REAL ANALYSIS I

Edited By
Dr. Sachin Kaushal

Printed by
EXCEL BOOKS PRIVATE LIMITED
A-45, Naraina, Phase-I,
New Delhi-110028
for
Lovely Professional University
Phagwara

SYLLABUS

Real Analysis I

Sr. No.	Content
1	Set Theory Finite, Countable and Uncountable Sets, Metric spaces ;Definition and examples
2	Compactness of k-cells and Compact Subsets of Euclidean, Space R^k, Perfect sets and Cantor's set, Connected sets in a metric space, Connected subset of Real line
3	Sequences I Metric Spaces, Convergent sequences and Subsequences, Cauchy sequence, complete metric space, Cantor's intersection theorem and Baire's Theorem, Branch contraction Principle.
4	Limit of functions, continuous functions, Continuity and compactness, continuity and connectedness, Discontinuities and Monotonic functions
5	Sequences and series; Uniform convergence, Uniform convergence and continuity, Uniform convergence and integration

CONTENT

Unit 1:	Sets and Numbers *Sachin Kaushal, Lovely Professional University*	1
Unit 2:	Algebraic Structure and Countability *Sachin Kaushal, Lovely Professional University*	22
Unit 3:	Matric Spaces *Sachin Kaushal, Lovely Professional University*	43
Unit 4:	Compactness *Sachin Kaushal, Lovely Professional University*	67
Unit 5:	Connectedness *Richa Nandra, Lovely Professional University*	73
Unit 6:	Completeness *Richa Nandra, Lovely Professional University*	78
Unit 7:	Convergent Sequence *Richa Nandra, Lovely Professional University*	84
Unit 8:	Completeness and Compactness *Sachin Kaushal, Lovely Professional University*	90
Unit 9:	Functions *Sachin Kaushal, Lovely Professional University*	100
Unit 10:	Limit of a Function *Sachin Kaushal, Lovely Professional University*	129
Unit 11:	Continuity *Richa Nandra, Lovely Professional University*	148
Unit 12:	Properties of Continuous Functions *Richa Nandra, Lovely Professional University*	159
Unit 13:	Discontinuities and Monotonic Functions *Richa Nandra, Lovely Professional University*	171
Unit 14:	Sequences and Series of Functions *Richa Nandra, Lovely Professional University*	179
Unit 15:	Uniform Convergence of Functions *Richa Nandra, Lovely Professional University*	193
Unit 16:	Uniform Convergence and Continuity *Sachin Kaushal, Lovely Professional University*	204
Unit 17:	Uniform Convergence and Differentiability *Sachin Kaushal, Lovely Professional University*	211

Unit 1: Sets and Numbers

Notes

CONTENTS

Objectives
Introduction
1.1 Sets and Functions
1.1.1 Sets
1.1.2 Functions
1.2 System of Real Numbers
1.2.1 Natural Numbers
1.2.2 Integers
1.2.3 Rational Numbers
1.3 The Real Line
1.4 Complex Numbers
1.5 Mathematical Induction
1.6 Summary
1.7 Keywords
1.8 Review Questions
1.9 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss the basic concepts of sets and functions
- Explain the system of real numbers
- Describe the representation of real numbers

Introduction

As we know that one of the main features of Mathematics is the identification of the subject matter, its analysis and its presentation in a satisfactory manner. In other words, the language should be a vehicle which carries ideas through the mind without affecting their meaning in any way. Set Theory comes closest to being such a language. Introduced between 1873 and 1895 by a famous German mathematician, George Cantor (1845-1918), Set Theory became the foundation of almost all the branches of Mathematics. Besides its universal appeal, it is quite amazing in its simplicity and elegance.

A rigorous presentation of Set Theory is not the purpose here because we believe that you are already familiar with it. We shall briefly recall some of its basic concepts which are essential for a systematic study of Real Analysis. Closely linked with the sets, is the notion of a function, which also you have learnt in your previous studies. In this unit, we shall review this as well as other related concepts which are necessary for our discussion.

Notes

'Real Analysis' is an important branch of Mathematics which mainly deals with the study of real numbers. What is, then, the system of the real numbers? We shall try to find an answer to this question as well as some other related questions in this unit. Also, we shall give the geometrical representation of the real numbers.

1.1 Sets and Functions

As you all know modern Mathematics is based on the ideas that are expressed in the language of sets and functions. Here you set knowledge of certain basic concepts of Set Theory which are quite familiar to you. These concepts will also serve an important purpose of recalling certain notations and terms that will be used throughout our discussion with you.

1.1.1 Sets

As you are used to the phrases like the 'team' of cricket players, the 'army' of a country, the 'committee' on the education policy, the 'panchayat' of a village, etc. The terms 'team', 'army', 'committee', panchayat', etc., indicate the notion of a 'collection' or 'totality' or 'aggregate' of objects. These are well-known examples of a set.

Therefore, our starting point is an informal description of the term 'set'. A set is treated as an undefined term just as a point in Geometry is undefined. However, for our purpose we say that a set is a well-defined collection of objects. A collection os well-defined of it is possible to say whether a given object belongs to the collection or not.

The following are some examples of sets:

1. The collection of the students registered in Excel Books.
2. The collection of the planets namely Jupiter, Saturn, Earth, Pluto, Venus, Mercury, Mars, Uranus and Neptune.
3. The collection of all the countries in the world. (Do you know how many countries are there in the world?)
4. The collection of numbers, 1, 2, 3, 4, ……

If we consider the collection of tall persons or beautiful ladies or popular leaders, then these collections are not well-defined and hence none of them forms a set. The reason is that the words 'tal' 'beautiful' or 'popular' are not well-defined. The objects constituting a set are called its elements or members or points of the set. Generally, sets are denoted by the capital letters A, B, C etc. and the elements are denoted by the small letters a, b, c etc. If S is any set and x is an element of S, we express it by writing that $x \in S$, where the symbol $\in$ means 'belongs to' or 'is a member of'. If x is not an element an element of a set S, we write $x \notin S$. For example , if S is the set containing 1, 2, 3, 4 only, then $2 \in S$ and $5 \in S$.

You know that there are two method of describing a set. One is known as the Tabular method and the other is the Set-Builder method. In the tabular method we describe a set by actually listing all the elements belonging to it.

Example: If S is the set consisting of all small letters of English alphabet, then we write

$$S = \{a, b, c,...,x, y, z\}.$$

If N is the set of all natural numbers, then we write

$$N = \{1, 2, 3....\}.$$

Notes

This is also called an explicit representation of a set.

In the set-builder method, a set is described by stating the property which determines the set as a well-defined collection. Suppose p denotes this property and x is an element of a set S. Then

$$S = \{x: x \text{ satisfies } p).$$

Example: The two sets S and N can be written as

$$S = \{x: x \text{ is a small letter of English alphabet}\}$$

$$N = \{n: n \text{ is a natural number}).$$

This is also called an implicit representation of a set.

Note that in the representation of sets, the elements of a set are not repeated. Also, the elements may be listed in any manner.

Example: Write the set S whose elements are all natural numbers between 7 and 12 including both 7 and 12 in the tabular as well as in the set-builder forms.

Solution: Tabular form is $S = \{7, 8, 9, 10, 11, 12, \}$.

Set-builder form is $S = \{ n \in N: 7 \leq n \leq 12, \}$.

The following standard notations are used for the sets of numbers:

$$N = \text{Set of all natural numbers}$$

$$= \{1, 2, 3....\}$$

$$= \{n{:}n \text{ is a natural number})$$

$$= \text{Set of all positive integers.}$$

$$Z = \text{Set of all integers}$$

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

$$= \{p{:}p \text{ is an integer}).$$

$$Q = \text{Set of all rational numbers}$$

$$= \{x : x = \frac{P}{q}, p \in Z, q \in Z, q \neq 0).$$

$$R = \text{Set of real numbers}$$

$$= \{x : x \text{ is a real number}).$$

We shall, however, discuss the development of the system of real numbers.

A set is said to be finite if it has a finite number of elements. A set is said to be infinite if it is not finite. We shall, however, give a mathematical definition of finite and infinite sets in Unit 2.

Note that an element of a set must be carefully distinguished from the set consisting of this element. Thus, for instance, you must distinguish

$$x, \{x\}, \{\{x\}\}$$

from each other

We talk of equality of numbers, equality of objects, etc.

