ such that $nx > y$.*

Proof. Proof by contradiction. If false, then $nx \le y$ for every $n \in \mathbb{N}$. Thus, y is an upper bound of the set $S = \{nx : n \in \mathbb{N}\}$. Let $M = \sup S$. Now for all n, $(n+1)x \le M$ or $nx \le M - x < M$. Hence, $M - x$ is an upper bound of S, a contradiction. $\square$

Corollary 10. *$\mathbb{N}$ is not bounded above in $\mathbb{R}$, that is, for any $x \in \mathbb{R}$, there exists $n \in \mathbb{N}$ such that $n > x$.*

Example 11. *Find the supremum and infimum of the set $S = \left\{\frac{m}{|m|+n} : m \in \mathbb{Z}, \ n \in \mathbb{N}\right\}$.*

Solution. Observe that $-1 < \frac{m}{|m|+n} < 1$, and for $m \in \mathbb{Z}$, $\frac{m}{|m|+1} \in S$. For $m > 0$, $\frac{m}{m+1}$ approaches 1 (whatever that means! We will see the precise definition in the next lecture), and $m < 0$, $\frac{m}{-m+1}$ approaches -1. Hence, we guess that $\sup S = 1$ and $\inf S = -1$. Let's prove this.

If $\beta < 1$, $1 - \beta > 0$. By AP (take $x = 1 - \beta, \ y = \beta$), $\exists \ n \in \mathbb{N}$ such that $n(1-\beta) > \beta$ $\Rightarrow \beta < \frac{n}{n+1} \in S$. Thus, $\sup S = 1$.

If $\alpha > -1$, $1+\alpha > 0$. By AP (take $x = 1+\alpha, \ y = -\alpha$), $\exists \ n \in \mathbb{N}$ such that $n(1+\alpha) > -\alpha$ $\Rightarrow \frac{-n}{n+1} < \alpha$. Hence, $\inf S = -1$.

References

[1] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, New York, 1965.

[2] R. Kaplan, *The Nothing That Is: A Natural History of Zero*, Oxford University Press, New York, 2000.

[3] W. Rudin, *Principles of Mathematical Analysis*, third ed., McGraw-Hill, New York-Auckland-Düsseldorf, 1976.

2

SEQUENCES AND THEIR CONVERGENCE

Definitions and Examples

We start with the definition of a sequence.

Definition 1. *Let X be nonempty set. A sequence in X is a function $f : \mathbb{N} \to X$. We let $x_n = f(n)$ and call x_n the n-th term of the sequence. Generally we denote f by (x_n) or as an infinite tuple $(x_1, x_2, \ldots, x_n, \ldots)$. A sequence in $\mathbb{R}$ is called a real sequence. Likewise, a sequence in $\mathbb{C}$ is called a complex sequence.*

Throughout this course, unless stated otherwise, (x_n) will always denote a real sequence.

Examples. Some examples of sequences:

(1) Fix $c \in \mathbb{R}$ and define $x_n = c$ for all n. The sequence $(c, c, c, \ldots)$ is called a constant sequence.

(2) $(n) = (1, 2, 3, \ldots)$.

(3) $\left(\frac{1}{n}\right) = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right)$.

(4) $((-1)^n) = (-1, 1, -1, 1, \ldots)$.

(5) Let $x_1 = x_2 = 1$ and define $x_n = x_{n-1} + x_{n-2}$ for all $n > 2$. The sequence is $(1, 1, 2, 3, 5, 8, 13, \ldots)$. This sequence is called Fibonacci sequence[1].

The sequence

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \cdots$$

getting closer and closer to the number 0. We say that this sequence converges to 0 or that the limit of the sequence is the number 0. How should this idea be properly defined? The definition which we learned in school was something like this

A sequence x_n converges to a number ℓ if the terms of the sequence get closer and closer to ℓ.

The above definition is not precise. It is too vague, and sometimes misleading. What about the sequence

$$0.1, 0.01, 0.02, 0.001, 0.002, 0.0001, 0.0002, \ldots?$$

This sequence should converge to 0 but the terms do not get steadily "closer and closer" but back off a bit at each second step. Also, the sequence

$$0.1, 0.11, 0.111, 0.1111, 0.1111, \ldots$$

is getting "closer and closer" to 0.2, but the sequence does not converge to 0.2. A smaller number ($\frac{1}{9}$, which it is also getting closer and closer to) is the correct limit.

[1]Also known as Fibonacci numbers. They appear in nature surprisingly often, for example, numbers of petals on a flower, number of spirals on a sunflower or a pineapple are typically Fibonacci numbers. The ratios of successive terms of the Fibonacci sequence tends to an irrational number called the golden ratio. The golden ratio appears in some patterns in nature, including the spiral arrangement of leaves.

Definition 2 (Limit of a Sequence). *Let (x_n) be a real sequence. We say that (x_n) converges if there exists a real number ℓ such that for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that*

$$|x_n - \ell| < \varepsilon \text{ for all } n \geq N.$$

In this case, (x_n) converges to ℓ. The number ℓ is called a limit of the sequence (x_n). We write $x_n \to \ell$ as $n \to \infty$ or $\lim_{n\to\infty} x_n = \ell$. A sequence that is not convergent is said to be divergent.

The definition of convergence of (x_n) can be written in terms of quantifiers as follows:

$$\text{“}\exists\, \ell \in \mathbb{R}\ (\forall\, \varepsilon > 0\ (\exists\, N \in \mathbb{N}\ (\forall\, n \geq N\ (|x_n - \ell| < \varepsilon))))\text{”}.$$

Exercise 3. *What does it mean to say that a sequence (x_n) does not converge? Write it in words and then in terms of quantifiers.*

Examples. (Convergent sequences)

(1) If $x_n = c$ for all n, then $x_n \to c$. In fact, given $\varepsilon > 0$, we may take $N = 1$, so $|x_n - c| = 0$ for all $n \geq 1 = N$.

(2) If $x_n = \frac{1}{n}$ for $n \in \mathbb{N}$, then it is easy to see that the sequence should have limit $\ell = 0$. We claim that $x_n \to 0$. To see this, given any $\varepsilon > 0$, $|x_n - \ell| = |\frac{1}{n} - 0| = \frac{1}{n}$. We want to choose an N such that for all $n \geq N$, $|x_n - \ell| < \varepsilon$, i.e., $\frac{1}{n} < \varepsilon$ or $n > \frac{1}{\varepsilon}$. Such an n exists by the Archimedean property (How?).

Thus, choose an integer N such that $N > \frac{1}{\varepsilon}$ by the Archimedean property. Then for all $n \geq N$, we have

$$|x_n - \ell| = |\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Hence, $\frac{1}{n} \to 0$.

(3) Consider $x_n = \frac{1}{2^n}$. If we look at the terms of the sequence, we can see that this sequence should converge to 0.

Let $\varepsilon > 0$. We have to find $N \in \mathbb{N}$ such that for all $n \geq N$

$$|x_n - \ell| = |\frac{1}{2^n} - 0| = \frac{1}{2^n} < \varepsilon.$$

Note that for all $n \in \mathbb{N}$, $2^n > n$ (Prove by induction!). Hence, $\frac{1}{2^n} < \frac{1}{n}$ for all $n \in \mathbb{N}$. Choose an integer N such that $N > \frac{1}{\varepsilon}$ by the AP, then for all $n \geq N$, we have

$$|x_n - \ell| = |\frac{1}{2^n} - 0| = \frac{1}{2^n} < \frac{1}{n} < \varepsilon.$$

Remark 4. *(1) If $|x_n - \ell| < \varepsilon$ for all $n \geq N$, then any $N_1 > N$ will also work. Thus, N is not unique.*

(2) From the the above examples we must have observed that the natural number N depends on the given $\varepsilon > 0$ while checking for convergence. That is why when we want to emphasize this, we sometime denote N by $N(\varepsilon)$.

Examples. (Divergent sequences)

(1) If $x_n = n$ for $n \in \mathbb{N}$, then it is not difficult to see intuitively that this (x_n) diverges. On the contrary, let $x_n \to \ell$. For $\varepsilon = 1$, $\exists\ N \in \mathbb{N}$ such that $|x_n - \ell| < 1$ for all $n \geq N$. In particular, for $n \geq N$, $x_n = n \in (\ell - 1, \ell + 1)$. This implies that $n < \ell + 1$ for all $n \in \mathbb{N}$, and hence $\mathbb{N}$ is bounded above. This is a contradiction. Therefore, (x_n) is divergent.

(2) If $x_n = (-1)^n$ for $n \in \mathbb{N}$, then (x_n) is divergent. Suppose $x_n \to \ell$. Choose $\varepsilon > 0$ such that $\varepsilon < 1$. Then there exists $N \in \mathbb{N}$ such that $x_n \in (\ell - \varepsilon, \ell + \varepsilon)$ for all $n \geq N$. In particular, $-1, 1 \in (\ell - \varepsilon, \ell + \varepsilon)$ $(x_{2N} = 1, x_{2N+1} = -1)$. Since, $1 < \ell + \varepsilon$ and $-1 > \ell - \varepsilon$, we have

$$2 < \ell + \varepsilon - (\ell - \varepsilon) = 2\varepsilon.$$

That is, $1 < \varepsilon$. This is a contradiction. Since ℓ is arbitrary, (x_n) is divergent.

Proposition 5 (Uniqueness of limit). *A convergent sequence has a unique limit.*

Proof. Exercise. □

Definition 6 (Bounded Sequence). *A sequence(x_n) is said to be bounded if there exists $C > 0$ such that $|x_n| \leq C$ for all $n \in \mathbb{N}$.*

In Example 2, the sequences in items $(1), (3)$ and (4) are bounded. The sequences in items (2) and (5) are unbounded.

Theorem 7 (Necessary condition for convergence). *A convergent sequence is bounded.*

Proof. Let $x_n \to \ell$ and $\varepsilon = 1$. There exists $N \in \mathbb{N}$ such that, for $n \geq N$, we have $|x_n - \ell| < 1$. This implies that

$$|x_n| \leq |x_n - \ell| + |\ell| < 1 + |\ell|, \text{ for } n \geq N.$$

If $C = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, 1 + |\ell|\}$, then $|x_n| \leq C$, and hence, (x_n) is bounded. □

Is the converse of the above proposition true? Consider the sequence $((-1)^n)$. It is bounded but not convergent.

Limit Theorems

In general, verifying the convergence directly from the definition is a difficult task. In this section, we will study few results which will enable us to find limits of many sequences, and some sufficient conditions for the convergence of a sequence. Given two sequences (x_n) and (y_n), the product sequence, denoted by $(x_n y_n)$, is a new sequence (t_n) such that $t_n = x_n y_n$.

Theorem 8 (Algebra of Convergent Sequences). *Let $x_n \to x$, $y_n \to y$ and $\alpha \in \mathbb{R}$. Then*

(1) $x_n + y_n \to x + y$.
(2) $\alpha x_n \to \alpha x$.
(3) $x_n y_n \to xy$.
(4) If $x \neq 0$ and $x_n \neq 0$ for all n, then $\frac{1}{x_n} \to \frac{1}{x}$.

The following theorem is immediate now.

Theorem 9. *The set* c *of all convergent sequences of real numbers is a vector space over* $\mathbb{R}$. *Moreover, The map* $T : \mathrm{c} \to \mathbb{R}$ *defined by*

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

is a linear transformation.

Example 10. *Let* $x_n = \frac{1}{1^2+1} + \frac{1}{2^2+2} + \cdots + \frac{1}{n^2+n}$. *As* $\frac{1}{n^2+n} = \frac{1}{n} - \frac{1}{n+1}$, *we have*

$$x_n = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{n} - \frac{1}{n+1} \to 1.$$

An easy and a very useful result is the following

Theorem 11 (Sandwich Theorem). *Let* (x_n), (y_n) *and* (z_n) *be sequences such that* $x_n \to \alpha$, $y_n \to \alpha$ *and* $x_n \le y_n \le z_n$ *for all* n. *Then* $z_n \to \alpha$.

Proof. Let $\varepsilon > 0$ be given. Since $x_n \to \alpha$ and $y_n \to \alpha$, there exist N_1 and N_2 such that

$$x_n \in (\alpha - \varepsilon, \alpha + \varepsilon) \text{ for all } n \ge N_1, \text{ and } y_n \in (\alpha - \varepsilon, \alpha + \varepsilon) \text{ for all } n \ge N_2.$$

If we let $N = \max\{N_1, N_2\}$, then for $n \ge N$, we have

$$\alpha - \varepsilon < x_n \le z_n \text{ and } z_n \le y_n < \alpha + \varepsilon.$$

That is, $z_n \in (\alpha - \varepsilon, \alpha + \varepsilon)$ for all $n \ge N$. It follows that $z_n \to \alpha$. □

Examples.

- Let $x \in \mathbb{R}$. For each $n \in \mathbb{N}$, choose any element x_n such that $x - \frac{1}{n} < x_n < x + \frac{1}{n}$) (Why such x_n exists!). It follows from the algebra of convergent sequences and sandwich theorem that $x_n \to x$.
- We have $\frac{\sin n}{n} \to 0$, as $-\frac{1}{n} \le \frac{\sin n}{n} \le \frac{1}{n}$.
- Let $x_n = \frac{1}{n^2+1} + \frac{2}{n^2+2} + \cdots + \frac{n}{n^2+n}$. Try to write down x_1, x_2 and x_3.
  Since $\frac{1}{n^2+n} \le \frac{1}{n^2+k} \le \frac{1}{n^2+1}$, $k = 1, 2, \ldots n$, we have

$$(1 + 2 + \cdots + n)\frac{1}{n^2+n} \le x_n \le (1 + 2 + \cdots + n)\frac{1}{n^2+1} \Rightarrow x_n \to \frac{1}{2}.$$

Some Important Limits

In this section, we will study some of the most often used sequences and their convergence.

Theorem 12. *(1) Let* $0 \le r < 1$. *Then* $r^n \to 0$.

(2) Let $|r| < 1$. *Then* $r^n \to 0$.

(3) Let $|r| < 1$. *Then* $nr^n \to 0$.

(4) Let $a > 0$. *Then* $a^{1/n} \to 1$.

(5) $n^{1/n} \to 1$.

(6) Fix $a \in \mathbb{R}$. *Then* $\frac{a^n}{n!} \to 0$.

Proof. (1) If $r = 0$, the result is obvious. If $0 < r < 1$, then $r = \frac{1}{h}$ for some $h > 0$. Using binomial theorem, we have

$$(1+h)^n = 1 + nh + \frac{n(n-1)}{2}h^2 + \cdots + h^n > nh,$$

since all terms are positive. This implies that $0 < r^n < \frac{1}{nh}$ for all n. By Sandwich theorem, $r^n \to 0$.

(2) Use part (1) and the fact that a sequence $\boxed{x_n \to 0 \iff |x_n| \to 0.}$

(3) It is enough to prove the result for $0 < r < 1$. Proceeding as in part (1),

$$(1+h)^n = 1 + nh + \frac{n(n-1)}{2}h^2 + \cdots + h^n > \frac{n(n-1)}{2}h^2,$$

since all terms are positive. Thus, $0 < nr^n < \frac{2}{h^2(n-1)}$ for all $n \geq 2$, and hence by Sandwich theorem, $nr^n \to 0$.

(4) If $a > 1$, write $a^{1/n} = 1 + h_n$, $h_n > 0$. This implies $a = (1+h_n)^n > nh_n$, and hence $0 < h_n < \frac{a}{n}$. This means $h_n \to 0$. Therefore, $a^{1/n} \to 1$ as desired.

If $0 < a < 1$, take $b = \frac{1}{a} > 1$.

(5) Assertions (5) and (6) are exercises.

□

Theorem 13. *Let $x_n \to x$. Let (s_n) be the sequence of arithmetic means defined by*

$$s_n = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Then $s_n \to x$.

Definition 14 (Sequences Diverging to $\pm\infty$)**.** *Let (x_n) be a sequence.*

- *We say that (x_n) diverges to $+\infty$ (or simply ∞), and write $\lim_{n\to\infty} x_n = \infty$ or $x_n \to \infty$ as $n \to \infty$,*

  if for any $r \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $x_n > r$ whenever $n \geq N$.
- *Likewise, (x_n) diverges to $-\infty$, and write $\lim_{n\to\infty} x_n = -\infty$ or $x_n \to -\infty$ as $n \to \infty$,*

  if for any $s \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $x_n < r$ whenever $n \geq N$.

Examples.

- Let $x_n = n$ and $y_n = 2^n$. Then both $x_n \to \infty$ and $y_n \to \infty$. (Prove!)
- Let $x_{2n-1} = 1$ and $x_{2n} = 2n$. This sequence is unbounded, and hence it is divergent (why!). However, it does not diverge to ∞. Try to write this in terms of quantifiers.

  If $r = 2$, then given any $N \in \mathbb{N}$, $2N - 1 \geq N$ and $x_{2N-1} < r$. Thus, $x_n \nrightarrow \infty$.
- $(n!)^{1/n}$ diverges to ∞. To see this, let $r > 0$ be given. Since $\frac{r^n}{n!} \to 0$, for $\varepsilon = 1$, there exists $N \in \mathbb{N}$ such that $\frac{r^n}{n!} < 1$ whenever $n \geq N$. That is, $r^n < n!$ or $(n!)^{1/n} > r$ for $n \geq N$. Therefore, $(n!)^{1/n} \to \infty$.

The following result, called ratio test for sequences, is very useful.

Theorem 15 (Ratio Test)**.** *Let (x_n) be a sequence such that $x_n > 0$ for all n. Let $\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = \lambda$. Then*

(1) If $\lambda < 1$*, then* $\lim_{n\to\infty} x_n = 0$.

(2) If $\lambda > 1$*, then* $\lim_{n\to\infty} x_n = \infty$.

Examples.

- Let $x_n = \frac{a^n}{n!}$, where $a \in \mathbb{R}$. Then $\frac{x_{n+1}}{x_n} = \frac{a}{n+1} \to 0$. It follows that $x_n \to 0$.
- Let $x_n = \frac{a^n}{n}$, where $a > 1$. Then $\frac{x_{n+1}}{x_n} = a$. It follows that $x_n \to \infty$.
- If $\lambda = 1$ in the previous theorem, the sequence (x_n) may converge or diverge. For example, $x_n = \frac{1}{n}$ and $x_n = n$.

Monotone Sequences

In this section we present a sufficient condition for the convergence of a sequence.

Definition 16. *We say that a sequence* (x_n) *is increasing (strictly increasing) if* $x_n \leq x_{n+1}$ $(x_n < x_{n+1})$ *for all* n*, and decreasing (strictly decreasing) if* $x_n \geq x_{n+1}$ $(x_n > x_{n+1})$ *for all* n*. A sequence* (x_n) *is said to be monotone if it is either increasing or decreasing.*

Note that any increasing sequence is bounded below by x_1, and a decreasing sequence is bounded above by x_1. Therefore, an increasing (decreasing) sequence is bounded if and only if it is bounded above (below).

Theorem 17 (Sufficient condition for convergence)**.** *(1) If a sequence* (x_n) *is increasing and bounded above, then it is convergent and*

$$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}.$$

(2) If a sequence (x_n) *is decreasing and bounded below, then it is convergent and*

$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

Example. Let $a > 0$ and $x_1 > 0$. Define $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$ for all $n \in \mathbb{N}$. We show that this sequence is bounded below and decreasing, hence convergent.

By AM-GM inequality,

$$x_{n+1} = \frac{x_n + \frac{a}{x_n}}{2} \geq \sqrt{a}. \text{ Further, } x_{n+1} - x_n = \frac{1}{2}\left(\frac{a - x_n^2}{x_n}\right) \leq 0.$$

Therefore, the sequence (x_n) is bounded below and decreasing. Can we find the limit?

Let $x_n \to \ell$. By the algebra of converging sequences, $\ell = \frac{1}{2}(\ell + \frac{a}{\ell})$ or $\ell^2 = a$. This means $\ell = \sqrt{a}$.

Problem 18. *Let* (x_n) *be bounded. Assume that* $x_{n+1} \geq x_n - 2^{-n}$*. Show that* (x_n) *is convergent.*

Solution. Since (x_n) be bounded, it is enough to show that sequence it is monotone. Let $y_n = x_n - \frac{1}{2^{n-1}}$. It is clear that (y_n) is bounded (How!). Moreover,

$$y_{n+1} - y_n = x_{n+1} - \frac{1}{2^n} - x_n + \frac{1}{2^{n-1}} \geq 0.$$

Hence, (y_n) is increasing. As (y_n) is bounded, it is convergent. This implies that (x_n) is convergent.

3 CAUCHY SEQUENCES AND SUBSEQUENCES

Cauchy Sequence

Let $x_n \to x$. Then for $\varepsilon > 0$, there exists a positive integer N such that $|x_n - x| < \varepsilon/2$ whenever $n \geq N$. Hence, for $n, m \geq N$ we have

$$|x_n - x_m| = |x_n - x + x - x_m| \leq |x_n - x| + |x_m - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

This means that the terms of a converging sequence (x_n) get close (in sense of distance) to each other as m, n increases. Such a sequence is called a Cauchy sequence.

Definition 1 (Cauchy sequence)**.** *A sequence (x_n) is said to be Cauchy if for each $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$, we have $|x_n - x_m| < \varepsilon$.*

It is clear that any convergent sequence is Cauchy (How!). What about the converse? It turns out that the converse is also true. An easy proof of this fact uses the concept of subsequences, and a fundamental result known as *Bolzano-Weierstrass theorem.* But before everything let us see the following remark.

Remark 2. *If (x_n) is Cauchy, then $x_{n+1} - x_n \to 0$ as $n \to \infty$. The converse, however, does not hold. For example, if $x_n = \sqrt{n}$, then*

$$x_{n+1} - x_n = \sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \to 0,$$

but (x_n) is not Cauchy. How do we prove that? Write in terms of quantifiers! Indeed, for $\varepsilon = 1$, and any $n_0 \in \mathbb{N}$, we have

$$|x_{4n_0} - x_{n_0}| = \sqrt{4n_0} - \sqrt{n_0} = \sqrt{n_0} \geq 1.$$

Lemma 3. *Any Cauchy sequence is bounded.*

Proof. Let (x_n) be Cauchy and $\varepsilon = 1$. Then there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$, we have $|x_n - x_m| < 1$. Take $m = N$. Then for $n \geq N$, $|x_n - x_N| < 1$. It follows that $|x_n| \leq |x_n - x_N| + |x_N| < 1 + |x_N|$ for all $n \geq N$. If we let $C = \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, |x_N|\}$, then $|x_n| \leq C$ for all $n \in \mathbb{N}$. □

The following result gives a useful sufficient condition for a sequence to be Cauchy.

Proposition 4. *Suppose $0 < \alpha < 1$ and (x_n) is a sequence satisfying the contractive condition:*

$$\boxed{|x_{n+2} - x_{n+1}| \leq \alpha |x_{n+1} - x_n|, \quad \text{for all } n \in \mathbb{N}.}$$

Then (x_n) is a Cauchy sequence.

Example 5. *Let $x_1 = 1$ and $x_{n+1} = 1 + \frac{1}{x_n}$ for $n > 1$. We note that*

$$|x_{n+2}-x_{n+1}| = |1+\frac{1}{x_{n+1}}-1-\frac{1}{x_n}| = \left|\frac{x_{n+1}-x_n}{x_{n+1}x_n}\right| \text{ and } |x_{n+1}x_n| = \left|\left(1+\frac{1}{x_n}\right)x_n\right| = |x_n+1| \geq 2.$$

This implies that $|x_{n+2} - x_{n+1}| \leq \frac{1}{2}|x_{n+1} - x_n|$. Hence (x_n) is Cauchy.

Subsequences

Definition 6. *Let $f : \mathbb{N} \to \mathbb{R}$ be a sequence. Then a subsequence is the restriction of f to an infinite subset S of $\mathbb{N}$.*

Remark 7. *(1) (**Well-Ordering Property**) Every nonempty subset S of $\mathbb{N}$ has a least element, i.e., there exists $\ell \in S$ such that $\ell \leq x$ for all $x \in S$.*

(2) Using item (1), one can prove that an infinite subset S of $\mathbb{N}$ can be listed as

$$\{n_1 < n_2 < \cdots < n_k < n_{k+1} < \cdots\}.$$

(3) In view of item (2), it is a standard practice to denote a subsequence of (x_n) as (x_{n_k}) where $n_1 < n_2 < \cdots < n_k < n_{k+1} < \cdots$. Note that $n_k \to \infty$ as $k \to \infty$.

(4) **A very useful observation:** $n_k \geq k$ **for all** k.

(5) A subsequence is formed by deleting some of the terms of the sequence and retaining the remaining terms in the same order. For example, if (x_n) is a sequence, then $(x_1, x_3, \ldots, x_{2k-1}, \ldots)$ and $(x_2, x_4, \ldots, x_{2k}, \ldots)$ are subsequences of (x_n). In the former case, $n_k = 2k - 1$, and in the later, $n_k = 2k$ (k varies from 1 to ∞).

(6) $(-1, -1, \ldots, -1, \ldots)$ and $(1, 1, \ldots, 1, \ldots)$ are both subsequences of $((-1)^n)$. From this example, it is clear that a divergent sequence may have convergent subsequences. Does this happen all the time?

(7) $(\frac{1}{k^3})$ and $(\frac{1}{2^k})$ are subsequences of $(\frac{1}{n})$. What is n_k in both cases?

Lemma 8. *A sequence (x_n) converges to x if and only if every subsequence of (x_n) converges to x.*

Proof. Exercise. □

A remarkable fact about monotonic subsequences is the following.

Proposition 9. *Every sequence has a monotonic subsequence.*

Lemma 8 says that if a sequence is convergent, then all its subsequences converge to the same limit. What happens if a sequence is divergent? Does it have a convergent subsequence? The next theorem, which is called the Bolzano-Weierstrass Theorem,[1] answers this question.

Theorem 10. *Every bounded sequence has a convergent subsequence.*

Proof. Let (x_n) be a bounded sequence. By Proposition 9, (x_n) has a monotone subsequence (x_{n_k}). Since (x_n) is bounded, so is its subsequence (x_{n_k}). By Theorem 23 of Lecture 3, (x_{n_k}) is convergent. □

Now we state a necessary and sufficient condition for convergence.

Theorem 11 (Cauchy's criterion)**.** *A sequence is convergent if and only if it is Cauchy.*

Proof. ($\Longrightarrow$) Already done.

($\Longleftarrow$) Let (x_n) be a Cauchy sequence. By Lemma 3, (x_n) is bounded. Bolzano-Weierstrass Theorem implies the existence of a converging subsequence (x_{n_k}). Let $x_{n_k} \to x$. So, we have a candidate for the limit of (x_n). We claim that $x_n \to x$.

Let $\varepsilon > 0$. Since $x_{n_k} \to x$, there exists N_1 such that

$$|x_{n_k} - x| < \varepsilon/2, \quad \text{for all } k \geq N_1. \tag{1}$$

For the same ε, since (x_n) is Cauchy, there exists N_2 such that

$$|x_n - x_m| < \varepsilon/2, \quad \text{for all } n, m \geq N_2. \tag{2}$$

Let $N = \max\{N_1, N_2\}$. Fix an $k \geq N$, and let $m = n_k$. It follows from Equation (1) that

$|x_m - x| \leq \varepsilon/2$, because $m = n_k \geq k \geq N \geq N_1$ (Check this!).

On the other hand, from Equation (2) it follows that

$|x_n - x_m| < \varepsilon/2$, whenever $n \geq N$ (Why $m \geq N_2$!). Everything is ready now! For $n \geq N$, we have

$$|x_n - x| \leq |x_n - x_m| + |x_m - x| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

This completes the proof. □

Cauchy's criterion is also referred as **Cauchy completeness of** $\mathbb{R}$.

[1] This theorem is considered as one of the most important tools in calculus. It is named after two great nineteenth-century mathematicians, Bernard Bolzano and Karl Weierstrass. These two mathematicians, the first German and the second Czech, rank with Cauchy among the founders of our subject.

4

CONTINUITY

In the previous lectures we studied real sequences, that is, real-valued functions defined on the subset $\mathbb{N}$ of $\mathbb{R}$. In this lecture we will consider real-valued functions whose domains are arbitrary subsets of $\mathbb{R}$.

Continuity of Functions

We first introduce the notion of continuity and see few examples.

Definition 1 ($\varepsilon - \delta$ Definition[1]). *Let $J \subset \mathbb{R}$. Let $f : J \to \mathbb{R}$ and $a \in J$. We say that f is continuous at a if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ whenever $x \in J$ and $|x - a| < \delta$.*

If f is continuous at each $a \in J$, then we say that f is continuous on J.

Remark 2. *The basic idea to show the continuity of f at a point a is to obtain an estimate of the form*

$$|f(x) - f(a)| \leq K_a|x - a|,$$

where K_a is a constant which may depend on a. This may not work always.

Examples.

- Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x$. We want to prove that f is continuous on $\mathbb{R}$.

  To check the continuity at any point $a \in \mathbb{R}$, we need to estimate $|f(x) - f(a)| = |x - a|$. If $\varepsilon > 0$ is given, we wish to find a $\delta > 0$ such that If $|x - a| < \delta \implies |f(x) - f(a)| = |x - a| < \varepsilon$. This suggests that we may take $\delta = \varepsilon$. Now, how do you write in the exam!!??

  Let $a \in \mathbb{R}$ and $\varepsilon > 0$. Let $\delta = \varepsilon$. Then for any x with $|x - a| < \delta$, we have $|f(x) - f(a)| = |x - a| < \delta = \varepsilon$. Thus, f is continuous at $a \in \mathbb{R}$. Since a is arbitrary, we conclude that f is continuous on $\mathbb{R}$.

- Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Let $a \in \mathbb{R}$ and $\varepsilon > 0$. We choose a $\delta < \min\{1, \frac{\varepsilon}{1+2|a|}\}$. For $x \in \mathbb{R}$ such that $|x - a| < \delta$, we have $|x - a| < 1$ so that

[1]Epsilon-delta proofs are first found in the works of Cauchy. The formal $\varepsilon - \delta$ definition of continuity is attributed to Bolzano and Weierstrass. For more details see [1].

$|x| \le |x-a| + |a| < 1 + |a|$. Now,

$$\begin{aligned}|f(x) - f(a)| = |x^2 - a^2| = |x+a||x-a| &\le (|x| + |a|)|x-a| \\ &\le (1 + 2|a|)|x-a| \\ &< (1+2|a|)\delta < \varepsilon.\end{aligned}$$

Thus, f is continuous at a and hence on $\mathbb{R}$.

But now you might be wondering: how did we know in advance that such a δ will work. For this we have to go back to Remark 2. We want an estimate of the form $|f(x) - f(a)| \le K_a|x-a|$. So,

$$|f(x) - f(a)| = |x^2 - a^2| = |x+a||x-a| \le (|x| + |a|)|x-a|. \tag{1}$$

We need to estimate $|x|$ in terms of a. Suppose we have already found a δ that works. This implies that $|x| \le |x-a| + |a| < \delta + |a|$. We know that if a δ works, then any $\delta' \le \delta$ also works. Assume $\delta < 1$ (If we find a $\delta > 1$, we can choose the minimum of this δ and 1). It follows that $|x| < 1 + |a|$. Equation (1) now becomes

$$|f(x) - f(a)| \le (|x| + |a|)|x-a| \le (1 + 2|a|)|x-a|.$$

Therefore, if we make sure that $(1+2|a|)|x-a| < \varepsilon$, we get $|f(x) - f(a)| < \varepsilon$. In other words, we have to ensure that $|x-a| < \frac{\varepsilon}{1+2|a|}$. We also wanted $|x| < 1 + |a|$. Thus, we need to take $\delta < \min\{1, \frac{\varepsilon}{1+2|a|}\}$.

Exercise 3. *Using $\varepsilon - \delta$ argument, prove that the following functions are continuous.*

(1) *For $a > 0$, let $f : [-a, a] \to \mathbb{R}$ be defined as $f(x) = x^2$.*
(2) *For $a > 0$, let $f : (a, \infty) \to \mathbb{R}$ be defined as $f(x) = \frac{1}{x}$.*

Sequential Criterion for Continuity

We can characterize continuity of a function using the theory of sequence limits. The next result provides a very useful criterion to check continuity of a function at a point.

Theorem 4. *Let $a \in J$, and let $f : J \to \mathbb{R}$ be a function. Then f is continuous at a if and only if for every sequence (x_n) in J with $x_n \to a$, we have $f(x_n) \to f(a)$.*

Proof. Suppose f is continuous at a and $x_n \to a$. Let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ whenever $|x-a| < \delta$. Since $x_n \to a$, for this $\delta > 0$, there exists a positive integer N such that $|x_n - a| < \delta$ for all $n \ge N$. This N serves our purpose. That is, for all $n \ge N$,

$$|x_n - a| < \delta \Longrightarrow |f(x_n) - f(a)| < \varepsilon \Longrightarrow f(x_n) \to f(a).$$

For the converse, let us assume that f is not continuous at a. Then there exists $\varepsilon > 0$ for which no $\delta > 0$ will satisfy the required $(\varepsilon - \delta)$ condition. For each $\delta = \frac{1}{n}, n \in \mathbb{N}$, there exist x_n such that $|x_n - a| < \frac{1}{n}$ with $|f(x_n) - f(a)| \geq \varepsilon$. This implies that $x_n \to a$ but $f(x_n) \nrightarrow f(a)$. This contradicts our hypothesis. Hence f is continuous at a. □

Example 5.