The question, therefore, arises: What is the notion of the equality of sets?

Notes

Definition 1: Equality of Sets

Any two sets are equal if that are identical. Thus the two sets S and T are equal, written as S = T if they consist of exactly the same elements. When the two sets S and T are unequal, we write

$$S \neq T.$$

It follows from the definition that S = T if any one of $x \in S$ implies $x \in T$ and $y \in T$ implies $y \in S$. Also S is different from T (S +T) if there is at least one element in one of them which is not in the other.

If every member of a given set S is also a member of T, then we say that S is a subset of T or "S is contained in T" and write:

$$S \subset T$$

or equivalently we say that "T contains S" or T is a superset of S, and write

$$T \supset S$$

The relation

$$S \not\subset T$$

means that S is not a subset of T i.e. there is at least one element in T which is not in S.

Thus, you can easily see that any two sets S and T are equal if and only if S is a subset of T and T is a subset of S i.e.

$$S = T \Leftrightarrow S \subset T \text{ and } T \subset S.$$

If $S \subset T$ but $T \not\subset S$, then we say that S is a proper subset of T. Note that $S \subset S$ i.e. every set is a subset of itself.

Another important concept is that of a set having no elements. Such a set, as you know, is called an empty set or a null set or a void set and is denoted by O.

You can easily see that there is only one empty set i.e. O is unique. Further O is a subset of every set.

Now why don't you try an exercise?

Task **Justify the following statements:**

1. The set N is a proper subset of Z.
2. The set R is not a subset of Q.
3. If A, B, C are any three sets such that $A \subset B$, and $B \subset C$, then $A \subset C$.

So far, we have talked about the elements and subsets of a given set. Let us now recall the method of constructing new sets from the given sets.

While studying subsets, we generally fix a set and consider the subsets of this set throughout our discussion. This set is usually called the Universal set. This Universal set may vary from situations to situations. For example, when we consider the subsets of R, then R is the Universal set. When we consider the set of points in the Euclidean plane, then the set of all points in the Euclidean plane is the Universal set. We shall denote the Universal set by X.

Notes

Now, suppose that the Universal set X is given as

$$X = \{1, 2, 3, 4, \}$$

and

$$S = \{1, 2, 3\}$$

is a subset of X. Consider a subset of X whose elements do not belong to S. This set is (4, 2}.

Such, a set, as you know is called the complement of S.

We define the complement of a set as follows:

Definition 2: Complement of a Set

Let X be the Universal set and S be a subset of X. The complement of the set S is the set of all those elements of the Universal set X which do not belong to S. It is denoted by S.

Thus, if S is an arbitrary set contained in the Universal Set X, then the complement of S is the set

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

Associated with each set S is the Power set P(S) of S consisting of all the subsets of S. It is written as

$$P(S) = \{A : A \subset S).$$

Now try the following exercise.

Let us consider the sets S and T given as

$$S = \{1, 2, 3, 4, 5\}, T = \{3, 4, 5, 6, 7\}.$$

Construct a new set {1, 2, 3, 4, 5, 6, 7). Note that all the elements of this set have been taken from S or T such that no element of S and T is left out. This new set is called the union of the sets S and T and is denoted by $S \cup T$.

Thus

$$S \cup T = \{1, 2, 3, 4, 5, 6, 7).$$

Again let us construct another set {3, 4, 5). This set consists of the elements that are common to both S and T i.e. a set whose elements are in both S and T. This set is called the intersection of S and T. It is denoted by $S \cap T$. Thus

$$S \cap T = \{3, 4, 5).$$

These notions of Union and Intersection of 'two sets' can be generalized for any sets in the following way: Note, all the sets under discussion will be treated as subsets of the Universal set X.

Definition 3: Union of Sets

Let S and T be given sets. The collection of all elements which belong to S or T is called the union of S and T. It is expressed as

$$S \cup T = \{x : x \in S \text{ or } x \in T\}.$$

Note that when we say that $x \in S$ or $x \in T$, then it means that x belong to S or x belong to T or x belong to both S and T.

Definition 4: Intersection of Sets

The intersection $S \cap T$ of the sets S and T is defined to be the set of all those elements which belong to both S and T i.e.

$$S \cap T = \{x : x \in S \text{ and } x \in T\}.$$

Notes

Note that the sets are disjoint or mutually exclusive when $S \cap T = 0$ i.e., when their intersection is empty.

You can now verify (or even prove) by means of examples the following laws of union and intersection of sets given in the next exercise.

Also, you can easily see that

$$A \cup A = A,\ A \cap A = A,\ A \cup \varnothing = A,\ A \cup \varnothing = \varnothing.$$

Given any two sets S and T, we can construct a new set in such a way that it contains only those elements of one of the sets which do not belong to the other. Such a set is called the difference of the given sets. There will be two such sets denoted by S—T and T—S. For example, let

$$S = \{2, 4, 8, 10, 11\},\ T = \{1, 2, 3.4).$$

Then

$$S - T = \{8, 10, 11),\ T - S = \{1, 3).$$

Thus, we can define the difference of two sets in the following way.

Definition 5: Difference of two Sets

Given two sets S and T, the difference S – T is a set consisting of precisely those members of S which are not in T.

Thus

$$S - T = \{x : x \in S \text{ and } x \notin T).$$

Similarly, we can define T – S.

Consider a collection of sets S_1, where i varies over some index set J. This simply means that to each element $i \in J$, there is a corresponding set S_i. For example, the collection $\{S_1, S_2, S_3,...)$ could be expressed as $\{S_i\}_{i \in N}$, where N is the index set.

With the introduction of an index set, the notions of the union and the intersection of sets can be extended to an arbitrary collection of sets. For example,

(i) $\bigcup_{i \in J} S_i = \bigcup_{i \in J} \{x : x \in S_i \text{ for at least one } i \in J\}.$

(ii) $\bigcap_{i \in J} S_i = \bigcap_{i \in J} \{x : x \in S_i \text{ for all } i \in J\}.$

(iii) $(\bigcup S_i)^c_{i \in J} = \bigcap_{i \in J} S_i^c.$

1.1.2 Functions

Let S be the set of Excel Books and let N be the set of all natural numbers. Assign to each book the number of pages the book contains. Here each book corresponds to a unique natural number. In other words, there is a correspondence between the books and the natural numbers, i.e., there is a rule or a mechanism by which we can associate to each book one and only one natural number. Such a rule or correspondence is named as a function or a mapping.

Definition 6: Function

Let S and T be any two non-empty sets. A function f from S to T denoted as $f: S \to T$ is a rule which assigns to each element of the set S, a unique element in the set T.

Notes

The set S is called the domain of the function f and T is called its co-domain. If an elements x in S corresponds to an element y in T under the function f, then y is called the image of x under f. This is expressed by writing $y = f(x)$. The set $\{f(x): x \in S\}$ which is a subset of T is called the range of f. If range of f = co-domain of f, then f is called onto or surjective function; otherwise f is called an into function.

Thus, a function $f: S \to T$ is said to be onto if the range of S is equal to its co-domain T.

Suppose $S = \{1, 2, 3, 4)$ and $T = \{1, 2, 3, 4, 5, 6)$ and $f: S \to T$ is defined by $f(n) = n+l$, $\forall\ n \in S$. Then the range of $f = \{2, 3, 4, 5)$. This shows that f is an into function. On the other hand, if $S = \{1, 2, 3, 4\}$, $T = \{1, 4, 9, 16\}$ and if $f: S \to T$ is defined by $f(n) = n^Z$, then f is onto. You can verify that the range of f is, in fact, equal to T.

Refer back to the example on the books in Excel Books. It just possible that two books may have the same number of pages. If it is so, then under this function, two different books shall have the same natural number as their image. However if for a function any two distinct elements in the domain have distinct images in the co-domain, then the function is called one-one or injective.

Thus a function f is said to be one-one if distinct elements in the domain of f have distinct image or in other words, if $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$, for any x_1, x_2 in the domain of f.

A function which is one-one and onto, is called a bijection or a 1-1 correspondence.

Example:

(i) Let $S = \{1, 2, 3)$ and $T = \{a, b, c\}$ and let $f: S \to T$ be defined as $f(1) = a$, $f(2) = b$, $f(3) = c$. Then f is one-one and onto.

(ii) Let $N = \{1, 2, 3, 4, ...\}$ and $f: N \to N$ be defined as $f(n) = n+1$. As 1 does not belong to the range of f, therefore f is not onto. However, f is one-one.

(iii) Let $S = (1, -1, 2, 3, -3)$ and let $T = (1, 4, 9)$. Define $f: S \to T$ by $f(n) = n^2\ \forall\ n \in S$. Then f is not one-one as $f(1) = f(-l) = 1$. However, f is onto.

Definition 7: Identity Function

Let S be any non-empty set. A function $f: S \to S$ defined by $f(x) = x$ for each x in S is called the identity function.

It is generally denoted by I_s. It is easy to see that I_s is one-one and onto.

Definition 8: Constant Function

Let S and T be any two non-empty sets. A function $f: S \to T$ defined by $f(x) = c$, for each x in S, where c is fixed element of T, is called a constant function.

For example $f: S \to R$ defined as $f(x)=2$, for every x in S, is a constant function. Is this function one-one and onto? Verify it.

Definition 9: Equality of Functions

Any two functions with the same domain are said to be equal if for each point of their domain, they have the same image. Thus if f and g are any two functions defined on an non-empty set S, then

$$f = g \text{ if } f(x) = g(x),\ \forall\ x \in S.$$

In other words, f = g if f and g are identical.

Notes

Let us now discuss another important concept in this section. This is about the composition or combination of two function. Consider the following situation:

Let S = {1, 2, 3, 4}, T = {1, 4, 9, 16}, N = {1, 2, 3, 4....} be any three sets Let f: S→T be defined by $f(x) = x^2$, $\forall\ x \in S$ and g: T→N be defined by g(x) = x +1, $\forall\ x \in T$. Then, by the function f, an element $x \in S$ is mapped to $f(x) = x^2$. Further by the function g the element f(x) is mapped to $f(x) + 1 = x^2 + 1$. Denote this as g(f(x)). Define a function h: S→N by h(x) = g(f(x)). This function h maps each x in S to some unique elements $g(f(x)) = x^2 + 1$ of N. The function h is called the composition or the composite of the functions f and g. Thus, we have the following definition:

Definition 10: Composite of Functions

Let f: S→T and g: T→V be any two functions. A function h: S→V denoted as h = gof and defined by

$$h(x) = (gof)\ (x) = g(f(x)),\ \forall\ x \in S$$

is called the composite of f and g.

Note that the domain of the composite function is the set S and its co-domain is the set V. The set T which contains the range of f is equal to the domain of g.

But in general, the composition of the two functions is meaningful whenever the range of the first is contained the domain of the second.

Example: Let S = T = {1, 2, 3, 4...}, Define

f(x) = 2x and g(x) = x + 5. Then

"gof is defined as (gof) (x) = g(f(x)) = g(2x) = 2x + 5.

Note that we can also define fog the composite of g and f. Here (fog) (x) = f(g(x)) = f(x + 5) = 2 (x + 5) = 2x = 10, Also (fog) (1) = 12 and (gof) (1) = 7. This shows that 'fog' need not be equal to 'gof'.

Let S = {1, 2, 3} and T = {a, b, c). Let f: S→T be f(1) = a, f(2) = b, f(3) = c. Define a function g: T→S as g(a) = 1, g(b) = 2 and g(c) = 3. Under the function g, the element f(x) in T is taken back to the element x in S. This mapping g is called the inverse of f and is given by g(f(x)) = x, for each in S. You may note that f(g(a)) = a, f(g(b)) = b and f(g(c)) = c. Thus, we have the following definition:

Definition 11: Inverse of a Function

Let S and T be two non-empty sets. A function f: S→T is said to be invertible if there exists a function g: T→S such that

$$(gof)(x) = x \text{ for each } x \text{ in } S,$$

and

$$(fog)(x) = x \text{ for each } x \text{ in } T.$$

In this case, g is said to be the inverse of f and we write it as $g = f^{-1}$.

Did u know? **Do all function possess inverses?**

No, all functions do not possess inverses. For example, let S = {1, 2, 3} and T = {a, b). If f: S→T is defined as f(1) = f(2) = a and f(3) = b, then f is not invertible. For, if g: S→T is inverse of f, then

(gof) (1) = g(f(l)) = g(a)

and (gof) (2) = g(f(2)) = g(a).

Therefore, 1 = 2, which is absurd.

Notes

This raises another question: Under what conditions a function is as an inverse? If a function f: S→T is one-one and onto, then it is invertible * conversely, if f is invertible, then f is both one-one and onto. Thus if a function is one-one and onto, then it must have an Inverse.

1.2 System of Real Numbers

You are quite familiar with some number systems and some of their properties. You will, perhaps recall the following properties:

(i) Any number multiplied by zero is equal to zero,

(ii) The product of a positive number with a negative number is negative,

(iii) The product of a negative number with a negative number is positive among takers.

To illustrate these properties, you will most likely use the natural numbers or integers or even rational numbers. The questions, which begin to arise are: What are these various types of numbers? What properties characterise the distinction between these various sets of numbers?

In this section, we shall try to provide answers to these and many other related questions. Since we are dealing with the course on Real Analysis, therefore we confine our discussion to the system of real numbers. Nevertheless, we shall make you peep into the realm of a still larger class of numbers, the so called complex numbers.

The system of real numbers has been evolved in different ways by different mathematicians. In the late 19th Century, the two famous German mathematicians Richard Dedekind [1815-1897] and George Cantor [1845-1918] gave two independent approaches for the construction of real numbers. During the same time, an Italian mathematician, G. Peano [1858-1932] defined the natural numbers by the well-known Peano Axioms. However, we start with the set of natural numbers as the foundation and build up the integers. From integers, we construct the rational numbers and then through the set of rational numbers, we reach the stage of real numbers. This development of number system culminates into the set of complex numbers. A detailed study of the system of numbers leads us to a beautiful branch of Mathematics namely. *The Number Theory*, which is beyond the scope of this course. However, we shall outline the general development of the system of the real numbers in this section. This is crucial to understand the characterization of the real numbers in terms of the algebraic structure to be discussed in Unit 2. Let us start our discussion with the natural numbers.

1.2.1 Natural Numbers

The notion of a number and its counting is so old that it is difficult to trace its origin. It developed much before the time of even the recorded history that its manner of development is based on conjectures and guesses. The mankind, even in the most primitive times, had some number sense. The man, at least, had the sense of recognizing 'more' and 'less', when some objects were added to or taken out from a small collection. Studies have shown that even some animals possess such a sense. With the gradual evolution of society, simple counting became imperative. A tribe had to count how many members it had, how many enemies and how many friends. A shepherd or a cowboy found it necessary to know if his flock of sheep or cows was decreasing or increasing in size. Various ways were evolved to keep such a count. Stones, pebbles, scratches on the ground, notches on a big piece of wood, small sticks, knots in a string or the fingers of hands were used for this purpose. As a result of several refinements of these counting methods, the numbers were expressed in the written symbols of various types called the digits. These digits were written differently according to the different languages and cultures of the societies. In the ultimate development, the numbers denoted by the digits 1, 2, 3, ... became universally acceptable and were named as natural numbers.

Notes

Different theories have been advanced about the origin and evolution of natural numbers. An axiomatic approach, as evolved by G. Peano, is often used to define the natural numbers. Some mathematicians like L. Kronecker [1823-1891] have remarked that the natural numbers are a creation of God while all else is the work of man.

However, we shall not go into the origin of the natural numbers. In fact, we accept that the natural numbers are a gift of nature to the mankind.

We denote the set of all natural numbers as

$$N = \{1,2, 3, ...\}.$$

One of the basic properties of these numbers is that there is a starting number 1. Then for each number there is a next number. This nextness property is an important idea that you may find fascinating with the natural numbers. You may think of any big natural number. Yet, you can always tell its next number. What's the next number after forty nine? After seventy seven? After one hundred twenty three? After three thousand and ninety nine? Thus you have an endless chain of natural numbers.

Some of the basic properties of the natural numbers are concerning the well-known fundamental operations of addition, multiplication, subtraction and division. You know that the symbol '+' is used for addition and the symbol 'x' is used for multiplication. If we add or multiply any two natural numbers, we again get natural numbers. We express it by saying that the set of natural numbers is closed with respect to these operations.