- *Let $f : \mathbb{R} \to \mathbb{R}$ be given by*

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q}, \\ 0, & \text{otherwise.} \end{cases}$$

This is known as Dirichlet's function. This function is not continuous at any point of $\mathbb{R}$.

Let $a \in \mathbb{Q}$. Choose a sequence of irrational numbers (x_n) converging to a. Then $f(x_n) = 0 \nrightarrow f(a) = 1$. Similarly, if $a \notin \mathbb{Q}$, we choose a sequence of rational numbers (x_n) converging to a. Then $f(x_n) = 1 \nrightarrow f(a) = 0$. So, f is not continuous at any $a \in \mathbb{R}$.

- *Let $f : [0, \infty) \to \mathbb{R}$ be given by*

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 1/q, & \text{if } x = p/q, \text{ where } p, q \in \mathbb{N} \text{ and } p, q \text{ have no common factors,} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

This function is known as Thomae's function. It is discontinuous at every rational in $[0, \infty]$. Let $a \in \mathbb{Q}$. Choose a sequence of irrational numbers (x_n) converging to a. Then $f(x_n) = 0 \nrightarrow f(a)$ as $f(a) \neq 0$. What happens at irrational numbers?

Problem 6. *Let $f : \mathbb{R} \to (0, \infty)$ be a function which satisfies $f(x + y) = f(x)f(y)$ for all $x, y \in \mathbb{R}$. If f is continuous at 0, show that f is continuous at every $a \in \mathbb{R}$.*

Solution. Since $f(0) = f(0)^2$, we have $f(0) = 1$. Moreover, $f(x - x) = f(x)f(-x) \implies f(-x) = \frac{1}{f(x)}$. Let $a \in \mathbb{R}$ and $x_n \to a$. It follows that $x_n - a \to 0$. As f is continuous at 0, $f(x_n - a) \to f(0) = 1$. But $f(x_n - a) = f(x_n)f(-a) = \frac{f(x_n)}{f(a)} \to 1$. This implies that $f(x_n) \to f(a)$ (Why?). Hence f is continuous at a.

References

[1] Judith V. Grabiner, *The Origins of Cauchy's Rigorous Calculus*, United States, Dover Publications, 2012.

5

LIMITS

In lecture 2, we have seen the definition of limit of a sequence. In this lecture, we shall define the concept of a limit of a function at a point in $\mathbb{R}$. For this, we must assume that the domain of the function satisfies certain conditions. Let $J \subseteq \mathbb{R}$ and $a \in \mathbb{R}$ such that J contains $(a-r, a)$ and $(a, a+r)$ for some $r > 0$. That is, the domain J contains an open interval around a except possibly the point a itself. So, in the sequel, when we write $f : J \to \mathbb{R}$, we always mean that J satisfies the above condition.

Definition 1. *Let $f : J \to \mathbb{R}$ be a function. We say that $\lim_{x\to a} f(x)$ exists if there exists $\ell \in \mathbb{R}$ such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\boxed{x \in J \text{ and } 0 < |x-a| < \delta \implies |f(x) - \ell| < \varepsilon.}$$

If $\lim_{x\to a} f(x)$ exists, we say that the limit of the function f as x tends to a (or as $x \to a$) is ℓ and write $\lim_{x\to a} f(x) = \ell$.

Remark 2. *Note that a need not be in the domain of f. Even if $a \in J$, ℓ need not be $f(a)$, and if $\ell = f(a)$, this information is irrelevant to us.*

Proposition 3. *With the notation of Definition 1, the limit ℓ is unique.*

Proof. Exercise. □

Some properties of limits

The proof of the following theorem is similar to the proof of the Theorem 5 in lecture 5.

Theorem 4. *$\lim_{x\to a} f(x) = \ell$ iff for every sequence (x_n) in $J \setminus \{a\}$ with $x_n \to a$, we have $f(x_n) \to \ell$.*

Example 5. *Let $f(x) = \cos\frac{1}{x}$, where $x \neq 0$. We show that $\lim_{x\to 0} f(x)$ does not exist. Indeed, consider $x_n = \frac{1}{\pi n}$ for $n \in \mathbb{N}$. Then $x_n \to 0$ but $f(x_n) = (-1)^n$ does not converge. Are we done here? Convince yourself that we are!*

The following simple lemma is an analogue of the sandwich theorem for the limits of functions.

Lemma 6. *Let f, g, h be defined on J. Assume that $f(x) \le h(x) \le g(x)$ for all $x \in J$, and $\lim_{x\to a} f(x) = \ell = \lim_{x\to a} g(x)$. Then $\lim_{x\to a} h(x) = \ell$.*

Proof. Exercise. □

We now relate the concepts of continuity and limit.

Proposition 7. *Let $f : J \to \mathbb{R}$ be a function and $a \in J$. Then f is continuous at a iff $\lim_{x\to a} f(x)$ exists and is equal to $f(a)$.*

Theorem 8 (Algebra of Limits)**.** *Let $f, g : J \to \mathbb{R}$ be given and $\alpha \in \mathbb{R}$. Assume $\lim_{x\to a} f(x) = \ell$ and $\lim_{x\to a} g(x) = m$. Then*

(1) $\lim_{x\to a}(f+g)(x) = \ell + m$.
(2) $\lim_{x\to a}(\alpha f)(x) = \alpha\ell$.
(3) $\lim_{x\to a}(fg)(x) = \ell m$.
(4) If $\ell \neq 0$, then $f(x) \neq 0$ for all $x \in (a-\delta, a+\delta)$ for some $\delta > 0$, $x \neq a$. Moreover, $\lim_{x\to a}(\frac{1}{f})(x) = \frac{1}{\ell}$.

One-sided and Infinite Limits

We now consider one-sided limits. In order to define one-sided limits such as $\lim_{x\to a^+} f(x)$, we need to restrict x to those $x > a$. That is

Definition 9 (Left and right hand limits)**.**
- $\lim_{x\to a^+} f(x) = \ell$ *exists if for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$x > a,\ \ x \in J \text{ and } 0 < |x-a| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon.$$

- $\lim_{x\to a^-} f(x) = \ell$ *exists if for every $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$x < a,\ \ x \in J \text{ and } 0 < |x-a| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon.$$

The relationship between the one-sided limits and the limit $\lim_{x\to a} f(x)$ is given below.

Theorem 10. $\lim_{x\to a} f(x)$ *exists iff* $\lim_{x\to a^+} f(x)$ *and* $\lim_{x\to a^-} f(x)$ *exist and are equal.*

Definition 11 (Limits at infinity)**.**
- *Let $f : (a, \infty) \to \mathbb{R}$. We say that a limit of f as x tends to infinity exists*

  if there is a real number ℓ such that for every $\varepsilon > 0$, there exists $r \in \mathbb{R}$ such that for $x > r$ we have $|f(x) - \ell| < \varepsilon$.

  We then write $f(x) \to \ell$ as $x \to \infty$ or $\lim_{x\to\infty} f(x) = \ell$.

  In terms of sequences,

  $\lim_{x\to\infty} f(x) = \ell$ *iff for any sequence (x_n) in (a, ∞) such that $x_n \to \infty$, we have $f(x_n) \to \ell$.*
- *Let $f : (-\infty, a) \to \mathbb{R}$. We say that $\lim_{x\to-\infty} f(x) = \ell$ if for every $\varepsilon > 0$, there exists $s \in \mathbb{R}$ such that for $x < s$ we have $|f(x) - \ell| < \varepsilon$.*

Definition 12 (infinite limits).

- *Let $f : J \to \mathbb{R}$. We say that $\lim_{x\to a} f(x) = \infty$ if for every $p > 0$, there exists $\delta > 0$ such that $x \in J$ and $0 < |x-a| < \delta \Longrightarrow f(x) > p$.*
- *Let $f : J \to \mathbb{R}$. We say that $\lim_{x\to a} f(x) = -\infty$ if for every $s < 0$, there exists $\delta > 0$ such that $x \in J$ and $0 < |x-a| < \delta \Longrightarrow f(x) < s$.*

Exercise 13. *Formulate a definition for the statements*

$$\lim_{x\to\infty} f(x) = \infty \quad \textit{and} \quad \lim_{x\to 0^+} f(x) = -\infty.$$

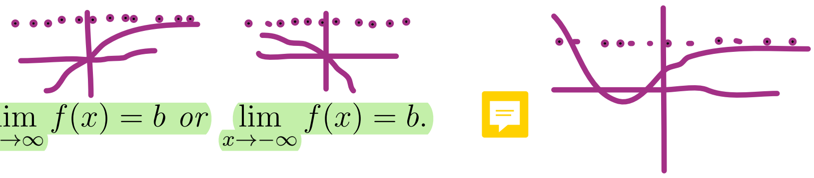
Example 14. $\lim_{x\to 0^+} \frac{1}{x} = \infty$. *Let $r > 0$. Choose $0 < \delta < \frac{1}{r}$. Then for $0 < x < \delta$, we have $f(x) = \frac{1}{x} > r$.*

Asymptotes

The concept of a limit at $\pm\infty$ is useful in considering asymptotes of curves. Roughly speaking, a straight line is considered to be an asymptote of a curve if it comes arbitrarily close to that curve. We use this information to draw a graph of a curve.

Definition 15. *Let $f : J \to \mathbb{R}$ be a function.*

- *A straight line $y = b$, $b \in \mathbb{R}$, is called* **horizontal asymptote** *of the curve $y = f(x)$ if either*



$$\lim_{x\to\infty} f(x) = b \quad or \quad \lim_{x\to-\infty} f(x) = b.$$

- *A straight line $y = ax+b$, where $a, b \in \mathbb{R}$ and $a \neq 0$, is called* **oblique asymptote** *of the curve $y = f(x)$ if either*

$$\lim_{x\to\infty} (f(x) - ax) = b \quad or \quad \lim_{x\to-\infty} (f(x) - ax) = b.$$

- *A straight line $x = c$, $c \in \mathbb{R}$, is called* **vertical asymptote** *of the curve $y = f(x)$ if one or more of the following holds:*

$$f(x) \to \infty \text{ as } x \to c^-,\ f(x) \to -\infty \text{ as } x \to c^-,$$
$$f(x) \to \infty \text{ as } x \to c^+,\ f(x) \to -\infty \text{ as } x \to c^+.$$

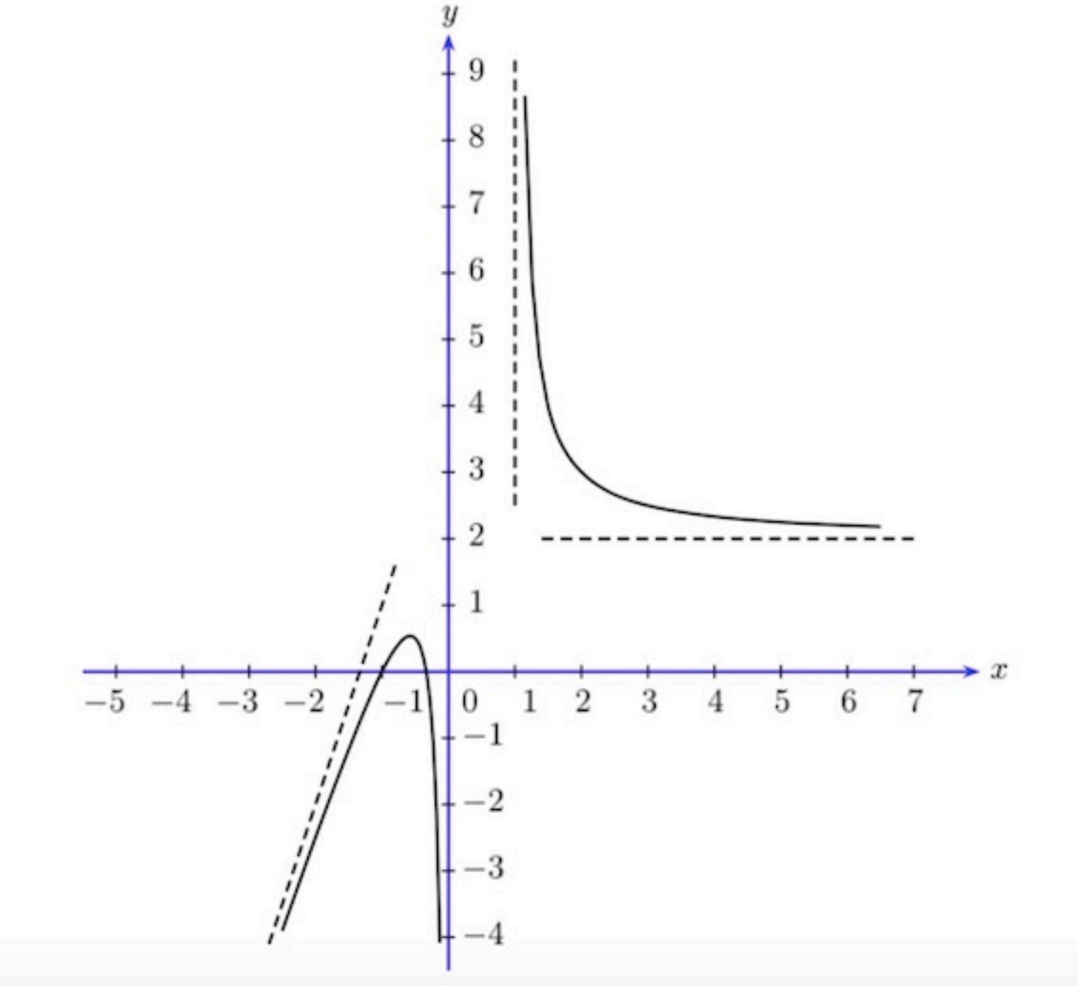
Example. Let $f : (-\infty, 0) \cup (1, \infty) \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} \frac{2x-1}{x-1} & \text{if } x > 1, \\ \frac{3x^2+4x+1}{x} & \text{if } x < 0. \end{cases}$$

For $x > 1$, $f(x) = 2 + \frac{1}{x-1}$, and so $\lim_{x\to\infty} f(x) = 2$. Therefore, the straight line $y = 2$ is a horizontal asymptote of the curve $y = f(x)$.

Also, as $x \to 1^+$, $f(x) \to \infty$. Hence, the straight line $x = 1$ is a vertical asymptote of the curve $y = f(x)$.

For $x < 0$, $f(x) = 3x + 4 + \frac{1}{x}$. Thus, $\lim_{x\to\infty}(f(x) - 3x) = 4$. It follows that the $y = 3x + 4$ is an oblique asymptote of the curve $y = f(x)$. Moreover, $f(x) \to -\infty$ as $x \to 0^-$. Hence, the straight line given by $x = 0$ is a vertical asymptote of the curve $y = f(x)$. The graph of f is shown below.



6

PROPERTIES OF CONTINUOUS FUNCTIONS

In this lecture will study the relations between the continuity of a function and its several geometric properties. Throughout this lecture, J will denote an interval.

1. Continuity and Boundedness

Definition 1. *Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$ be a function.*

(1) *f is said to be bounded above on D if $\exists\, \alpha \in \mathbb{R}$ such that $f(x) \le \alpha$ for all $x \in D$. Any such α is called an upper bound for f.*

(2) *f is said to be bounded below on D if $\exists\, \beta \in \mathbb{R}$ such that $f(x) \ge \beta$ for all $x \in D$. Any such β is called a lower bound for f.*

(3) *f is said to be bounded on D if it is bounded above on D and also bounded below on D.*

Remark 2. (1) *f is bounded if and only if $\exists\, \gamma \in \mathbb{R}$ such that $|f(x)| \le \gamma$ for all $x \in D$. Such a γ is called a bound for the absolute value of f.*

(2) *Geometrically, f is bounded above means that the graph of f lies below some horizontal line, while f is bounded below means that its graph lies above some horizontal line.*

Examples

(1) Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions defined by $f(x) = -x^2$ and $g(x) = x^2$. Then f is bounded above and g is bounded below on $\mathbb{R}$. Neither of these functions is bounded on $\mathbb{R}$.

(2) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \frac{x^2}{x^2+1}$ is bounded on $\mathbb{R}$. We can easily observe that $0 \le f(x) < 1$. Moreover, the bounds 0 and 1 are optimal, i.e.,

$$\inf\{f(x) : x \in \mathbb{R}\} = 0 \text{ and } \sup\{f(x) : x \in \mathbb{R}\} = 1. \text{ (Prove!)}$$

(3) In the above example, we see that the infimum of f is attained, i.e., $\exists\, c\ (= 0) \in \mathbb{R}$ such that $\inf\{f(x) : x \in \mathbb{R}\} = f(c)$. What to you think about the supremum? Is it attained?

Definition 3. *Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$ be a function.*

(1) *f attains its upper bound on D if $\exists\, c \in D$ such that $\sup\{f(x) : x \in \mathbb{R}\} = f(c)$.*

(2) *f attains its lower bound on D if $\exists\, d \in D$ such that $\inf\{f(x) : x \in \mathbb{R}\} = f(d)$.*

(3) *f attains its bounds on D if it attains its upper bound and lower bound on D.*

A bounded function need not be continuous. For example, the Dirichlet's function defined in Lecture 5 is bounded but not continuous. Also, a continuous function need not be bounded. For instance, let $D_1 = [0, \infty)$ and $f_1(x) = x$ for $x \in D_1$, or $D_2 = (0, 1]$ and $f_2(x) = \frac{1}{x}$ for $x \in D_2$. It is clear that both f_1 and f_2 are continuous. The function f_1 is unbounded because its domain D_1 is unbounded. To understand why f_2 is unbounded, we need the following concept.

Definition 4. *Let $D \subseteq \mathbb{R}$. We say that D is a closed set if*

$$(x_n) \text{ any sequence in } D \text{ and } x_n \to x \Rightarrow x \in D.$$

The interval $(0, 1]$ is not a closed set, since $(\frac{1}{n}) \in (0, 1]$ and $\frac{1}{n} \to 0$, but $0 \notin (0, 1]$. The interval $[a, b]$ is closed. To see this, consider any sequence (x_n) in $[a, b]$ such that $x_n \to x$. Since $a \le x_n \le b$ and $x_n \to x$, we have $a \le x \le b$ (Why?). Hence $[a, b]$ is a closed set.

Theorem 5. *Let D be a closed and bounded subset of $\mathbb{R}$, and $f : D \to \mathbb{R}$ be a continuous function. Then f is bounded and attains its bounds on D.*

Proof. Suppose f is not bounded on D. Then for every $n \in \mathbb{N}$, $\exists\ x_n \in D$ such that $|f(x_n)| > n$ (Why?). The sequence (x_n) is bounded since D is bounded. By the Bolzano-Weierstrass Theorem, (x_n) has a convergent subsequence (x_{n_k}). If $x_{n_k} \to x$, then $x \in D$ since D is closed. By the continuity of f, $f(x_{n_k}) \to f(x)$. Being convergent, the sequence $(f(x_{n_k}))$ is bounded. This contradicts the fact that $|f(x_{n_k})| > n_k$ for every $k \in \mathbb{N}$ and $n_k \to \infty$ as $k \to \infty$. Hence, f is bounded.

To show that f attains its bounds, let $m = \inf\{f(x) : x \in D\}$ and $M = \sup\{f(x) : x \in D\}$ (Why m and M exist?). There exists a sequence (x_n) in D such that $f(x_n) \to M$. Since D is bounded, the sequence is bounded and has a convergent subsequence, say (x_{n_k}). Let $x_{n_k} \to x$. Then $x \in D$ because D is closed. By the continuity of f, $f(x_{n_k}) \to f(x)$. This implies that $f(x) = M$ (Why?) Thus, f attains its upper bound. The proof for the lower bound case is similar. $\square$

2. Continuity and Monotonicity

Definition 6. *$f : J \to \mathbb{R}$ be a function.*

(1) f is monotonically increasing (strictly increasing) on J if

$$x, y \in J,\ x < y \Rightarrow f(x) \le f(y) (f(x) < f(y)).$$

(2) f is monotonically decreasing (strictly decreasing) on J if

$$x, y \in J,\ x < y \Rightarrow f(x) \ge f(y) (f(x) > f(y)).$$

(3) f is monotonic (strictly monotonic)) on J if f is either increasing (strictly increasing) or decreasing (strictly decreasing) on J.

We can easily find an example of a function that is monotonic but not continuous. For example, $f(x) = [x]$ for $x \in \mathbb{R}$. (Exercise. Discuss the points of continuity of f). Similarly, a continuous function may not be monotonic, $f(x) = |x|$ for $x \in \mathbb{R}$.

Theorem 7. *Let $f : J \to \mathbb{R}$ be a function that is strictly monotonic on J. Then $f^{-1} : f(J) \to \mathbb{R}$ is continuous.*

Remark 8. *(1) If f is strictly monotonic on J, then f is one-one and its inverse $f^{-1} : f(J) \to \mathbb{R}$ is well defined (Verify!).*

(2) In the above theorem, the function f need not be continuous, and the range of f need not be an interval.

Exercise 9. *Consider the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x + [x]$. Show that f is strictly increasing on $\mathbb{R}$ and f^{-1} is continuous on $f(\mathbb{R})$ even though f is not continuous at any $m \in \mathbb{Z}$.*

3. Continuity and Intermediate Value Property (IVP)

Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a)$ and $f(b)$ are of opposite signs. If you draw the graphs of few such functions, you will see that the graph meets the x-axis.

Theorem 10 (Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a) < \lambda < f(b)$. Then $\exists\ c \in (a, b)$ such that $f(c) = \lambda$.*

3.1. Some Applications of IVP.

(1) Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(x) \neq 0$ for any $x \in [a, b]$. Then either $f > 0$ or $f < 0$.

(2) Let $f : \mathbb{R} \to \mathbb{R}$ be continuous taking values in $\mathbb{Z}$ or in $\mathbb{Q}$. Then f is constant.

(3) Let $f : [a, b] \to \mathbb{R}$ be a non-constant continuous function. Then $f([a, b])$ is an interval. To see this, let $y_1, y_2 \in f([a, b])$. Assume $y_1 < y < y_2$. We need to show that $y \in f([a, b])$. Let $x_1, x_2 \in [a, b]$ such that $f(x_1) = y_1$, $f(x_2) = y_2$. By IVP, $\exists$ x between x_1 and x_2 such that $f(x) = y$.

Theorem 11. *Let $f : J \to \mathbb{R}$ be a one-one continuous function. Then f is strictly monotone.*

7

DIFFERENTIABILITY

The basic idea of differential calculus is to study the local behaviour of a function at a point by its first order (linear) approximation at the same point.

Throughout J will denote an interval, and $c \in J$. Let $f : J \to \mathbb{R}$ be a function. We want to approximate $f(x)$ for x near c. If $E(x) = f(x) - a - b(x-c)$ is the error by taking the value of $f(x)$ as $a + b(x-c)$ near c, we want the error to go to zero much faster than x goes to c. That is, $\lim_{x\to c} \frac{f(x)-a-b(x-c)}{x-c} = 0$. If this is true, then $\lim_{x\to c}(f(x) - a - b(x-c)) = 0$ (Why?). This implies that $\lim_{x\to c} f(x) = a$. If f is continuous at c, then $f(c) = a$. Thus, we can approximate f at c if there exists a real number b such that $\lim_{x\to c} \frac{f(x)-f(c))}{x-c} = b$. If this happens, we say that f is differentiable at c, and denote the real number b by $f'(c)$.

Definition 1. *Let $f : J \to \mathbb{R}$. Then f is said to be differentiable at c if $\exists\ b \in \mathbb{R}$ such that $\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = b$. In this case, the value of the limit, that is, b is denoted by $f'(c)$ and is called the derivative of f at c.*

If f is differentiable at every point of J, we say that f is differentiable on J. In such a case, we obtain a new function from J to $\mathbb{R}$ given by $c \mapsto f'(c)$. This function is denoted by f' and is called the derivative function of f. We sometime denote f' by $\frac{df}{dx}$ or $\frac{dy}{dx}$ when $y = f(x)$. Likewise, $f'(c)$ is often denoted by $\frac{df}{dx}|_{x=c}$ or $\frac{dy}{dx}|_{x=c}$.

Remark 2. *(1) In $\epsilon - \delta$ form, we say that f is differentiable at c if for every $\epsilon > 0$, $\exists$ a $\delta > 0$ such that*

$$x \in J,\ 0 < |x-c| < \delta \Rightarrow |f(x) - f(c) - b(x-c)| < \epsilon|x-c|.$$

(2) It is useful to use the variable h for the increment $x - c$. So, f is differentiable at c, if $\exists\ b \in \mathbb{R}$ such that $\lim_{h\to 0} \frac{f(c+h)-f(c)}{h} = b$.

Examples

(1) Let $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$. Then $\lim_{h\to 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0} \frac{|h|}{h}$. Let $x_n = \frac{1}{n}$ and $y_n = -\frac{1}{n}$. Then $x_n \to 0$ and $y_n \to 0$. But $f(x_n) = \frac{|x_n|}{x_n} \to 1$ and $f(y_n) \to -1$. Hence, $\lim_{h\to 0} \frac{|h|}{h}$ does not exist, and therefore, f is not differentiable at 0. On the other hand, verify that f is differentiable at each $c \in \mathbb{R}$, $c \neq 0$, and $f'(c) = 1$ if $c > 0$ and $f'(c) = -1$ if $c < 0$.

(2) Let $f(x) = x^2 \sin\frac{1}{x}$ when $x \neq 0$ and $f(0) = 0$. At the point 0, $\lim_{h\to 0} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0} \frac{h^2 \sin(1/h)}{h} = 0$ (Why?). Thus, f is differentiable at 0. It is clear that f

differentiable at other points as it is the product of two differentiable functions, and $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$. Is the function f' differentiable at 0.

1. Properties of Derivatives

The following result gives a powerful, and a very simple characterization of differentiability of a function at a point

Proposition 3 (Carathéodory's Lemma). *The function $f : J \to \mathbb{R}$ is differentiable at c iff $\exists$ a function $f_1 : J \to \mathbb{R}$ such that*

$$f(x) - f(c) = (x - c) f_1(x), \tag{1}$$

(LHS=RHS)

and f_1 is continuous at c. In such a case, $f'(c) = f_1(c)$.

Proof. ($\Rightarrow$) Define $f_1 : J \to \mathbb{R}$ by

$$f_1(x) = \begin{cases} \frac{f(x)-f(c)}{x-c}, & \text{if } x \in J \setminus \{c\}, \\ f'(c), & \text{if } x = c. \end{cases}$$

Then f_1 satisfies the required properties.

($\Leftarrow$) Putting $x = c + h$ in Equation (1), we have

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h} = \lim_{h \to 0} f_1(c+h) = f_1(c).$$

Thus, f is differentiable at c and $f'(c) = f_1(c)$. □

The function f_1 is called an increment function associated with f and c.

Corollary 4. *If $f : J \to \mathbb{R}$ is differentiable at $c \in J$, then f is continuous at c.*

Theorem 5 (Algebra of Differentiable Functions). *Let $f, g : J \to \mathbb{R}$ be differentiable at $c \in J$. Then the following hold:*

1. *$f + g$ is differentiable at c with $(f + g)'(c) = f'(c) + g(c)$.*
2. *αf is differentiable at c with $(\alpha f)'(c) = \alpha f'(c)$ for any $\alpha \in \mathbb{R}$.*
3. *fg is differentiable at c with $(fg)'(c) = f(c)g'(c) + f'(c)g(c)$.*
4. *If f is differentiable at c with $f'(c) \neq 0$, then $\frac{1}{f}$ is differentiable at c and $(\frac{1}{f})'(c) = -\frac{f'(c)}{f(c)^2}$.*

Corollary 6. *Let $D(J)$ (respectively $C(J)$) denotes the set of differentiable (respectively continuous) functions on J. Then $D(J)$ is a vector subspace of $C(J)$.*

To find the derivative of a composite function $u = g(y)$ where $y = f(x)$, we use the Chain rule or Substitution rule. Roughly speaking, the Chain rule is $\frac{du}{dx} = \frac{du}{dy} \cdot \frac{dy}{dx}$. Can you cancel out dy? Then there would be nothing to prove!

Theorem 7 (Chain Rule). *Let $f : J \to \mathbb{R}$ and $g : J_1 \to \mathbb{R}$ be functions such that $f(J) \subset J_1$, an interval. If f is differentiable at c, and g is differentiable at $f(c)$ then $g \circ f$ is differentiable at c and $(g \circ f)'(c) = g'(f(c)) \cdot f'(c)$.*

Proof. Let $f_1 : J \to \mathbb{R}$ be the increment function associated with f and c, and let $g_1 : J_1 \to \mathbb{R}$ be the increment function associated with g and $f(c)$. Then

$$f(x) - f(c) = (x-c)f_1(x) \;\forall\; x \in J, \;\; g(y) - g(f(c)) = (y - f(c))g_1(y) \;\forall\; y \in J_1.$$

Since $f(J) \subset J_1$, from the above equations we have

$$g(f(x)) - g(f(c)) = [f(x) - f(c)]g_1(f(x)) = (x-c)g_1(f(x))f_1(x) \;\forall\; x \in J.$$

We see that the function $(g_1 \circ f) \cdot f_1 : J \to \mathbb{R}$ is continuous at c (How?). Therefore, by Carathéodory's Lemma, $g \circ f$ is differentiable at c and $(g \circ f)'(c) = g_1(f(c))f_1(c) = g'(f(c))f'(c)$. $\square$

Problem 8. *Let $f : (0,\infty) \to \mathbb{R}$ satisfy $f(xy) = f(x) + f(y)$ for all $x, y \in (0,\infty)$. Suppose f is differentiable at $x = 1$. Show that f is differentiable at every $x \in (0,\infty)$. Find out $f'(x)$.*

Solution. We first observe that $f(1) = 0$, $f(\frac{1}{x}) = -f(x)$ and $f(\frac{x}{y}) = f(x) - f(y)$. Now,

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{1}{h} f\left(\frac{x+h}{h}\right) = \lim_{k\to 0} \frac{f(1+k)}{kx} \\ &= \frac{1}{x} \lim_{k\to 0} \frac{f(1+k) - f(1)}{k} = \frac{1}{x} f'(1). \end{aligned}$$

$\square$

8

MEAN VALUE THEOREM

1. Tangents and Rate of Change

There are two important ways of looking at derivatives. One interpretation is physical and the other is geometric.

Let $f : [a, b] \to \mathbb{R}$ be a differentiable function. We can interpret the difference quotient $\frac{f(x+h)-f(x)}{h}$ as the *average rate of change* of f over the interval from x to $x + h$. The *instantaneous rate of change* of f with respect to x at x_0 is the derivative $f'(x_0) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$. Thus, instantaneous rates are limits of average rates.

Problem 1. *The area A of a circle is related to its diameter D by the equation $A = \frac{\pi}{4}D^2$. How fast does the area change w.r.t the diameter when the diameter is 10 m.*

Solution. The rate of change of the area w.r.t the diameter is $\frac{dA}{dD} = \frac{\pi D}{2}$. When $D = 10$ m, the area changes w.r.t the diameter at the rate of 5π m. □

If we think of the domain as the time interval and $f(x_0 + h) - f(x_0)$ as the distance travelled by a particle in h units of time, then the velocity is $\frac{f(x_0+h)-f(x_0)}{h}$. The derivative which is the limit of these velocities as $h \to 0$ is called the instantaneous velocity of the motion of the particle at the instant $x = x_0$.

The geometric interpretation of the derivative $f'(x)$ is that it is the slope of the tangent line at $(x, f(x))$ to the graph $\{(x, f(x)) : x \in [a, b]\}$.

Problem 2. (1) *Find the slope of the curve $y = \frac{1}{x}$ at any point $x = a \neq 0$. What is the slope of at $x = -1$.*

(2) *At which point(s), the slope equals $-\frac{1}{4}$.*

(3) *What happens to the tangent to the curve at the point $(a, \frac{1}{a})$ as a changes.*

Solution. (1) Here, $f(x) = \frac{1}{x}$. Then $f'(x) = -\frac{1}{x^2}$. Hence, the slope at $(a, \frac{1}{a})$ is $-\frac{1}{a^2}$.

(2) $-\frac{1}{a^2} = -\frac{1}{4} \Rightarrow a = 2$ or $a = -2$. Thus, the curve has slope $-\frac{1}{4}$ at the points $(2, \frac{1}{2})$ and $(-2, -\frac{1}{2})$.

(3) The slope $-\frac{1}{a^2}$ is always negative if $a \neq 0$. As $a \to 0^+$ (or 0^-), the slope approaches $-\infty$ and the tangent becomes increasingly steep. As a moves away from 0 in either direction, the slope approaches 0 and the tangent levels off to become horizontal.

□

2. Mean Value Theorem

Definition 3 (Local Exremum). *(1) Let J be an interval and $f : J \to \mathbb{R}$ be a function. We say that a point $c \in J$ is a point of local maximum if there exists $\delta > 0$ such that $(c-\delta, c+\delta) \subset J$ and $f(x) \leq f(c)$ for all $x \in (c-\delta, c+\delta)$. A local minimum is defined similarly.*

(2) A point c is said to be a local extremum if it is either a local maximum or a local minimum.

(3) A point x_0 is called a point of global maximum on J if $f(x) \leq f(x_0)$ for all $x \in J$. Global minimum and then global extremum are defined similarly.

Examples

(1) Let $f : [a, b] \to \mathbb{R}$ be defined as $f(x) = x$. Then b is a point of global maximum but not a local maximum. Also, a is a point of global minimum but not a local minimum.

(2) Consider $f : [-2\pi, 2\pi] \to \mathbb{R}$ where $f(x) = \cos x$. The point $x = 0$ is a local maximum as well as a global maximum. What do you think about the points $x = \pm 2\pi$.