However, if you subtract 2 from 2, then what you get is not a natural number. It is a number which we call zero denoted as '0'. The word, zero, in fact is a translation of the Sanskrit 'shunya'. It is universally accepted that the concept of the number zero was given by the ancient Hindu mathematicians. You come across with certain concrete situations indicating the meaning of zero. For example, the temperature of zero degree is certainly not an absence of temperature.

After having fixed the idea of the number zero, it should not be difficult for you to understand the notion of negative natural numbers. You must have heard the weather experts saying that the temperature on the top of the hills is minus 5 degrees written as –5°. What does it mean? The simple and straight explanation is that –5 is the negative of 5 i.e. –5 is a number such that 5 + (–5) = 0. Hence –5 is a negative natural number. Thus for each natural n, there is a unique number –n, called the negative of n such that

$$n + (-n) = 0.$$

1.2.2 Integers

You have seen that in the set N of natural numbers, if we subtract 2 from 2 or 3 from 2, we do not get back natural numbers. Thus set of natural numbers is not closed with respect to the operation of subtraction. After the operation of subtraction is introduced, the need to include 0 and negative numbers becomes apparent. To make this operation valid, we must enlarge the system of natural numbers, by including in it the number 0 and all the negative natural numbers. This enlarged set consisting of all the natural numbers, zero and the negatives of natural numbers, is called the set of integers. It is denoted as

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

Now you can easily verify that the set of integers is closed with respect to the operations of addition, multiplication and subtraction.

The integers 1, 2, 3 .... are also called positive integers which are in fact natural numbers. The integers –1, –2, –3,.... are called negative integers which are actually the negative natural numbers.

Notes

The number 0, however, is neither a positive integer nor a negative one. The set consisting of all the positive integers and 0 is called the set of non-negative integers. Similarly we talk of the set of non-positive integers. Can you describe it?

1.2.3 Rational Numbers

If you add or multiply the integers 2 and 3, then the result is, of course, an integer in each case. Also if you subtract 2 from 2 or 2 from 3, the result once again in each case, is an integer. What do you get, when you divide 2 by 3? Obviously, the result is not an integer. Thus if you divide an integer by a non-zero integer, you may not get an integer always. You may get the numbers of the form

$\frac{1}{2}, \frac{1}{3}, \frac{-2}{3}, \frac{-4}{5}, \frac{5}{6} \ldots$ so on.

Such numbers are called rational numbers.

Thus the set Z of integers is inadequate when the operation of division is introduced. Therefore, we enlarge the set Z to that of all rational numbers. Accordingly, we get a bigger set which includes all integers and in which division by non-zero integers is possible. Such a set is called the set of rational numbers. Thus a rational number is a number of the form $\frac{p}{q}$, $q \neq 0$, where p and q are integers. We shall denote the set of all rational numbers by Q. Thus,

$$Q = \{x = \frac{p}{q},\ P \in Z, q \in Z, q \neq 0).$$

Now if you add or multiply any two rational numbers you again get a rational number. Also if you subtract one rational number from another or if you divide one rational number by a non-zero rational, you again get a rational numbers in each case. Thus the set Q of rational numbers looks to be a highly satisfactory system of numbers in the sense that the basic operations of addition, multiplication, subtraction and division are defined on it. However, Q is also inadequate in many ways. Let us now examine this aspect of Q.

Consider the equation $x^2 = 2$. We shall show that there is no rational number which satisfies this equation. In other words, we have to show that there is no rational number whose square is 2. We discuss this question in the form of the following example:

Example: Prove that there is no rational number whose square is 2.

Solution: If possible, suppose that there is a rational number x such that $x^2 = 2$. Since x is a rational number, therefore x must be of the form

$$x = \frac{p}{q},\ p \in Z, q \in Z, q \neq 0.$$

Note that the integers p and q may or may not have a common factor. We assume that p and q have no common factor except 1.

Squaring both sides, we get

$$\frac{p^2}{q^2} = 2.$$

Then we have

$$p^2 = 2q^2.$$

Notes

This means that p^2 is even and hence p is even (verify it). Therefore, we can write p = 2k for some integer k. Accordingly, we will have

$$p_2 = 4k^2 = 2q_2$$

or

$$q^2 = 2k^2.$$

Thus p and q are both even. In other words, p and q have 2 as a common factor. This contradicts our supposition that p and q have no common factor.

Hence there is no rational number whose square is 2.

Why don't you try the following similar exercises?

Thus you have seen that there are numbers which are not rationals. Such numbers are called irrational. In other words, a number is irrational if it cannot be expressed as p/q, $p \in Z$, $q \in Z$, $q \neq 0$. In this way, $\sqrt{2}, \sqrt{3}, \sqrt{5}$, etc. are irrational numbers. In fact, such numbers are infinite. Rather, you will see in Unit 2 that such numbers are even uncountable. Also you should not conclude that all irrational numbers can be obtained in this way. For example, the irrational numbers e and π are not of this form. We denote by I, the set of all irrational numbers.

Thus, we have seen that the set Q is inadequate in the sense that there are number which are not rationals.

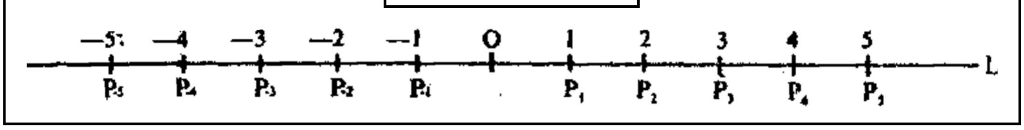
A number which is either rational or irrational is called a real number. The set of real numbers is denoted by R. Thus the set R is the disjoint union of the sets of rational and irrational numbers i.e. $R = Q \cup I$, $Q \cap I = O$.

Now in order to visualise a clear picture of the relationship between the rationals and irrationals, their geometrical representation as points on a line is of great help. We discuss this in the next section.

1.3 The Real Line

Draw a straight line L as shown in the Figure 1.1.

Figure 1.1

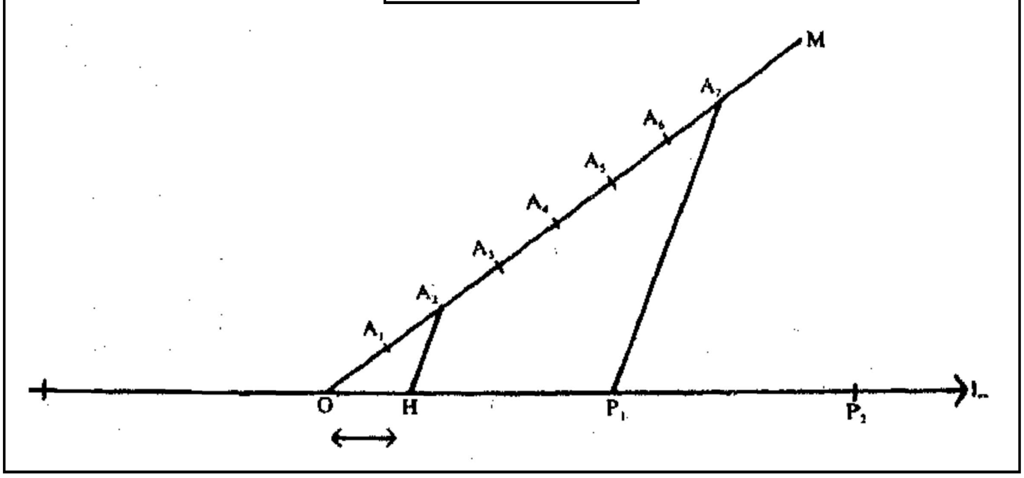


Choose a point O on L and another point P, to the right of O. Associate the number O (zero) to the point O and the number 1 to the point P_1. We take the distance between the points P and P_1 as a unit length. We mark a succession of points P_2, P_3, …… to the right of P_1 each at a unit distance from the preceding one. Then associate the integers 2, 3, ... to the points P_2, P_3, ..., respectively. Similarly, mark the points P_{-1}, P_{-2},..., to the left of the point O, Associate the integers –1, –2,… to the points P_{-1}, P_{-2},…. Thus corresponding to each integer, we have associated a unique point of the line L.

Now associate every rational number to a unique point of L. Suppose you want to associate the rational number $\frac{2}{7}$ to a point on the line L. Then $\frac{2}{7} = 2 \times \frac{1}{7}$ i.e., one unit is divided into seven parts, out of which 2 are to be taken. Let us see how you do it geometrically.

Notes

Figure 1.2



Through O, draw a line OM inclined to the line L. Mark the points A, A,..., A_7 on the line OM at equal distances. Join P_1A,. Now if you draw a line through A, parallel to P_1P_7 to meet the line L in H. Then H corresponds to the rational number $\frac{2}{7}$ i.e., OH = $\frac{2}{7}$.

You can do likewise for a negative rational number. Such points, then, will be to the left of O.

By having any line through O, you can see that the point P does not depend upon chosen line OM. Thus, you have associated every rational number to a unique point on the line L.

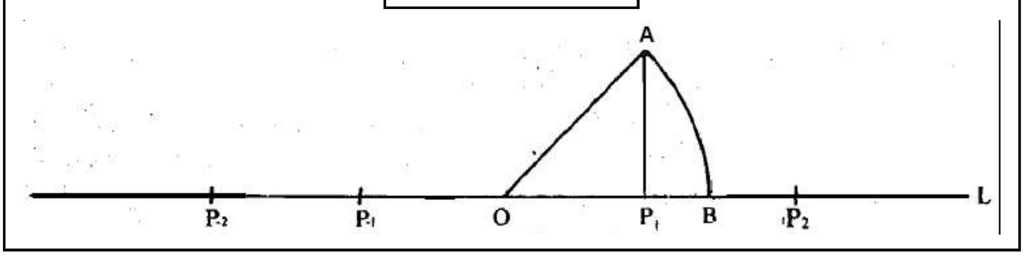
Now arises the important question:

Have you used all the points of the line L while representing rational numbers on it?

The answer to this question is NO. But how? Let us examine this.

At the point P, draw a line perpendicular to the line and mark A such that P_1A = 1 unit. Cut off OB = OA on the line, as shown in the Figure 1.3.

Figure 1.3



Then B is a point which correspond to a number whose square is 2. You have already seen that there is no rational number whose square is 2. In fact, the length OA = $\sqrt{2}$ by Pythagorean Theorem. In other words, the irrational number $\sqrt{2}$ is associated with the point B on the line L. In this way, it can be shown that every irrational number can be associated to a unique point on the line L.

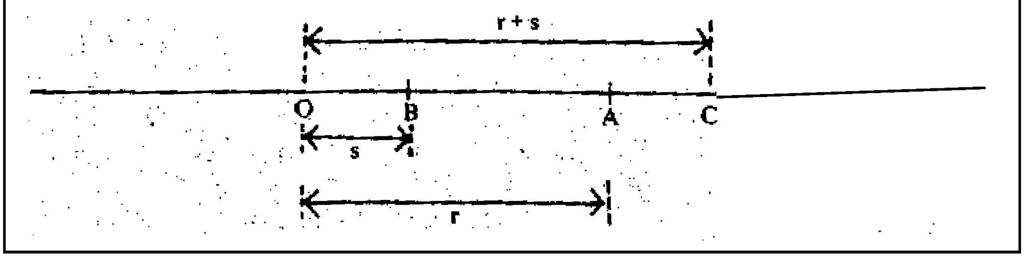
Thus, it can be shown that to every real number, there corresponds a unique point on the line L. In other words, all the real numbers are represented as points on a line. Is the converse true? That is to say, does every point on the line correspond to a unique real number? This is true but we are not going to prove it here. Therefore, hence onwards, we shall say that every real number

Notes

corresponds to a unique point on the line and conversely every point on the line corresponds to a unique real number. In this sense, the line L is called the Real Line.

Now let L be the real line.

Figure 1.4



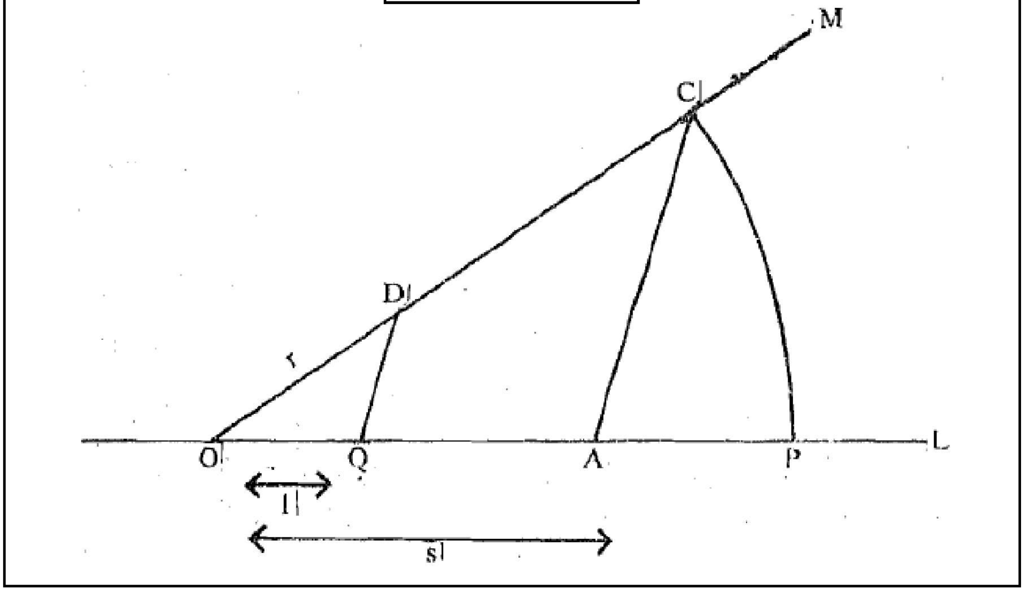
We may define addition (+) and multiplication (.) of real numbers geometrically as follows:

Suppose A represents a real number r and B represents a real number s so that OA = r and OB = s. Shift OB so that O coincides with A. The point C which is the new position of B is defined to represent r + s. See the Figure 1.4.

The construction is valid for positive as well as negative values of r and s. A real number r is said to be positive if r corresponds to a point on the line L on the right of the point O. It is written as r > 0. Similarly, r is said to be negative if it corresponds to a point on the left of the point O and is written as r < 0. Thus if r is a real number then either r is zero or r is positive or r is negative i.e. either r = 0 or r > 0 or r < 0. You should try the following exercise:

What about the product r.s of two real numbers r and s? We shall consider the case when r and s are both positive real numbers.

Figure 1.5



Though O draw some other line OM. On L, let A represent the real number s. On OM take a point D so that OD = r. Let Q be a point on I, so that OQ = 1 unit. Join QD. Through A draw a straight line parallel to QD to meet OM at C. Cut off OP on the line equal to OC. Then F represents the real number r.s.

Notes

Suppose s is a positive real number and r is a negative real number. Then, there exists a number r such that r = –r′ where r′ is a positive real numbers. Therefore, the product rs can be defined on L as

$$rs = (-r')s = -(r's).$$

Similarly you can state that rs = r(–s′) = ⁻ (rs′) where s is negative and s = –s′ for some positive s′, while r is positive.

If both r and s are negative and r = –r′ and s = –s′ where r′ and s′ are positive real numbers, then we define

$$rs = r's' = (-r)\,(-s).$$

We can also similarly define 0, r = r! 0 = 0 for each real number r.

1.4 Complex Numbers

So far, we have discussed the system of real numbers. We have yet, another system of numbers. For example, if you have to find the square root of a negative real number say –5, then you will write at as $\sqrt{-1}, \sqrt{5}$. You know that $\sqrt{5}$ is a real number but what about $\sqrt{-1}$? Again you can verify that a simple equation $x^2 + 1 = 0$ does not have a solution in the set of real numbers because the solution involves the square root of a negative real number. As a matter of fact, the problem is to investigate the nature of the number $\sqrt{-1}$ which we denote by such that $i^2 = -1$. Let us discuss the following example to know the nature of i.

Example: Show that i is not a real number.

We claim that i is not a real number. If it is so, then either i = 0 or i > 0 or I < 0.

If i = 0, then $i^2 = 0$. This implies that –1 = 0 which is absurd. If i > 0, then $i^2 > 0$ which implies that –1 > 0. This is also absurd. Finally, if i < 0, then again $i^2 > 0$ which implies that –1 > 0. This again is certainly absurd. Thus i is not a real number. This number 'i' is called the imaginary number. The symbol 'i' is called 'iota' in Greek language. This generates another class of numbers, the so called complex numbers.