Proposition 4. *Let $f : J \to \mathbb{R}$ be differentiable at $c \in J$. If f has local extremum at c, then $f'(c) = 0$.*

Proof. Suppose f has a local maximum at c. Then $\exists\ \delta > 0$ such that $f(x) \leq f(c)$ for all $x \in (c-\delta, c+\delta)$. In other words, for all $h \in (-\delta, \delta)$, we have $f(c+h) \leq f(c)$ and $f(c-h) \leq f(c)$. Since f is differentiable at c,

$$f'(c) = \lim_{h \to 0^+} \frac{f(c+h) - f(c)}{h} \leq 0 \text{ and } f'(c) = \lim_{h \to 0^-} \frac{f(c+h) - f(c)}{h} \geq 0.$$

This implies that $f'(c) = 0$. □

Proposition 5 (Rolle's Theorem). *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and if $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Exercise. □

Geometrically, Rolle's Theorem says that there exists a point $c \in (a, b)$ such that the tangent at $(c, f(c))$ is parallel to x-axis.

Rolle's Theorem together with IVP are used to check the existence and uniqueness of roots of continuous functions in certain intervals as illustrated in the following examples.

Examples

(1) Let $f(x) = x^3 + p + q$ for $x \in \mathbb{R}$, where $p, q \in \mathbb{R}$, $P > 0$. We observe that $f(x) \to \infty$ as $x \to \infty$ and $f(x) \to -\infty$ as $x \to -\infty$. Hence by IVP, there $a \in \mathbb{R}$ such that $f(a) = 0$. Thus, f has at least one real root. Suppose there is $b \in \mathbb{R}$ such that $f(b) = 0$. Rolle's Theorem implies the existence of $c \in (a, b)$ such that $f'(c) = 0$. But $f'(x) = 3x^2 + p \neq 0$ for any $x \in \mathbb{R}$ since $p > 0$.

(2) If $f(x) = x^4 + 2x^3 - 2$ for $x \in \mathbb{R}$, then $f(0) = -2 < 0$ and $f(1) = 1 > 0$. Therefore, by IVP f will have at least one root in $[0, 1]$. Moreover, $f'(x) = 4x^3 + 6x^2 > 0$ for $x \in (0, \infty)$. So, f has at most one root in $[0, \infty)$. This implies that f has a unique root in $[0, \infty)$.

Now we state the most important result in differentiation.

Theorem 6 (Mean Value Theorem (MVT))**.** *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that*

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

Proof. Define $F : [a, b] \to \mathbb{R}$ by

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

Now, apply Rolle's Theorem to the function F. □

Remark 7. *(1) The mean value theorem (in short, MVT) is also known as Lagrange's mean value theorem. MVT is crucial in characterizing constant functions, monotonic functions, and convex/concave functions. Such characterizations can only be obtained using MVT.*

(2) If we write $b = a + h$, then MVT could be stated as follows:

$$f(a + h) = f(a) + hf'(a + \theta h) \text{ for some } \theta \in (0, 1). \text{ (Verify!)}$$

2.1. Applications of MVT.

(1) Let $f : [a, b] \to \mathbb{R}$ be differentiable such that $f'(x) = 0$ for all $x \in [a, b]$. Then f is constant.

Proof. For any $x, y \in [a, b]$, by MVT, we have $f(y) - f(x) = f'(z)(y - x)$ for some z between x and y. Since $f'(z) = 0$, we get $f(x) = f(y)$ for all $x, y \in [a, b]$. This implies that f is a constant function. (Try to prove this ab initio!) □

(2) Let $f : [a, b] \to \mathbb{R}$ be differentiable. If $f'(x) \geq 0$ (respectively, $f'(x) > 0$) for all $x \in [a, b]$, then f is increasing (respectively, strictly increasing). We have a similar result for decreasing functions.

Proof. For $x, y \in [a, b]$ such that $x < y$, we have by MVT, $f(y) - f(x) = f'(z)(y - x)$ for some z between x and y. Since $f'(z) \geq 0$ (respectively, $f'(z) > 0$) and $y - x > 0$, we get $f(y) \geq f(x)$ (respectively, $f(y) > f(x)$). Thus, f is increasing (respectively, increasing). □

(3) MVT is quite useful in proving certain inequalities. For example, can you find out which is greater, e^π or π^e. Let us prove a more general inequality.

■ If $e \leq a < b$, then $a^b > b^a$.

Proof. Let $0 < x < y$ and $f(x) = \log x$ on $[x, y]$. Using MVT, try to prove

$$\frac{y - x}{y} < \log \frac{y}{x} < \frac{y - x}{y}. \tag{1}$$

Now, using (1), we have

$$\frac{b - a}{b} < \log \frac{b}{a} < \frac{b - a}{a}.$$

Since $a \log \frac{b}{a} < b - a$, we have $\frac{b^a}{a^a} = e^{a \log \frac{b}{a}} < e^{b-a}$. That is, $b^a < e^{b-a} a^a$. If $e \leq a$, then $e^t \leq a^t$ for $t \geq 0$ (Why?). Thus, we conclude that

$$b^a < e^{b-a} a^a < a^{b-a} a^a = a^b.$$

□

Theorem 8 (Cauchy's Mean Value Theorem (CMVT))**.** *If $f, g : [a, b] \to \mathbb{R}$ are continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that*

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

Proof. Exercise. □

9

L'HÔPITAL'S RULES

1. Indeterminate Forms

In this lecture we will discuss limit theorems that involve cases which cannot be determined by previous limit theorems.

Let $f, g : [a, b]$ be two functions such that $g(x) \neq 0$ for all $x \in [a, b]$, $x \neq c$, and $c \in [a, b]$. We have seen that if $\lim_{x\to c} f(x) = A$ and $\lim_{x\to c} g(x) = B$, and if $B \neq 0$, then $\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{A}{B}$. If $B = 0$, then no conclusion was deducted. Consider the following two situations:

(1) If $B = 0$, and $\lim_{x\to c} \frac{f(x)}{g(x)}$ exists, then $A = \lim_{x\to c} f(x) = \lim_{x\to c} \frac{f(x)}{g(x)} g(x) = 0$.

(2) If $g(x) > 0$ for all $x \in [a, b]$, $x \neq c$, $A > 0$ and $B = 0$, then $\lim_{x\to c} \frac{f(x)}{g(x)} = \infty$ (Try to prove this!). Similarly, if $A < 0$ and $B = 0$, then $\lim_{x\to c} \frac{f(x)}{g(x)} = -\infty$.

The case $A = B = 0$ is still remaining. In this case, the limit of the quotient $\frac{f}{g}$ is said to be "indeterminate", and depending on the functions f and g the limit may not exist or may be any real number. The symbolism $\frac{0}{0}$[1] is used to refer to this situation. For instance, if $\lambda \in \mathbb{R}$, then for $f(x) = \lambda x$ and $g(x) = x$,

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{\lambda x}{x} = \lambda.$$

Thus, the indeterminate form $\frac{0}{0}$ can lead to any real number λ as a limit.

Other indeterminate forms are represented by the symbols $\frac{\infty}{\infty}$, $0 \cdot \infty$, 0^0, 1^∞, ∞^0 and $\infty - \infty$. These notations correspond to the indicated limiting behaviour of the functions f and g. We will focus on the indeterminate forms $\frac{0}{0}$ and $\frac{\infty}{\infty}$. Other indeterminate cases are usually reduced to the form $\frac{0}{0}$ or $\frac{\infty}{\infty}$ by taking logarithms, exponentials, or algebraic manipulations.

2. L'Hôpital's Rules

Most calculus books discuss L'Hôpital's Rules for $\frac{0}{0}$ and $\frac{\infty}{\infty}$ indeterminate forms separately. In the next theorem, we will give a general statement of L'Hôpital's Rule which covers the above two forms and some other cases as well.

---

[1] The limit theorem for the $\frac{0}{0}$ case, later known as L'Hôpital's (pronounced as Lowpeetal) Rule, was actually discovered by Bernoulli. It appeared in the first textbook on differential calculus book written by L'Hôpital's in 1696. Bernoulli was L'Hôpital's teacher.

Theorem 1 (L'Hôpital's Rule). *Let J be an open interval. Let either $a \in J$ or a is an endpoint of J (It is possible that $a = \pm\infty$). Assume that*

(1) $f, g : J \setminus \{a\} \to \mathbb{R}$ is differentiable,
(2) $g(x) \neq 0 \neq g'(x)$ for $x \in J \setminus \{a\}$,
(3) $A = \lim_{x\to a} f(x) = \lim_{x\to a} g(x)$, where A is either 0 or ∞, and
(4) $B = \lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists either in $\mathbb{R}$ or $B = \pm\infty$.

Then

$$\boxed{\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} = B.}$$

Proof. We apply Cauchy mean value theorem to prove the case where $A = 0$, $a \in \mathbb{R}$, and $B \in \mathbb{R}$. Interested readers may refer to [1] for the proof of other cases.

Define $f(a) = 0 = g(a)$ to make f and g continuous on J. Let (x_n) be a sequence in J such that $x_n > a$ or $x_n < a$ for all n and $x_n \to a$. By CMVT, $\exists\, c_n$ between a and x_n such that

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f(a)}{g(x_n) - g(a)} = \frac{f'(c_n)}{g'(c_n)}.$$

?

Now, $x_n \to a$ implies that $c_n \to a$. By hypothesis, the sequence $\frac{f'(c_n)}{g'(c_n)} \to B$. Thus, $\frac{f(x_n)}{g(x_n)} \to B$, and hence $\lim_{x\to a} \frac{f(x)}{g(x)} = B$. $\square$

Before looking at some examples, note that when a is an endpoint of J, we get the statement of L'Hôpital's Rule for left and right hand limits.

Examples

(1) $\lim_{x\to 2} \frac{\sqrt{x^2+5}-3}{x^2-4} = \lim_{x\to 2} \frac{\frac{x}{\sqrt{x^2+5}}}{2x} = \frac{\frac{2}{3}}{4} = \frac{1}{6}$.
(2) $\lim_{x\to\infty} (x^3 + 4x^2 + 13x + 1)^{1/3} - x = \lim_{y\to 0^+} \frac{(1+4y+13y^2+y^3)^{1/3}-1}{y} = \lim_{y\to 0^+} \frac{1}{3}(1 + 4y + 13y^2 + y^3)^{-2/3}(4 + 26y + 3y^2) = \frac{4}{3}$.

Examples

(1) $\lim_{x\to\infty} \frac{x^2+2x+3}{3x^2+2x+1} = \lim_{x\to\infty} \frac{2x+2}{6x+2} = \lim_{x\to\infty} \frac{2}{6} = \frac{1}{3}$.
(2) $\lim_{x\to\infty} \frac{x^3}{x^2-1} - \frac{x^3}{x^2+1} = \lim_{x\to\infty} \frac{2x^3}{x^4-1} = \lim_{x\to\infty} \frac{6x^2}{4x^3} = \lim_{x\to\infty} \frac{3}{2x} = 0$.

Other Indeterminate Forms.

(1) $\lim_{x\to 0^+} x \log x = \lim_{x\to 0^+} \frac{\log x}{\frac{1}{x}} = \lim_{x\to 0^+} \frac{\frac{1}{x}}{\frac{-1}{x^2}} = 0$.
(2) $\lim_{x\to 0^+} \left(\frac{1}{x} - \frac{1}{\sin x}\right) = \lim_{x\to 0^+} \frac{\sin x - x}{x \sin x} = \lim_{x\to 0^+} \frac{\cos x - 1}{\sin x + x\cos x} = \lim_{x\to 0^+} \frac{-\sin x}{2\cos x - x\sin x} = 0$.
(3) $\lim_{x\to\infty} (1 + \frac{1}{x})^x = \lim_{x\to\infty} e^{x\log(1+\frac{1}{x})}$. We note that $\lim_{x\to\infty} x\log(1+\frac{1}{x}) = \lim_{x\to\infty} \frac{\log(1+\frac{1}{x})}{\frac{1}{x}} = \lim_{x\to\infty} \frac{1}{1+\frac{1}{x}} = 1$. Thus, $\lim_{x\to\infty}(1+\frac{1}{x})^x = e$.

References

[1] Tom M. Apostol, Calculus. Vol. I: One-variable calculus, with an introduction to linear algebra, Second edition, 1967.

10

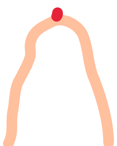
LOCAL EXTREMA AND POINTS OF INFLECTION

In lecture 8, we have seen a necessary condition for local maximum and local minimum. In this lecture we will see some sufficient conditions.

In the following results we assume $f : (a, b) \to \mathbb{R}$ and $c \in (a, b)$.

1. Sufficient conditions for a local extremum

We will state results for local maximum, and results for local minimum are similar.



Theorem 1. *Let f be continuous at c. If for some $\delta > 0$, f is increasing on $(c - \delta, c)$ and decreasing on $(c, c + \delta)$, then f has a local maximum at c.*

Proof. Choose x and y such that $c - \delta < x < y < c$. Then $f(x) \le f(y)$. The continuity of f at c implies that $f(x) \le \lim_{y \to c^-} f(y) = f(c)$. Similarly, if $c < y < x < c + \delta$, then $f(x) \le \lim_{y \to c^+} f(y) = f(c)$. This proves the result. $\square$

Corollary 2. (1) ***(First Derivative Test for Local Maximum)*** *Let f be continuous at c. If for some $\delta > 0$*

$$f'(x) \ge 0 \ \forall\, x \in (c - \delta, c) \text{ and } f'(x) \le 0 \ \forall\, x \in (c, c + \delta),$$

then f has a local maximum at c.

(2) ***(Second Derivative Test for Local Maximum)*** *If f is twice differentiable at c and satisfies $f'(c) = 0$ and $f''(c) < 0$, then f has a local maximum at c.*

Remark 3. *An easy way to remember the First Derivative Test (for local minimum and local maximum) is as follows:*

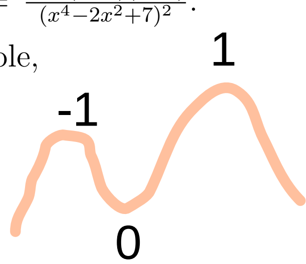
f' changes from $-$ to $+$ at $c \Rightarrow f$ has a local minimum at c,

f' changes from $+$ to $-$ at $c \Rightarrow f$ has a local maximum at c.

Examples

(1) Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = \frac{1}{x^4 - 2x^2 + 7}$. We have $f'(x) = \frac{-4x(x-1)(x+1)}{(x^4 - 2x^2 + 7)^2}$. Thus, $f'(x) = 0$ when $x = -1, 0, 1$. Now, consider the following table,

Interval	$(-\infty, -1)$	$(-1, 0)$	$(0, 1)$	$(1, \infty)$
Sign of f'	$+$	$-$	$+$	$-$



So, we conclude that f has a local minimum at $x = 0$ and a local maxima at $x = -1$ and $x = 1$.

(2) Consider $f : (-1, 1) \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2 & \text{if } 0 < |x| < 1, \\ -1, & \text{if } x = 0. \end{cases}$$

We note the conditions of the first derivative test is not satisfied. In fact, f is differentiable on $(-1, 0)$ and $(0, 1)$ and f' changes sign from $-$ to $+$ at $x = 0$ but f is not continuous at $x = 0$. Nevertheless, $f(0) < f(x)$ for all nonzero $x \in (-1, 1)$, and thus f has a strict local minimum at $x = 0$.

(3) If $f : \mathbb{R} \to \mathbb{R}$ is defined as $f(x) = x^4$. Then $f(0) = 0 < f(x)$ for all nonzero $x \in \mathbb{R}$). Therefore, f has a strict local minimum at $x = 0$. Note that $f'(0) = 0$, but $f''(0)$ is not positive.

2. Convex Sets and Convex Functions

Let V be a vector space over $\mathbb{R}$.

Definition 4. *A set $C \subseteq V$ is said to be convex if the line segment between any two points in C lies in C, i.e.,*

x y

if for any $x, y \in C$ and any $t \in [0, 1]$, we have $tx + (1-t)y \in C$.

Example. It is clear that the unit disc is convex in $\mathbb{R}^2$. However, the unit circle is not convex. Any interval in $\mathbb{R}$ is a convex set.

Definition 5. *Let $C \subseteq V$ be a convex set. A function $f : C \to \mathbb{R}$ is said to be* ***convex*** *if for all $x, y \in C$ and for all $t \in [0, 1]$, we have*

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

If for $t \in (0, 1)$, the above inequality is strict, the f is said strictly convex.

We say that f is ***concave*** *if the reverse inequality in ($\boxed{\star}$) holds.*

Theorem 6 (Derivative Test for Convexity)**.** *Assume that $f : [a, b]$ is differentiable on (a, b). If f' is increasing on (a, b), then f is convex on $[a, b]$. In particular, if f'' exists and non-negative on (a, b), then f is convex.*

kind of graph => convex

Example. Let $f(x) = x^3 - 6x^2 + 9x$. We have $f'(x) = 3(x-1)(x-3)$ and $f''(x) = 6x - 12$. We see that $f''(x) > 0$ if $x > 2$ and $f''(x) < 0$ if $x < 2$. Hence, f is convex for $x > 2$ and concave for $x < 2$.