The basic idea of extending the system of real numbers to the system of complex numbers arose due to the necessity of finding the solutions of the equations, like $x^2 + 1 = 0$ or $x^2 + 2 = 0$ and so on. The first contribution to the notion of such a number was made by the most celebrated Indian Mathematician of the 9th century, Mahavira, who showed that a negative real number does not have a square root in the set of real numbers. But it was an Italian mathematician, G. Cardon [1501-1576] who used imaginary numbers in his work without bothering about their existence. Due to notable contributions made by a large number of mathematicians, the system of complex numbers came into existence in the 18th century. Since we are dealing with real numbers, therefore, we shall not go into the detailed discussion of complex numbers. However, we shall give a brief introduction to the system of complex numbers. We denote the set of complex numbers as

$$C = \{z = a + i\,b, \text{ a and b real numbers}\}$$

In a complex number, z = a + i b, a is called its real part and b is called its imaginary part.

Any two complex numbers $z_1 = a_1 + i\,b_1$ and $z_2 = a_2 + i\,b_2$ are equal if only their corresponding real and imaginary parts are equal.

If $z_1 = a_1 + i\,b_1$ and $z_2 = a_2 + i\,b_2$ are any two complex numbers, then we define addition (+) and multiplication (.) as follows:

Notes

$$z_1 + z_2 = (a_1 + a_2) + i\,(b_1 + b_2)$$

$$z_1.z_2 = (a_1a_2 - b_1b_2) + i\,(a_1b_1 - a_2b_2).$$

The real numbers represent points on a line while complex numbers are identified as points on the plane.

Before concluding this section, we would like to mention yet another classification of numbers as enunciated by some mathematicians. Consider the number $\sqrt{2}$. This is an example of what is called an Algebraic Number because it satisfies the equation

$$x^2 - 2 = 0.$$

A number is called an Algebraic Number if it satisfies a polynomial equation

$$a_0x^n + a_1x^{n-1}\ t \ldots. + a_{n-1}\ x + a,, + a_n = 0$$

where the coefficients a_0, a_1, a_2,.... a, are integers, a, $\neq 0$ and $n > 1$. The rational numbers are always algebraic numbers. The numbers defined in terms of the square root etc., are also treated as algebraic numbers. But there are some real numbers which are not algebraic. Such numbers are called the Transcendental numbers. The numbers π and ε are transcendental numbers.

You may think that the operations of algebraic operations viz. addition, multiplication, etc. are the only aspects to be discussed about the set of real numbers. But certainly there are some more important aspects of the set of real numbers as points on the real line. We shall discuss these aspects in Unit 3 namely the point sets of the real line called also the topology of the real line. But prior to that, we shall discuss the structure of real numbers in Unit 2.

We conclude this unit by talking briefly about an important hypothesis-closely linked with the system of natural numbers. This is called the Principle of Induction.

1.5 Mathematical Induction

The Principle of Induction and the natural numbers are inseparable. In Mathematics, we often deal with the proofs of various theorems and formulas. Some of these are derived by the direct proofs, while some others can be proved by certain indirect methods. Consider, for example, the validity of the following two statements:

(i) The number 4 divides $5^n - 1$ for every natural number n.

(ii) The sum of the first n natural numbers is $\frac{n(n+1)}{2}$ i.e.

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}.$$

In fact, you can provide most of the verifications for both statements in the following way:

For (i), if n = 1, then $5^n - 1 = 5 - 1 = 4$ which is obviously divisible by,

if n = 2, then $5^2 - 1 = 24$, which is also divisible by 4;

if n = 6, then $5^6 - 1 = 15624$, which is indeed divisible by 4.

Similarly for (ii) if n = 10 then 1 + 2 + ... 4 – 10 = 55, while the formula

$$\frac{n(n+1)}{2} = 55 \text{ when } n = 10.$$

Again, if n = 100; then also you can verify that in each way, the sum of the first hundred natural numbers is 5050 i.e.

Notes

$$\frac{n(n+1)}{2} = 5050 \text{ for } n = 100.$$

What do these statements have in common and what do they indicate? The answer is obvious that each statement is valid for every natural number.

Thus to a great extent, a large number of theorems, formulas, results etc. whose statement involves the phrase, "for every natural number n" are those for which an indirect proof is most appropriate. In such indirect proofs, clearly a criterion giving a general approach is applied. One such criterion is known as the criterion of Mathematical Induction. The principle of Mathematical Induction is Stated (without proof) as follows:

Principle of Mathematical Induction

Suppose that, for each $n \in N$, P (n) is a statement about the natural number n. Also, suppose that

(i) P(1) is true,

(ii) if P(n) is true, then P(n + 1) is also true.

Then P(n) is true for every $n \in N$.

Let us illustrate this principle by an example:

Example: The sum of the first n natural numbers is $\frac{n(n+1)}{2}$

Solution: In other words, we have to show that for each $n \in N$,

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

$$S_n = 1 + 2 + 3 + \ldots + n$$

$$= \sum_{k=1}^{n} k.$$

Let P(n) be the statement that

$$S_n = \frac{n(n+1)}{2}$$

We, then, have $S_i = 1$ and $\frac{1(1+1)}{2} = 1$. Hence P(1) is true.

This proves part (i) of the Principle of Mathematical Induction. Now for (ii), we have to verify that if P(n) is true, then P(n + 1) is also true. For this, let us assume that P(n) is true and establish that P(n + 1) is also true. Indeed, if we assume that

$$S_n = \frac{n(n+1)}{2},$$

then we claim that

$$S_{n+1} = \frac{(n+1)(n+2)}{2}$$

Indeed

$$S_{n+1} = 1 + 2 + 3 + \ldots + n + (n + 1)$$

$$= S_n + (n + 1)$$

Notes

$$= \frac{1}{2} n(n + l) + (n + 1)$$

$$= \frac{(n+1)(n+2)}{2}$$

Thus P(n + 1) is also true.

Similarly, by using the Principle of Induction, you can prove that

(i) the sum of the squares of the first n natural numbers is $\frac{1}{6}$ n(n + l) (2n +1); and

(ii) the sum of the cubes of the first n natural numbers is $\frac{1}{4}$ n^2 $(n + 1)^2$.

Self Assessment

Choose appropriate answer for the following:

1. The complement of the set S is the set of all those element of ....which do not belong to S. It is denoted by S.

   (a) universal set (b) empty set

   (c) union set (d) intersection set

2. Let S and T all two sets. The collection of all elements which belong to S or T is called ....

   (a) universal (b) union

   (c) intersection (d) Difference of two set

3. The intersection .... of sets S and T is defined to be the set of all those elements which belong to both S and T.

   (a) $S \cup T$ (b) $S \neq T$

   (c) $S \cap T$ (d) $S \leq T$

4. If let S = {1, 2, 3} and T = {a, b, c} and let f : S $\rightarrow$ T be defined as f(1) = a, f(2) = b, f(b) = c. Then f is.....................

   (a) one-one (b) onto

   (c) one-one and onto (d) one-one and surjection

5. The set S is called the domain of the function f and T is called its ....

   (a) Range (b) pre-domain

   (c) co-domain (d) bijection

1.6 Summary

- We have recalled some of the basic concepts of sets and functions in section 1.2. A set is a well-defined collection of objects. Each object is an element or a member of the set. Sets are generally designated by capital letters and the members by small letters enclosed with braces. There are two ways to indicate the members of a set. The tabular method or listing method in which we list each element of a set individually and the set-builder method

Notes

which gives a verbal description of the elements or a property that is common to all the elements of a set.