Examples of convex functions

- e^x is strictly convex on $\mathbb{R}$.
- $x \log x$ is strictly convex on $(0, \infty)$.

- $f(x) = x^4$ is strictly convex but $f''(0) = 0$

The following result is one of the reasons why convex functions are very useful in applications especially in optimization problems.

Theorem 7. *If $f : (a, b) \to$ is convex and $c \in (a, b)$ is a local minimum, then c is a minimum for f on (a, b). That is, local minima of convex functions are global minima.*

3. Points of Inflection

Definition 8. *Let $f : (a, b) \to \mathbb{R}$ be a function and $c \in (a, b)$. The point c is said to be a **point of inflection** for f if there is $\delta > 0$ such that f is convex in $(c - \delta, c)$, while f is concave in $(c, c + \delta)$, or vice versa, that is, f is concave in $(c - \delta, c)$, while f is convex in $(c, c + \delta)$.*

Examples For the function $f(x) = x^3$ on $\mathbb{R}$, 0 is a point of inflection.

Theorem 9. *Let $f : (a, b) \to \mathbb{R}$ be a function and $c \in (a, b)$.*

*(1) [**Necessary Condition for a Point of Inflection**] Let f be twice differentiable at c. If c is a point of inflection for f, then $f''(c) = 0$.*

*(2) [**Sufficient Condition for a Point of Inflection**] Let f be thrice differentiable at c. If $f''(c) = 0$ and $f'''(c) = 0$, then c is a point of inflection for f.*

Examples

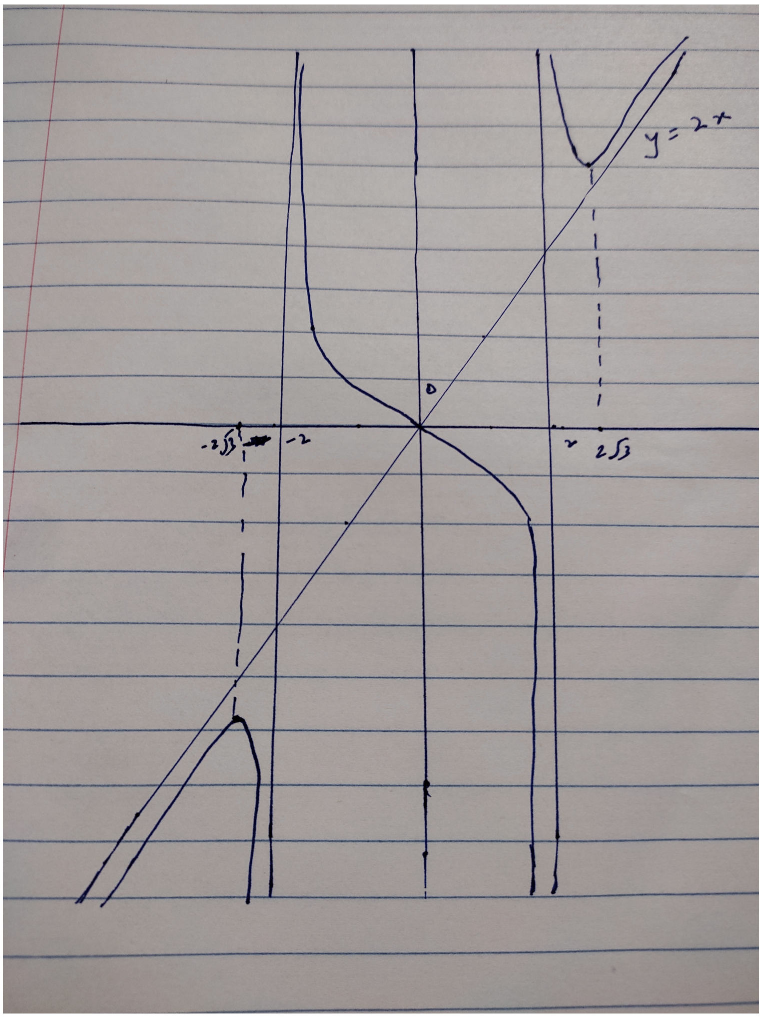
- For the function $f(x) = x^4$, 0 is not a point of inflection, though, $f''(0) = 0$.
- For the function $f(x) = x^5$, 0 is a point of inflection, but $f'''(0) = 0$.

Problem 10. *Sketch the graph of the function $f(x) = \frac{2x^3}{x^2-4}$ after finding the intervals of decrease/increase, intervals of concavity/convexity, points of local minima/local maxima, points of inflection and asymptotes.*

Solution. We note that

$$f(x) = 2x + \frac{8x}{x^2 - 4}, \quad f'(x) = \frac{2x^2(x^2 - 12)}{(x^2 - 4)^2} \text{ and } f''(x) = \frac{16x(x^2 + 12)}{(x^2 - 4)^3}.$$

Verify that $x = 2$, $x = -2$ and $y = 2x$ are the asymptotes. Moreover, the function is increasing on $(-\infty, -2\sqrt{3})$ and $(2\sqrt{3}, \infty)$. The function is decreasing on $(-2\sqrt{3}, -2)$, $(-2, 2)$ and $(2, 2\sqrt{3})$. Furthermore, the function is convex on $(-2, 0)$ and $(2, \infty)$ and concave on $(-\infty, -2)$ and $(0, 2)$. The point of inflection is 0. The sketch of the graph is shown below.



y = 2x
0
$-2\sqrt{3}$
-2
2
$2\sqrt{3}$

□

To do

11

THE PICARD AND NEWTON METHODS

In this lecture we will study a numerical method which is a technique to find the approximate solution of the equation $f(x) = 0$. This is the next best alternative to finding an exact solution. In fact, finding an exact solution is a very difficult problem even for the nicest functions, namely polynomials[1].

Consider the following two problems.

(1) Let $f : [a, b] \to [a, b]$ be a function. Does there exist a point $x \in [a, b]$ such that $f(x) = x$. Such a point is called a fixed point of f. Fixed point theory is one of the most powerful tools of modern mathematics.

(2) Let $f : [a, b] \to \mathbb{R}$ be a function. Does there exist a point $x \in [a, b]$ such that $f(x) = 0$.

We first observe that these two problems are interrelated in the sense that if you solve one, you solve the other. If $f : [a, b] \to [a, b]$, and if we choose $F(x) = f(x) - x$, then finding a fixed point of f is equivalent to finding a solution to $F(x) = 0$. On the other hand, if $f : [a, b] \to \mathbb{R}$ and we set $F(x) = x + h(x)g(x)$ where $h : [a, b] \to \mathbb{R}$ is chosen such that $h(x) \neq 0$ and $x + h(x)g(x) \in [a, b]$ for all $x \in [a, b]$, then finding a solution of $f(x) = 0$ is equivalent to finding a fixed point of $F : [a, b] \to [a, b]$.

Proposition 1. *If $f : [a, b] \to [a, b]$ is continuous, then f has a fixed point.*

Proof. Define $F : [a, b] \to \mathbb{R}$ by $F(x) = f(x) - x$. Now,

$$F(a) = f(a) - a \geq 0 \text{ and } F(b) = f(b) - b \leq 0.$$

By IVP, there exists $c \in [a, b]$ such that $F(c) = 0$, i.e., $f(c) = c$. $\square$

1. Picard method

Suppose that a function $f : [a, b] \to [a, b]$ has a fixed point. The question is that can we find it. In general this is so easy. So, we try to find it approximately. A simple and effective method is given by Picard which is described in the algorithm below.

[1] For linear and quadratic equations, there are simple and well-known formulas for their solutions. For cubic and quadratic equations, there are complicated formulas due to Cardan and Ferrari, which express the solutions in terms of the coefficients of the polynomial. For a general polynomial equation of degree 5 or more, Abel proved that no such formula exists. Galois' theory provides a much more complete answer to this question, by explaining why it is possible to solve some equations, and why it is not possible for most equations of degree 5 or higher.

Algorithm 1. Given any $x_0 \in [a,b]$, define (x_n) by

$$x_n = f(x_{n-1}) \text{ for } n \in \mathbb{N}.$$

Such a sequence is called Picard sequence for the function f (with its initial point x_0). It is clear that if (x_n) is convergent and f is continuous, then the limit x of (x_n) is a fixed point of f. Indeed,

$$f(x) = f(\lim_{n\to\infty} x_n) = f(\lim_{n\to\infty} x_{n-1}) = \lim_{n\to\infty} f(x_{n-1}) = \lim_{n\to\infty} x_n = x.$$

The next theorem gives a sufficient condition for the convergence of a Picard sequence.

Theorem 2 (Picard Convergence Theorem)**.** *If $f : [a,b] \to [a,b]$ is continuous on $[a,b]$ and differentiable on (a,b) with $|f'(x)| < 1$ for all $x \in (a,b)$, then f has a unique fixed point. Furthermore, any Picard sequence for f is convergent and converges to the unique fixed point of f.*

Example. Let $f : [0,2] \to \mathbb{R}$ be defined as $f(x) = (1+x)^{1/5}$. Then $0 \leq f(x) \leq 3^{1/5} < 2$ for all $x \in [0,2]$. Thus, f maps the interval $[0,2]$ into itself. Moreover, $|f'(x)| = \frac{1}{5(1+x)^{4/5}} \leq \frac{1}{5} < 1$ for $x \in [0,2]$. Picard convergence theorem implies that f has a unique fixed point. In other words, the sequence (x_n) defined by $x_{n+1} = (1+x_n)^{1/5}$ converges to a root of $x^5 - x - 1 = 0$ in the interval $[0,2]$.

2. Newton–Raphson method

Suppose we know that $f : [a,b] \to \mathbb{R}$ is such that the equation $f(x) = 0$ has a solution. We have discussed earlier that it is difficult to find an exact solution. So, we find an approximate solution. A method given by Newton is used to achieve this.

Algorithm 2. Choose any $x_0 \in [a,b]$ such that $f'(x_0)$ exists and $f'(x_0) \neq 0$. Given any $n \in \mathbb{N}$ and $x_{n-1} \in [a,b]$ such that $f'(x_{n-1}) \neq 0$, let

$$x_n = x_{n-1} - \frac{f(x_n - 1)}{f'(x_{n-1})}.$$

Such a sequence (x_n) is called a Newton sequence for the function f (with its initial point x_0). It is easy to verify that if a Newton sequence (x_n) for f is convergent and f' is bounded, then the limit x of (x_n) satisfies $f(x) = 0$. We can see that if we choose $f(x) = x - \frac{f(x)}{f'(x)}$, then algorithm 2 is a particular case of algorithm 1.

Example. Suppose $f(x) = x^2 - a$, $a > 0$. Since $f'(x) = 2x$, the Newton sequence for the function f is $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$. We have seen in Lecture 2 that this sequence converges to $\sqrt{a}$.

We end this lecture by giving a sufficient condition for the convergence of a Newton sequence.

Proposition 3. *Let $f : [a,b] \to \mathbb{R}$ be such that $f(c) = 0$ for some $[a,b]$. If f' is nonzero and monotonic throughout $[a,b]$, then c is the unique solution of $f(x) = 0$ in $[a,b]$ and the Newton sequence with any initial point $x_0 \in [a,b]$ converges to c.*

12

TAYLOR'S THEOREM

Taylor Polynomials

Sometimes we can approximate complicated functions with simpler ones that give the accuracy we want for specific applications and are easier to work with. The approximating functions discussed here are called linearizations, and they are based on tangent lines. Other approximating functions, such as polynomials will be discussed afterwards.

Consider the function $f(x) = x^2$. The tangent to the curve $y = x^2$ at $(1,1)$ is $y = 2x - 1$. If we plot the graph of the curve and its tangent, we can see that the tangent line lies close to the curve near the point of tangency. If we zoom the graph of f near $(1,1)$, the graph becomes flatter and it almost resembles its tangent. Thus, for a small interval around the point 1 on x-axis, the y-values along the tangent line gives a good approximations to the y-values on the curve.

In general, the tangent to $y = f(x)$ at a point c, where f is differentiable, is the line

$$\boxed{L(x) = f(c) + f'(c)(x - c).}$$

As long as this line remains close to the graph of f, $L(x)$ provides a good approximation to $f(x)$. The approximating function $L(x)$ is called the linearization of f at c. This is the standard linear approximation of f at c, and the point c is the center of the approximation.

Note that the linearization $L(x)$ of f at c is a polynomial of degree one. The question here is that can we get a better approximation if we take a polynomial of higher degree. The answer is yes and this is what Taylor's theorem talks about[1].

Definition 1. *Let $f : [a,b] \to \mathbb{R}$ be a function which is n-times differentiable at $c \in (a,b)$. The Taylor polynomial of order n generated by f at c is the polynomial*

$$\boxed{P_n(x) = f(c) + f'(c)(x - c) + \frac{f''(c)}{2!}(x - c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x - c)^n.}$$

Example 2. *Find the Taylor polynomial generated by $f(x) = e^x$ at $c = 0$.*

Solution. Since $f^{(n)}(x) = e^x$ and $f^{(n)}(0) = 1$ for every n, the Taylor polynomial of order n at $c = 0$ is

$$P_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}.$$

[1] Taylor's theorem is named after the English mathematician Brook Taylor. He also introduced Taylor series which will be discussed later. Both Taylor's theorem and Taylor series are among the most useful results in calculus.

Taylor's Theorem

Theorem 3. *Assume that $f : [a,b] \to \mathbb{R}$ is such that $f^{(n)}$ is continuous on $[a,b]$ and $f^{(n+1)}(x)$ exists on (a,b). Fix $x_0 \in [a,b]$. Then for each $x \in [a,b]$ with $x \neq x_0$, there exists c between x and x_0 such that*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}. \qquad (\star)$$

Remark 4. *(1) Note that the MVT corresponds to the case $n = 0$ of Taylor's Theorem. The case $n = 1$ is sometimes called the Extended Mean Value Theorem. In both cases we take $x_0 = a$ and $x = b$.*

(2) The right-hand side of Equation $(\star)$ is called the n-th order Taylor expansion (or formula) for f around x_0.

(3) The term $R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$ is called the remainder term of order n. It is known as Lagrange's form of remainder in Taylor's formula.

(4) Usually, the n-th order Taylor polynomial $P_n(x)$ of f around x_0 provides a progressively better approximation to f around x_0 as n increases. The remainder term is the "error term" if we wish to approximate f near x_0 by $P_n(x)$. If we assume that $f^{(n+1)}$ is bounded by M on (a,b), then R_n goes to 0 much faster than $(x-x_0)^n \to 0$, since $\left|\frac{R_n(x)}{(x-x_0)^n}\right| \leq \frac{M}{(n+1)!}|x-x_0|$.

Corollary 5. *Let $f : [a,b] \to \mathbb{R}$ and n be a nonnegative integer. Then f is a polynomial function of degree $\leq n$ iff $f^{(n+1)}$ exists and is identically zero on $[a,b]$.*

Proof. Exercise. □

Example 6. *Using Taylor's theorem show that for any $k \in \mathbb{N}$ and for all $x > 0$,*

$$x - \frac{1}{2}x^2 + \cdots + \frac{1}{2k}x^{2k} < \log(1+x) < x - \frac{1}{2}x^2 + \cdots + \frac{1}{2k+1}x^{2k+1}.$$

Solution. By Taylor's theorem, there exists $c \in (0,x)$ such that (x0 = 0)

$$\log(1+x) = x - \frac{1}{2}x^2 + \cdots + \frac{(-1)^{n-1}}{n}x^n + \frac{(-1)^n}{n+1}\frac{1}{(1+c)^{n+1}}x^n.$$

We observe that for any $x > 0$,

$$\frac{(-1)^n}{n+1}\frac{1}{(1+c)^{n+1}}x^n = \begin{cases} > 0 & \text{if } n = 2k \\ < 0 & \text{if } n = 2k+1. \end{cases}$$

Problem 7. *Let $x_0 \in (a,b)$ and $n \geq 2$. Suppose $f', f", \ldots, f^{(n)}$ are continuous on (a,b) and $f'(x_0) = \cdots = f^{(n-1)}(x_0) = 0$. Show that if n is even and $f^{(n)}(x_0) > 0$, then f has a local minimum at x_0. Similarly, if n is even and $f^{(n)}(x_0) < 0$, show that f has a local maximum at x_0.*

Solution. Suppose that $f^{(n)}(x_0) > 0$ and n is even. Since $f^{(n)}$ is continuous at x_0, there exists a neighbourhood U of x_0 such that $f^{(n)}(x) > 0$ for all $x \in U$. By Taylor's theorem, for $x \in U$, there exists c between x and x_0 such that $f(x) = f(x_0) + \frac{f^{(n)}(c)}{n!}(x - x_0)^n$. It follows that $\frac{f^{(n)}(c)}{n!}(x - x_0)^n > 0$ ($\because$ $c \in U$ and n is even). This implies that $f(x) > f(x_0)$ for all $x \in U$. Hence, x_0 is a local minimum.