- A set with a limited number of elements is a Finite set. A set with an unlimited number of elements is an infinite set. A set with no elements is a null-set. A set S is a subset of a set T if every element of S is in T. The set S is said to be a proper subset of T if every element of S is in T and there is at-least one element of T which does not belong to S. The sets S and T are equal if S is a subset of T and T is a subset of S. The null set is a subset of every set and every set is a subset of itself.
- The union of two sets S and T, written as $S \cup T$, includes elements of S and T without repetitions. The intersection of S and T, written as S n T, includes all those elements that are common to both S and T. The complement of a set S in a Universal set X is denoted as S^c and it consists of all those elements of X which do not belong to S. The laws with respect to union, intersection and complement have been asked in the form of exercises. Also, these notions have been extended to an arbitrary family of sets.
- A function $f: S \rightarrow T$ is a rule by which you can associate to each element of S, a unique element of T. The set S is the domain and the set T is the co-domain of f. The set $\{f(x):x \in S\}$ is the Range of f, where f(x) is an image of x under f. The function f is one-one if $f(x_1) = f(x_2) \Rightarrow x, = x$, for any x, x, in the domain of f. It is said to be onto if the range of f is equal to the domain of f. A function f is said to be a one-one correspondence if it is both one-one and onto. A function $i: S \rightarrow S$ defined by $i(x) = x, \forall x \in S$ is called an identity function, while a function $f; S \rightarrow T$ is said to be constant if $f(x) = c, x \in S$, c being a fixed element of T.
- Any two functions with the same domain are said to be equal if they have the same image for each point of the domain. The composite of the functions $f: S \rightarrow T$ and $g: T \rightarrow V$ is a function denoted as 'g o f': $S \rightarrow V$ and defined by (g o f) (x) = g(f(x)). The function $f: S \rightarrow T$ is said to be invertible if there exists a function $g: T \rightarrow S$ such that both 'g o f' and 'f o g' are identity functions. Also, a function is invertible if it is both one-one and onto. The inverse of f exists if f is invertible and it is denoted as f.
- We have discussed the development of the system of numbers starting from the set of natural numbers. These are the following:

Natural Numbers (Positive Integers):

$$N = \{1, 2, 4 ....\}$$

Integers:

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

Rational Numbers:

$$Q = \{\frac{p}{q} : p \in z, q \in Z, q \neq 0\}$$

Real Numbers:

R = Disjoint Union of Rational and Irrational Numbers

$$R = Q \cup I, Q \cap I = \phi$$

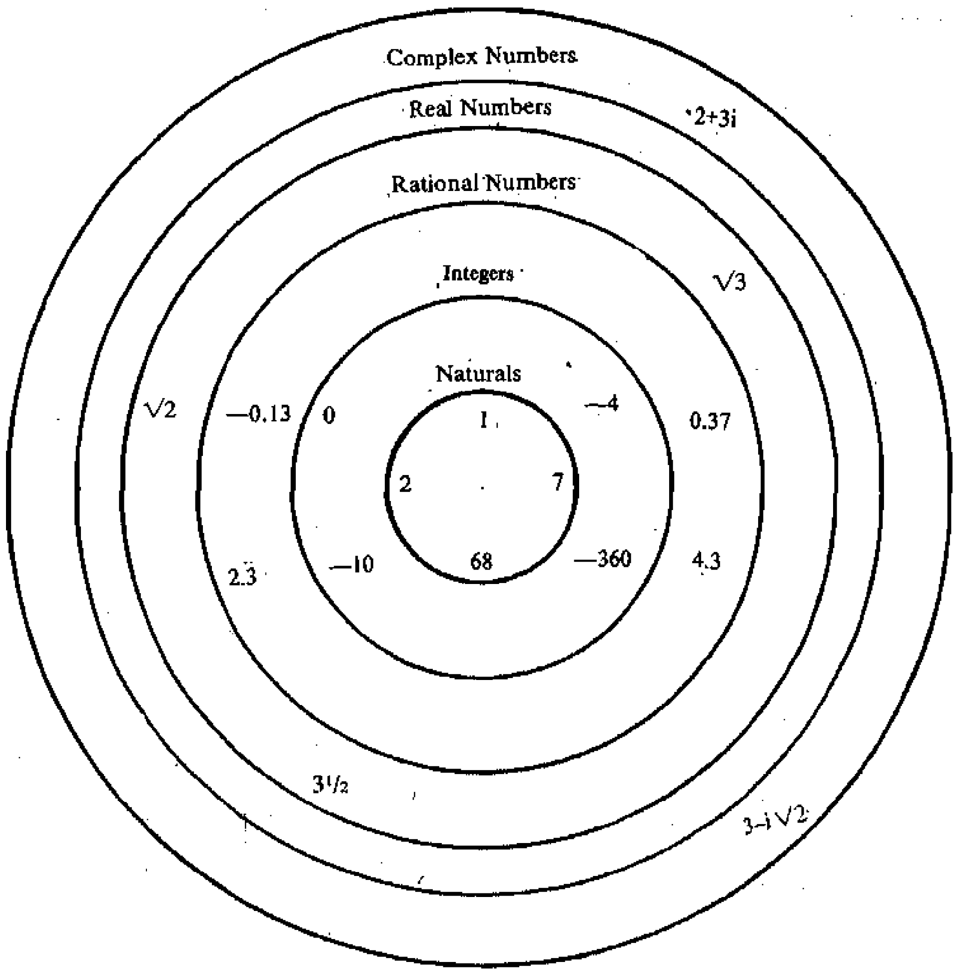
Complex Numbers:

$$C = \{z = x + iy : x \in R, y \in R\}, i = \sqrt{-1}.$$

Notes

A mathematical development of the number systems is depicted in Figure 1.6:

Figure 1.6: A Mathematical Development of Numbers Systems



- We have discussed the geometrical representation of the real numbers and stated the continuum Hypothesis. According to this, every real number can be represented by a unique point on the line and every point on the line corresponds to a unique real number. In view of this, we call this line as the Real Line.

1.7 Keywords

Constant Function: Let S and T be any two non-empty sets. A function f: S→T defined by f(x) = c, for each x in S, where c is fixed element of T, is called a constant function.

Co-domain: The set S is called the domain of the function f and T is called its co-domain.

Finite: A set with a limited number of elements is a Finite set.

Function: A function f: S→T is a rule by which you can associate to each element of S, a unique element of T.

1.8 Review Questions

1. Write the following in the set-builder form:

   $A = \{2, 4, 6, ...\}$

   $A = \{1, 3, 5, ...\}$

2. Write the following in the tabular form:

   A = {x:x is a factor of 15}

   A = {x:x is a natural number between 20 and 30}

   A = {x:x is a negative integer}

3. Let X be a universal set and let S be a subset of X. Prove that

   (i) $P(0) = \{\varnothing\}$

   (ii) $(S^c)^c = S$.

4. Let A, B and C be any three sets. Then prove the following:

   (i) $A \cup B = B \cup A, A \cap B = B \cap A$ (Commutative laws).

   (ii) $A \cup (B \cup C) = (A \cup B) \cup C, A \cap (B \cap C) = (A \cap B) \cap C$ (Associative laws).

   (iii) $A\cup (B \cap C) = (A \cup B) \cap (A \cup C)$

   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributive laws).

   (iv) $(A \cup B)^C = A^C \cap B^C, (A \cap B)^C = A^C \cup B^C$ (DeMorgan laws).

5. Justify that

   (i) N is a proper subset of Z.

   (ii) Z is a proper subset of Q.

Answers: Self Assessment

1.	(a)	2.	(b)
3.	(c)	4.	(c)
5.	(c)		

1.9 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10, Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol : Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik : Mathematical Analysis.

H.L. Royden : Real Analysis, Ch. 3, 4.

Notes

Unit 2: Algebraic Structure and Countability

CONTENTS

Objectives
Introduction
2.1 Order Relations in Real Numbers
  2.1.1 Intervals
  2.1.2 Extended Real Numbers
2.2 Algebraic Structure
  2.2.1 Ordered Field
  2.2.2 Complete Ordered Field
2.3 Countability
  2.3.1 Countable Sets
  2.3.2 Countability of Real Numbers
2.4 Summary
2.5 Keywords
2.6 Review Questions
2.7 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss the order relation – extended real number system
- Explain the field structure of the set of real numbers
- Describe the order-completeness
- Discuss countability to various infinite sets

Introduction

In Unit 1 we have discussed the construction of real numbers from the rational numbers which, in turn, were constructed from integers. In this unit, we show that the set of real numbers has an additional property which the set of rational numbers does not have, namely it is a complete ordered field. The questions, therefore, that arise are: What is a field? What is an ordered field? What does it mean for an ordered field to be complete? In order to answer these questions we need a few concepts and definitions, e.g., those of order inequalities and intervals in R. We shall discuss these concepts. Also in this unit, we shall explain the extended real number system.