13

SERIES

We know that the set of real numbers is a group under the usual addition. The associativity of real numbers allows us to add any finite real numbers $x_1, x_2, \ldots, x_n$ as $x_1 + x_2 + \cdots + x_n$. In this lecture we will learn whether we can add infinitely many real numbers. In other words, if (x_n) is a sequence, then what is the meaning of the symbol $\sum_{n=1}^{\infty} x_n = x_1 + x_2 + \cdots + x_n + \cdots$. Such expression is called an infinite series, or just s series. What will happen if we try to add them term by term. For example, consider the sequence, $x_n = (-1)^{n-1}$. Then

$$\begin{aligned}
1 - 1 + 1 - 1 + \cdots &= (1-1) + (1-1) + \cdots \\
&= 0 + 0 + \cdots + 0 + \cdots \\
&= 0, \\
1 - 1 + 1 - 1 + \cdots &= 1 + (-1+1) + (-1+1) + \cdots \\
&= 1 + 0 + 0 + \cdots + 0 + \cdots \\
&= 1.
\end{aligned}$$

This absurdity shows that we should define the sum of infinite real numbers in a rigorous way so as to avoid this. In your school days, you have learned the geometric series $1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^n} + \cdots = 2$. What do we mean by this?

Convergence and Sum of an Infinite Series

Definition 1. *Let (x_n) be a sequence of real numbers. Define the sequence*

$$S_n = x_1 + x_2 + \cdots + x_n.$$

*Then (S_n) is called the **sequence of (n-th) partial sum** of the series $\sum_{n=1}^{\infty} x_n$.*

*We say that the infinite series $\sum_{n=1}^{\infty} x_n$ is **convergent** if the sequence (S_n) of partial sums is convergent.*

The limit of (S_n), say S, is called the sum of the series. We denote this fact by the symbol $\sum_{n=1}^{\infty} x_n = S$.

*We say that the series $\sum_{n=1}^{\infty} x_n$ is **divergent** if the sequence of its partial sums is divergent.*

*The series $\sum_{n=1}^{\infty} x_n$ is said to be **absolutely convergent** if the infinite series $\sum_{n=1}^{\infty} |x_n|$ is convergent.*

If a series is convergent but not absolutely convergent, then it is said to be **conditionally convergent***.*

You need to remember the following remark by heart.

Remark 2. *The ONLY way to deal with an infinite series is through its sequence partial sums and by using the definition of the sum of an infinite series.*

You need to be careful when dealing with infinite series. Mindless algebraic manipulations may lead to absurdities as shown in the beginning of this lecture.

Examples.

(1) **(Geometric Series)** Let $x \in \mathbb{R}$ such that $|x| < 1$. Consider the series $\sum_{n=0}^{\infty} x^n$. The sequence of partial sums is

$$S_n = 1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Since $|x| < 1$, we have $x^{n+1} \to 0$, and therefore, $S_n \to \frac{1}{1-x}$. Thus, $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ if $|x| > 1$.

(2) The series $\sum_{n=0}^{\infty} \log(\frac{n+1}{n})$ diverges because $S_n = \log(n+1)$ diverges.

(3) **(Telescoping Series)** Let (x_n) and (y_n) be two sequences such that $x_n = y_{n+1} - y_n$. If S_n is the sequence of partial sums of $\sum_{n=0}^{\infty} x_n$, then

$$S_n = x_1 + x_2 + \cdots + x_n = (y_2 - y_1) + (y_3 - y_2) + \cdots + (y_{n+1} - y_n) = y_{n+1} - y_1.$$

This implies that $\sum_{n=0}^{\infty} x_n$ converges if and only if the sequence (y_n) converges. In this case, $\sum_{n=0}^{\infty} x_n = \lim_{n\to\infty} y_n - y_1$.

(4) Consider $\sum_{n=0}^{\infty} \frac{n}{n^4+n^2+1}$. Observe that $x_n = \frac{n}{n^4+n^2+1} = \frac{1}{2}\left[\frac{1}{n^2-n+1} - \frac{1}{n^2+n+1}\right] = \frac{1}{2}(y_n - y_{n+1})$, where $y_n = \frac{1}{2}\left(\frac{1}{n^2-n+1}\right)$. Hence, $S_n = \frac{1}{2} - \left(\frac{1}{n^2-n+1}\right) \to \frac{1}{2}$.

Necessary condition for convergence

Theorem 3 (The nth Term Test)**.** *If $\sum_{n=0}^{\infty} x_n$ converges, then $x_n \to 0$.*

Proof. Let $\sum_{n=0}^{\infty} x_n = S$. Then $S_{n+1} - S_n = x_{n+1} \to S - S = 0$. $\square$

Examples.

(1) If $|x| \geq 1$, the the geometric series $\sum_{n=0}^{\infty} x^n$ diverges because $x^n \nrightarrow 0$.

(2) $\sum_{n=0}^{\infty} \sin n$ diverges because $\sin n \nrightarrow 0$.

(3) $\sum_{n=0}^{\infty} \log(\frac{n+1}{n})$ diverges, however, $\log(\frac{n+1}{n}) \to 0$.

Necessary and sufficient condition for convergence

Theorem 4. *Suppose $x_n \geq 0$ for all n. Then $\sum_{n=0}^{\infty} x_n$ converges iff (S_n) is bounded above.*

Proof. Exercise. □

Example. The Harmonic series $\sum_{n=0}^{\infty} \frac{1}{n}$ diverges. To see this, we will show that the sequence of partial sums (S_n) is not bounded above. It is enough to show that the subsequence S_{2^k} is not bounded above (Why?). Observe that

$$
\begin{aligned}
S_1 &= 1 \\
S_2 &= 1 + \frac{1}{2} \\
S_{2^2} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) \\
&> 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) \\
&= 1 + \frac{1}{2} + 2 \cdot \frac{1}{4} = 1 + 2 \cdot \frac{1}{2} \\
S_{2^3} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) \\
&> 1 + \frac{1}{2} + 2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{2^3} = 1 + 3 \cdot \frac{1}{2}.
\end{aligned}
$$

This implies that

$$\boxed{S_{2^k} > 1 + k \cdot \frac{1}{2}.}$$

Thus, (S_n) is not bounded above, and hence the series $\sum_{n=0}^{\infty} \frac{1}{n}$ diverges.

Let's see another proof. Assume that (S_n) is convergent. Then (S_n) is Cauchy. It follows that for $\varepsilon = \frac{1}{2}$, there exists $N \in \mathbb{N}$ such that $|S_{2m} - S_m| < \frac{1}{2}$ for all $m \geq N$. But

$$|S_{2m} - S_m| = \frac{1}{m+1} + \frac{1}{m+2} + \cdots + \frac{1}{2m} > \frac{1}{2m} + \frac{1}{2m} + \cdots + \frac{1}{2m} > \frac{1}{2}.$$

This is a contradiction. Hence, (S_n) cannot converge.

Algebra of Convergent Series

Given two series (whether or not convergent) $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$, and a scalar $\alpha \in \mathbb{R}$, we define

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

The following theorem shows that the set of all convergent series is a vector space over $\mathbb{R}$. The proof is straightforward and you should go for it.

Theorem 5. *Let $\sum_{n=1}^{\infty} x_n = x$ and $\sum_{n=1}^{\infty} y_n = y$. Then*

(1) *Their sum* $\sum_{n=1}^{\infty}(x_n + y_n)$ *is convergent and* $\sum_{n=1}^{\infty}(x_n + y_n) = x + y$.

(2) *The series* $\alpha \sum_{n=1}^{\infty} x_n$ *is convergent and* $\alpha \sum_{n=1}^{\infty} x_n = \alpha x$.

We now give an important result about absolutely convergent series.

Theorem 6. *An absolutely convergent series is convergent.*

Proof. Let S_n and Γ_n denote the partial sums of $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} |x_n|$ respectively. For $n > m$, we have

$$|S_n - S_m| = |\sum_{k=m+1}^{n} x_k| \leq \sum_{k=m+1}^{n} |x_k| = \Gamma_n - \Gamma_m,$$

which converge to 0 as (Γ_n) is convergent. Therefore, (S_n) is Cauchy. □

We will see later that the converse of the above result does not hold.

14

CONVERGENCE TESTS FOR SERIES

In this lecture we will give several tests to determine the convergence of a series.

Theorem 1 (Comparison Test). *Suppose $|a_n| \le b_n$ for all $n \ge k$ for some k.*

(1) If $\sum b_n$ is convergent, then $\sum a_n$ is absolutely convergent and $|\sum a_n| \le \sum b_n$.

(2) If $\sum |a_n|$ is divergent, then $\sum b_n$ is also divergent.

Examples.

(1) $\sum \frac{1}{n!}$ is convergent as $n^2 < n!$ for $n \ge 4$.

(2) $\sum \frac{1}{\sqrt{n}}$ diverges because $\frac{1}{n} \le \frac{1}{\sqrt{n}}$.

(3) $\sum \frac{2^n+n}{3^n+n}$ converges as $\frac{2^n+n}{3^n+n} \le \frac{2^n+2^n}{3^n} = 2(\frac{2}{3})^n$.

Theorem 2 (Limit Comparison Test). *Suppose $a_n, b_n > 0$ eventually (i.e., $\exists\ k \in \mathbb{N}$ such that $a_n, b_n > 0\ \forall\ n \ge k$), and $\frac{a_n}{b_n} \to \ell$ as $n \to \infty$. Then*

(1) If $\ell > 0$, then $\sum a_n$ is convergent if and only if $\sum b_n$ is convergent.

(2) If $\ell = 0$ and $\sum b_n$ converges, then $\sum a_n$ also converges.

(3) If $\ell = \infty$ and $\sum b_n$ diverges, then $\sum a_n$ also diverges.

Examples. $\to a_n$

(1) $\sum \frac{2^n+n}{3^n-n}$ converges. If we take $b_n = (\frac{2}{3})^n$, then $\frac{a_n}{b_n} \to 1$.

(2) $\sum \sin(\frac{1}{n})$. Take $b_n = \frac{1}{n}$.

(3) $\sum \frac{1}{(\log n)^p}$ is divergent for $p > 0$. Let $b_n = \frac{1}{n}$. Then $\frac{a_n}{b_n} = \frac{1/(\log n)^p}{1/n} = \frac{n}{(\log n)^p} \to \infty$. ?

Theorem 3 (Cauchy condensation test). *Let $a_n \ge 0$ and $a_{n+1} \le a_n$. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges.*

Examples.

(1) $\sum \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \le 1$ (Verify).

(2) $\sum \frac{1}{n(\log n)^p}$ converges if $p > 1$ and diverges if $p \le 1$.

Theorem 4 (Cauchy's Root Test). *(1) If $|a_n|^{1/n} \le \alpha$ eventually for some $\alpha < 1$, then $\sum a_n$ is absolutely convergent.*

(2) If $|a_n|^{1/n} \ge 1$ for infinitely many n, then $\sum a_n$ is divergent.

(3) In particular, if $|a_n|^{1/n} \to \ell$ where $\ell \in \mathbb{R}$ or $\ell = \infty$, then $\sum a_n$ is absolutely convergent when $\ell < 1$, and it is divergent when $\ell > 1$.

Examples.

(1) $\sum \frac{1}{(\log n)^n}$ converges because $a_n^{1/n} = \frac{1}{\log n} \to 0$.

(2) $\sum (\frac{n}{n+1})^{n^2}$ converges as $a_n^{1/n} = \frac{1}{(1+\frac{1}{n})^n} \to \frac{1}{e} < 1$.

Theorem 5 (D'Alembert's Ratio Test). *Suppose $a_n \neq 0$ for all n.*

(1) If $|\frac{a_{n+1}}{a_n}| \leq \alpha$ eventually for some $\alpha < 1$, then $\sum a_n$ is absolutely convergent.

(2) If $|\frac{a_{n+1}}{a_n}| \geq 1$ eventually, then $\sum a_n$ is divergent.

(3) In particular, if $|\frac{a_{n+1}}{a_n}| \to \ell$ where $\ell \in \mathbb{R}$ or $\ell = \infty$, then

$\sum a_n$ is absolutely convergent when $\ell < 1$, and it is divergent when $\ell > 1$.

Remark 6. *Both Root test and Ratio test are inconclusive if $\ell = 1$.*

Examples.

(1) $\sum \frac{1}{n!}$ converges because $\frac{a_{n+1}}{a_n} \to 0$.

(2) $\sum (\frac{n^n}{n!})^{n^2}$ diverges as $\frac{a_{n+1}}{a_n} = (1 + \frac{1}{n})^n \to e > 1$.

Theorem 7 (Dirichlet's Test). *Let (a_n) and (b_n) be sequences such that (a_n) is monotonic, $a_n \to 0$, and the sequence (B_n) defined by $B_n = \sum_{k=1}^n b_k$ is bounded. Then the series $\sum a_n b_n$ is convergent.*

Corollary 8 (Leibniz Test). *Let (a_n) be a monotonic sequence such that $a_n \to 0$. Then $\sum (-1)^{n-1} a_n$ is convergent.*

Examples. For $p > 0$, both the series $\sum \frac{(-1)^{n-1}}{n^p}$ and $\sum \frac{(-1)^{n-1}}{(\log n)^p}$ are convergent.

Rearrangements of a Series

Let $\sum a_n$ and a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ be given. Define $b_n = a_{\sigma(n)}$. Then the new series $\sum b_n$ is said to be a rearrangement of the series $\sum a_n$.

Theorem 9 (Riemann's Theorem). *A conditionally convergent series can be made to converge to any arbitrary real number or even made to diverge by a suitable rearrangement of its terms.*

The above theorem should convince us the danger of manipulating a series without paying attention to rigorous analysis. The next theorem tells us when we can manipulate the terms of the series whichever way we want.

Theorem 10 (Rearrangement of Terms). *A series $\sum a_n$ is absolutely convergent if and only if every rearrangement of it is convergent. In this case, the sum of a rearrangement is unchanged.*

15

POWER SERIES

In this lecture we will define a class of functions that are very important in Calculus. So far, you have encountered many transcendental functions[1] such as trigonometric, logarithmic and exponential functions. You also use to write such functions in the form of series since your school days. This series is known as Taylor series. We will see that many classical functions admits a Taylor series and we will study its convergence.

Definition 1. *A power series around a is an expression of the form $\sum_{n=0}^{\infty} a_n(x-a)^n$ where $a_n, a, x \in \mathbb{R}$.*

If we let $\tilde{x} = x - a$, then the power series around a can be reduced to a power series around 0. We are interested in finding $x \in \mathbb{R}$ for which the series $\sum_{n=0}^{\infty} a_n x^n$ is convergent.

Consider the following three power series.

(1) $\sum_{n=1}^{\infty} n^n x^n$. This series converges only at $x = 0$. Indeed, if $x \neq 0$, we can choose $N \in \mathbb{N}$ such that $\frac{1}{N} < |x|$. Thus, for all $n \geq N$, we have $|(nx)^n| > 1$ and hence the series is divergent.

(2) $\sum_{n=0}^{\infty} x^n$. We have seen that this series converges absolutely for $|x| < 1$.

(3) $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. Using ratio test we can show that this series converges absolutely for all $x \in \mathbb{R}$.

The general phenomenon is given in the following lemma.

Lemma 2 (Abel's Lemma)**.** *Let $x_0 \in \mathbb{R}$ such that $\{a_n x_0^n : n \in \mathbb{N}\}$ is bounded. Then $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent for every $x \in \mathbb{R}$ with $|x| < |x_0|$. In particular, if $\sum_{n=0}^{\infty} a_n x_0^n$ is convergent, then $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent for all x such that $|x| < |x_0|$.*

[1] A function $f : D(\subseteq \mathbb{R}) \to \mathbb{R}$ is said to be an algebraic function if $y = f(x)$ satisfies an equation whose coefficients are polynomials, i.e., $p_n(x)y^n + p_{n-1}(x)y^{n-1} + \cdots + p_1(x)y + p_0(x) = 0$ for $x \in D$, where $n \in \mathbb{N}$ and $p_0(x), p_1(x), \ldots, p_n(x)$ are polynomials such that $p_n(x)$ is nonzero. For example, $y = f(x) = \sqrt[n]{x}$ is an algebraic function since y satisfies the equation $y^n - x = 0$ for $x \in [0, \infty)$. A real-valued function that is not algebraic is called a transcendental function. A real number α is called an algebraic number if it satisfies a nonzero polynomial with integer coefficients. Numbers that are not algebraic are called transcendental numbers. For example, $\sqrt{2}, \sqrt{3}, \sqrt[5]{7}$ and $\sqrt{2} + \sqrt{3}$ are algebraic numbers. Any rational number is algebraic. The numbers e and π are transcendental. For more details, see [1] and [2].

Theorem 3. *A power series $\sum_{n=0}^{\infty} a_n x^n$ is either absolutely convergent for all $x \in \mathbb{R}$ or there is a unique nonnegative real number R such that the series is absolutely convergent for all x with $|x| < R$ and is divergent for $|x| > R$.*

Definition 4. *We say that the* **radius of convergence** *of a power series is ∞ if the power series is absolutely convergent for all $x \in \mathbb{R}$; otherwise, it is defined to be the unique nonnegative real number R such that the power series is absolutely convergent for all x with $|x| < R$ and is divergent for $|x| > R$.*

Let $S = \{x \in \mathbb{R} : \sum_{n=0}^{\infty} a_n x^n \text{ is convergent }\}$. Then $S \neq \emptyset$ (Why?).

Exercise. The radius of convergence and the set S for various power series are given below. Verify?

Power Series	Radius of convergence	S
$\sum_{n=1}^{\infty} n^n x^n$	0	$\{0\}$
$\sum_{n=1}^{\infty} x^n/n!$	∞	$(-\infty, \infty)$
$\sum_{n=0}^{\infty} x^n$	1	$(-1, 1)$
$\sum_{n=0}^{\infty} x^n/n^2$	1	$[-1, 1]$
$\sum_{n=0}^{\infty} x^n/n$	1	$[-1, 1)$
$\sum_{n=0}^{\infty}(-1)^n x^n/n$	1	$(-1, 1]$

The following result is useful in calculating the radius of convergence of a power series.

Proposition 5. *Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series and R be its radius of convergence.*

(1) If $|a_n|^{1/n} \to \ell$, then

$$R = \begin{cases} 0 & \text{if } \ell = \infty, \\ \infty & \text{if } \ell = 0, \\ 1/\ell & \text{if } \ell > 0. \end{cases}$$

(2) If $a_n \neq 0$ eventually and $|\frac{a_{n+1}}{a_n}| \to \ell$, then the above conclusion holds.

Taylor and Maclaurin Series

Let $f : [a, b] \to \mathbb{R}$ be an infinitely differentiable function and $c \in (a, b)$. In one of previous lectures, we have defined the Taylor polynomial of order n generated by f at c as the polynomial

$$P_n(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k.$$

The power series

$$f(c)+f'(c)(x-c)+\frac{f''(c)}{2!}(x-c)^2+\cdots+\frac{f^{(n)}(c)}{n!}(x-c)^n+\cdots=\sum_{n=0}^{\infty}\frac{f^{(n)}(c)}{n!}(x-c)^n$$

is called the Taylor series of f around c. If $c=0$, then the Taylor series of f around c is called the **Maclaurin series** of f.

Let $R_n(x)=f(x)-P_n(x)$ for $x\in[a,b]$. If $R_n(x)\to 0$ for some $x\in[a,b]$, then the nth partial sum $P_n(x)$ of the Taylor series of f around c converges to $f(x)$, i.e., $f(x)$ is the sum of the Taylor series.

Now, the following two questions arise.

- At what points $x\in[a,b]$, the Taylor series of f around c converge?
- If the Taylor series of f converges for some x, does it converge to $f(x)$?

If $x=c$, then $P_n(c)=f(c)$ for all $n=0,1,\dots$, and hence, $R_n(c)=0$. Therefore, the Taylor series of f converges to $f(c)$ at $x=c$. Consider the following examples.

(1) Let $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=|x|$. Then the Taylor series of f around 1 is $1+(x-1)$. But this is not equal to $f(x)$ when $x<0$.

(2) As an extreme case, consider $f:\mathbb{R}\to\mathbb{R}$ defined by

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

Show that f is infinitely differentiable on $\mathbb{R}$ with $f^{(n)}(0)=0$ for all $n=0,1,\dots$. Thus, the Maclaurin series of f is identically zero and it does not converge to $f(x)$ at $x\neq 0$.

Taylor's theorem helps us to show the convergence of a Taylor series of f to $f(x)$ in the following way. Recall that Taylor's theorem says that for $x\in[a,b]$ with $x\neq c$ and each $n\in\mathbb{N}$,

$$R_n(x)=\frac{f^{(n+1)}(c_{x,n})}{(n+1)!}(x-c)^{n+1}\text{ for some } c_{x,n} \text{ between } c \text{ and } x.$$

It is clear that if $R_n(x)\to 0$, then the Taylor series of f around c converges to $f(x)$.

Problem 6. *Show that the Maclaurin series of $f(x)=e^x$ converges to $f(x)$ for all $x\in\mathbb{R}$.*

Solution. By Taylor's theorem, there exists $c_{x,n}$ between 0 and x such that

$$e^x=1+x+\frac{x^2}{2!}+\cdots+\frac{e^{c_{n,x}}}{(n+1)!}x^{n+1}.$$

Since e^x is increasing, $e^{c_{n,x}}$ lies between $e^0=1$ and e^x. If $x<0$, then $c_{n,x}<0$, and $e^{c_{n,x}}<1$. If $x=0$, then $e^x=1$ and $R_n(x)=0$. If $x>0$, then $c_{n,x}>0$ and $e^{c_{n,x}}<e^x$.

Thus,

$$\left|\frac{e^{c_{n,x}}}{(n+1)!}x^{n+1}\right| \le \frac{|x|^{n+1}}{(n+1)!} \text{ if } x \le 0, \text{ and } \left|\frac{e^{c_{n,x}}}{(n+1)!}x^{n+1}\right| \le e^x \frac{x^{n+1}}{(n+1)!} \text{ if } x > 0.$$

Finally, $\lim_{n\to\infty} \frac{x^{n+1}}{(n+1)!} = 0$ for every x. Hence, $\lim_{n\to\infty} R_n(x) = 0$. Therefore, the Maclaurin series of e^x converges to e^x for every x. Thus,

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

□

Exercise 7. *Show that for every $x \in \mathbb{R}$,*

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

and

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots .$$

References

[1] A. Baker, *Transcendental Number Theory*, Cambridge University Press, Cambridge, 1975.

[2] G. Chrystal, *Algebra: An Elementary Text-book for the Higher Classes of Secondary Schools and for Colleges*, sixth ed., parts I and II, Chelsea, New York, 1959.

RIEMANN INTEGRATION

The Riemann Integral

The notion of integration was developed much earlier than differentiation. The main idea of integration is to assign a real number A, called the "area", to the region bounded by the curves $x = a$, $x = b$, $y = 0$, and $y = f(x)$. To proceed formally, we introduce the following concept.

Definition 1. *By a partition P of $[a, b]$ we mean a finite ordered set $\{x_0, x_1, \ldots, x_n\}$ of points in $[a, b]$ such that*

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

Examples.

(1) The simplest partition of $[a, b]$ is given by $P_1 = \{a, b\}$.

(2) For $n \in \mathbb{N}$, let $P_n = \{x_0, x_1, \ldots, x_n\}$, where

$$x_i = a + \frac{i(b-a)}{n}, \quad \text{for } i = 0, 1, \ldots, n.$$

Then P_n is a partition that subdivides the interval $[a, b]$ into n subintervals, each of length $(b - a)/n$. What happens when n becomes larger?

Let $f : [a, b] \to \mathbb{R}$ be a bounded function, and $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. Define

$$m = \inf\{f(x) : x \in [a, b]\} \text{ and } M = \sup\{f(x) : x \in [a, b]\}.$$

For $i = 1, 2, \ldots, n$,

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\} \text{ and } M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}.$$

Clearly,

$$m \le m_i \le M_i \le M \text{ for all } i = 1, 2, \ldots, n. \text{ (Prove!)}$$

Definition 2. *Given a bounded function $f : [a, b] \to \mathbb{R}$ and a partition $P : \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$, the lower Riemann[1] sum of f with respect to the partition P is defined as*

$$L(P, f) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}).$$

[1] Bernhard Riemann was a German mathematician who made contributions to analysis, number theory, and differential geometry. A work which Riemann did in 1859 is referred to as the Riemann hypothesis. Anyone who solves the Riemann hypothesis will earn a **million dollar**!

The upper Riemann sum of f with respect to the partition P is defined as

$$U(P,f)=\sum_{i=1}^{n} M_i(x_i-x_{i-1}).$$

Lemma 3. *Let $f:[a,b]:\to\mathbb{R}$ be a bounded function. Then for any partition P of $[a,b]$, we have*

$$m(b-a)\le L(P,f)\le U(P,f)\le M(b-a).$$

Proof. Exercise. □

Definition 4. *The upper and lower Riemann integral of f over $[a,b]$ is respectively defined as*

$$L(f)=\underline{\int_a^b} f(x)\mathrm{d}x=\sup\{L(P,f): P \textit{ is a partition of} [a,b]\},$$

and

$$U(f)=\overline{\int_a^b} f(x)\mathrm{d}x=\inf\{L(P,f): P \textit{ is a partition of} [a,b]\}.$$

Does these two numbers exist? If the upper and lower Riemann integrals are equal, we say that f is Riemann integrable or simply integrable. In this case, the common value of $L(f)=U(f)$ is called the Riemann integral of f (on $[a,b]$) and is denoted by

$$\int_a^b f(x)\mathrm{d}x \textit{ or simply } \int_a^b f.$$

Examples.

(1) Let $f:[a,b]:\to\mathbb{R}$ be a constant function with $f(x)=c$ for all x. Let P be any partition of $[a,b]$. Then $m_i=M_i=c$ for all i and $L(P,f)=U(P,f)=c(b-a)$. Thus, $\underline{\int_a^b} f(x)\mathrm{d}x=c(b-a)=\overline{\int_a^b} f(x)\mathrm{d}x$, and hence f is Riemann integrable. This implies that $\int_a^b f(x)\mathrm{d}x=c(b-a)$.

(2) Let $\lambda>1$ be a real number. Let $f:[0,1]:\to\mathbb{R}$ be defined by

$$f(x)=\begin{cases}1, & \text{if } 0\le x<1,\\ \lambda, & \text{if } x=1.\end{cases}$$

Let us find the upper and lower Riemann integrals. Let $P=\{x_0,\dots,x_n\}$ be a partition of $[0,1]$. We observe that $m_i=1$ for $i=1,\dots,n$. Moreover, $M_i=1$ for $i=1,\dots,n-1$ and $M_n=\lambda$. It follows that

$$L(P,f)=1 \text{ and } \underline{\int_0^1} f(x)\mathrm{d}x=1.$$

On the other hand, $U(P,f)=x_{n-1}+\lambda(1-x_{n-1})$ and

$$\overline{\int_0^1} f(x)\mathrm{d}x=\inf\{U(P,f): P \text{ is a partition of } [0,1]\}=\inf(1,\lambda]=1 \text{ (Verify!)}.$$

This shows that f is integrable and $\int_0^1 f(x)\mathrm{d}x=1$.

(3) Consider the Dirichlet's function, i.e., the function defined by $f : [0, 1] \to \mathbb{R}$ defined by $f(x) = 1$ if $x \in \mathbb{Q}$ and 0 otherwise. Let $P = \{x_0, \dots, x_n\}$ be a partition of $[0, 1]$. In any subinterval $[x_{i-1}, x_i]$, there exist some rational numbers and irrational numbers. Hence, $m_i = 0$ and $M_i = 1$. So, $L(P, f) = 0$ and $U(P, f) = 1$. It follows that $\underline{\int_a^b} f(x)\mathrm{d}x = 0$ and $\overline{\int_a^b} f(x)\mathrm{d}x = 1$. Thus, f is not integrable on $[0, 1]$.

Integrable Functions

Definition 5. *Given a partition P of $[a, b]$, we say that a partition P^* of $[a, b]$ is a refinement of P if $P \subset P^*$. Given partitions P_1 and P_2 of $[a, b]$, the partition $P^* = P_1 \cup P_2$ is called the common refinement of P_1 and P_2.*

Proposition 6. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then we have the following:*

(1) If P is partition of $[a, b]$, and P^* is a refinement of P, then

$$L(P, f) \le L(P^*, f) \text{ and } U(P^*, f) \le U(P, f),$$

and consequently,

$$U(P^*, f) - L(P^*, f) \le U(P, f) - L(P, f).$$

(2) If P_1 and P_2 are partitions of $[a, b]$, then $L(P_1, f) \le U(P_2, f)$.

(3) $\underline{\int_a^b} f(x)\mathrm{d}x \le \overline{\int_a^b} f(x)\mathrm{d}x$

In the following result we present a necessary and sufficient condition for the existence of the integral of a bounded function.

Theorem 7 (Riemann's criterion for integrability)**.** *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. The f is integrable if and only if for every $\varepsilon > 0$, there is a partition P_ε of $[a, b]$ such that*

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

The proof of the following corollary is immediate from the previous theorem.

Corollary 8. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Suppose (P_n) is a sequence of partitions of $[a, b]$ such that $U(P_n, f) - L(P_n, f) \to 0$. Then f is integrable.*

Example. Let $f(x) = x^2$ on $[0, 1]$. Let $\varepsilon > 0$ be given. Choose a partition P such that $\max\{x_i - x_{i-1} : 1 \le i \le n\} < \varepsilon/2$. Since f is increasing, we have

$$m_i = f(x_{i-1}) = x_{i-1}^2 \text{ and } M_i = f(x_i) = x_i^2.$$

This implies that

$$\begin{aligned} U(P,f)-L(P,f) &= \sum_{i=1}^{n} x_i^2(x_i-x_{i-1}) - \sum_{i=1}^{n} x_{i-1}^2(x_i-x_{i-1}) \\ &= \sum_{i=1}^{n}(x_i-x_{i-1})(x_i+x_{i-1})((x_i-x_{i-1}) \\ &< \sum_{i=1}^{n}\left[\left(\frac{\varepsilon}{2}\right)\times 2\right](x_i-x_{i-1}), \text{ since } 0\leq, x_{i-1}, x_i \leq 1 \\ &= \varepsilon\sum_{i=1}^{n}(x_i-x_{i-1}) = \varepsilon. \end{aligned}$$

Hence, f is integrable.

We will apply the Riemann's criterion for integrability to prove the following theorem.

Theorem 9. *Let $f : [a,b] \to \mathbb{R}$ be a function.*

(1) If f is monotone, then it is integrable.
(2) If f is continuous, then it is integrable.

Proof. We proof part (1). The the proof of part (2) is left to the reader.

Suppose f is monotonically increasing (the proof is similar in the other case). Choose a partition P such that $x_i - x_{i-1} = \frac{b-a}{n}$ for each i. Then $M_i = f(x_i)$ and $m_i = f(x_{i-1})$. Thus, for large n, we have

$$U(P,f)-L(P,f) = \frac{b-a}{n}\sum_{i=1}^{n}(f(x_i)-f(x_{i-1})) = \frac{b-a}{n}\sum_{i=1}^{n}(f(b)-f(a)) < \varepsilon.$$

Hence, f is integrable. □

We end this lecture with the following problem. We encourage students to understand and verify each step.

Problem 10. *Let $f : [0,1] \longrightarrow \mathbb{R}$ be defined as*

$$f(x) = \begin{cases} x, & x \text{ is rational} \\ 0, & x \text{ is irrational.} \end{cases}$$

Evaluate the upper and lower integrals of f and show that f is not integrable.

Solution. Let $P = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[0,1]$. Since there exists an irrational number in each subinterval $[x_{i-1}, x_i]$, $L(P,f) = 0$, and hence $\underline{\int_0^1} f(x)\mathrm{d}x = 0$.

Now,

$$
\begin{aligned}
U(P,f) &= \sum_{i=1}^{n} x_i(x_i - x_{i-1}) \\
&= \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_{i-1}x_i \\
&\geq \sum_{i=1}^{n} x_i^2 - \frac{1}{2}\sum_{i=1}^{n}(x_{i-1}^2 + x_i^2) \text{ (Using AM-GM inequality)} \\
&= \frac{1}{2}\sum_{i=1}^{n}(x_{i-1}^2 - x_i^2) = \frac{1}{2}. \\
\implies \overline{\int_0^1} f(x)\mathrm{d}x &\geq \frac{1}{2}.
\end{aligned}
$$

For each $n \in \mathbb{N}$, consider $P_n = \left\{0, \frac{1}{n}, \frac{2}{n}, \cdots, \frac{n-1}{n}, \frac{n}{n} = 1\right\}$.

Then

$$
\begin{aligned}
U(P_n, f) = \frac{n+1}{2n} &= \frac{1}{2} + \frac{1}{2n} \\
\implies \quad & \inf\left\{U(P_n, f) : n \in \mathbb{N}\right\} = \frac{1}{2} \\
\implies \quad & \overline{\int_0^1} f(x)\mathrm{d}x \leq \frac{1}{2}.
\end{aligned}
$$

Therefore, $\overline{\int_0^1} f(x)\mathrm{d}x = \frac{1}{2}$. $\square$