You know that a given set is either finite or infinite. In fact a set is finite, if it contains just n elements where n is some natural number. A set which is not finite is called an infinite set. The elements of a finite set can be counted as one, two, three and so on, while those of an infinite set can not be counted in this way. Can you count the elements of the set of natural numbers? Try it. We shall show that this notion of counting can be extended in certain sense to even infinite sets.

2.1 Order Relations in Real Numbers

Notes

We have demonstrated that every real number can be represented as a unique point on a line and every point on a line represents-% unique real number. This helps us to introduce the notion of inequalities and intervals on the real line which we shall frequently use in our subsequent discussion through out the course.

You know that a real number x is said to be positive if it lies on the right side of O, the point which corresponds to the number 0 (zero) on the real line. We write it as $x > 0$. Similarly, a real number x is negative, if it lies on the left side of O. This is written as $x < 0$. If $x > 0$, then x is a non-negative real number. The real number x is said to be non-positive if $x \leq 0$.

Let x and y be any two real numbers. Then, we say that x is greater than y if $x - y > 0$. We express this by writing $x > y$. Similarly x is less than y if $x - y < 0$ and we write $x < y$. Also x is greater than or equal to y ($x \geq y$) if $x - y \geq 0$. Accordingly, x is less than or equal to y ($x \leq y$) if $x - y \leq 0$. Given any two real numbers x and y, exactly one of the following can hold:

either (i) $x > Y$

or (ii) $x < y$

or (iii) $x = y$.

This is called the law of trichotomy. The order relation $\leq$ has the following properties:

Property 1

For any x, y, z in R,

(i) If $x \leq y$ and $y \mid x$, then $x = y$,

(ii) If $x \leq y$ an $y \leq z$, then $x \leq z$,

(iii) If $x \leq y$ then $x + z \mid y + z$,

(iv) If $x < y$ and $o \leq z$, then $x z \leq y z$.

The relation satisfying (i) is called anti-symmetric. It is called transitive if it satisfies (ii). The property (iii), shows that the inequality remains unchanged under addition of a real number. The property (iv) implies that the inequality also remains unchanged under multiplication by a non-negative real number. However, in this case the inequality gets reversed under multiplication by a non-positive real number. Thus, if $x \mid y$ and $z \leq 0$, then $xz \geq yz$. For instance, if $z = -1$, we see that

$$-2 \leq 4 \Rightarrow 2\,(-1) \geq 4\,(-1) \Rightarrow -2 \geq -4.$$

We state the following results without proof:

- There lie an infinite number of rational numbers between any two distinct rational numbers.
- As a matter of fact, something more is true.
- Between any two real numbers, there lie infinitely many rational (irrational) numbers. Thus there lie an infinite number or real numbers between any two given real numbers.

2.1.1 Intervals

Now that the notion of an order has been introduced in R, we can talk of some special subsets of R defined in terms of the order relation. Before we formally define subset, we first introduce the notion of 'betweenness', which we have already used intuitively in the previous results. If 1, 2, 3 are three real numbers, then we say that 2 lies between 1 and 3. Thus, in general, if a, b and c are any three real numbers such that $a\ 5 \leq b \leq c$ then we say that b lies 'between' a and c. Closely related to notion of betweenness is the concept of an interval.

Notes

Definition 1: Interval

An interval in R is an non-empty subset of R which has the property that, whenever two numbers a and b belong to it, all numbers between a and b also belong to it.

The set N of natural numbers is not an interval because while 1 and 2 belong to N, but 1.5 which lies between 1 and 2, does not belong to N.

We now discuss various forms of an interval.

Let a, b $\in$ R with a $\leq$ b.

(i) Consider the set {x $\in$ R : a $\leq$ x $\leq$ b}. This set is denoted by]a, b[, and is called a closed interval. Note that the-end points a and b are included in it.

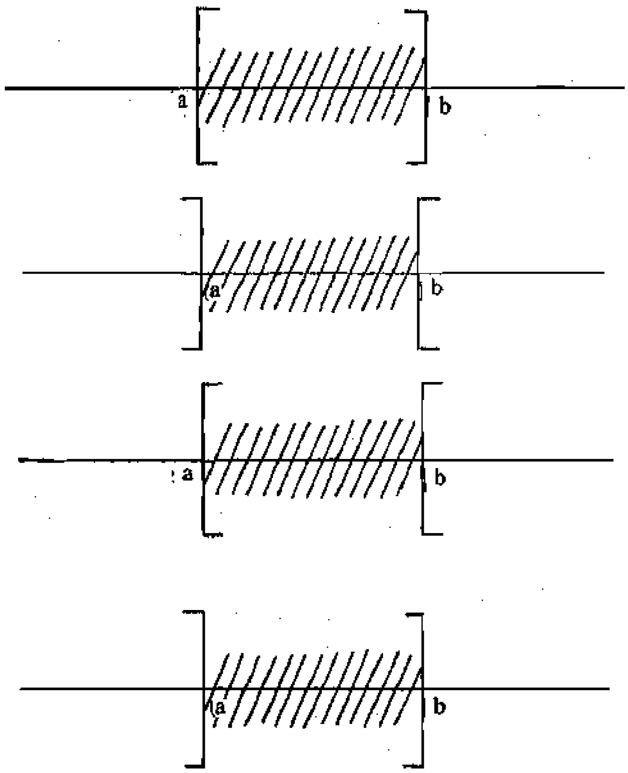
(ii) Consider the set{x $\in$ R : a < x < b}. This set is denoted by [a, b], and is called an open interval. In this case the end points a and b are not included in it,

(iii) The interval {x $\in$ R: a $\leq$ x < b} is denoted by [a, b[.

(iv) The interval {x $\in$ R : a < x $\leq$ b} is denoted by]a, b].

You can see the graph of all the four intervals in the Figure 2.1.

Figure 2.1



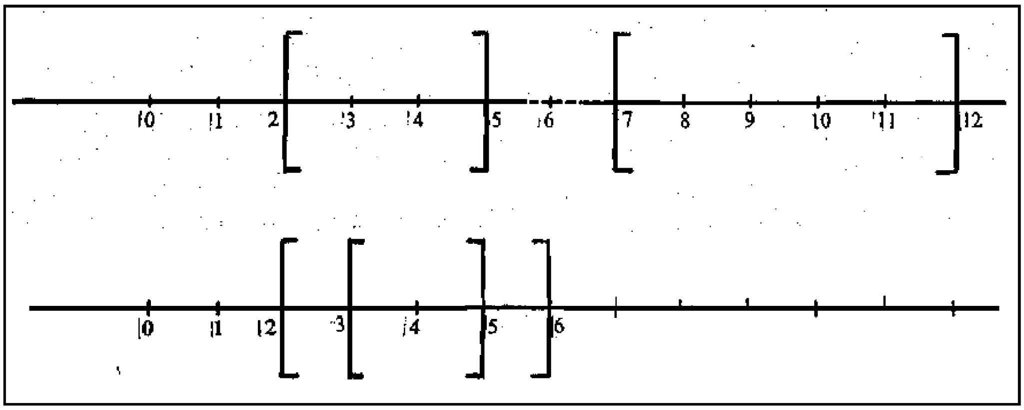
Intervals of these types are called bounded intervals. Some authors also call them finite intervals. But remember that these are not finite sets. In fact these are infinite sets except for the case [a, a] = {a}.

You can easily verify that an open interval]a, b[as well as]a, b] and [a, b[are always contained in the closed interval [a, b].

Notes

Example: Test whether or not the union of any two intervals is an interval.

Solution: Let [2, 5] and [7, 12] be two intervals. Then [2, 5] ∪ [7, 12] is not an interval as can be seen on the line in Figure below.



However, if you take the intervals which are not disjoint, then the union is an interval. For example, the union of [2, 5] and [3, 6] is [2, 6] which is an interval. Thus the union of any two intervals is an interval provided the intervals are not disjoint.

2.1.2 Extended Real Numbers

The notion of the extended real number system is important since we need it in this unit as well as in the subsequent units.

You are quite familiar with the symbols + ∞ and – ∞. You often call these symbols are 'plus infinity' and 'minus infinity', respectively. The symbols + ∞ and – ∞ are extremely useful. Note that these are not real numbers.

Let us construct a new set R* by adjoining – ∞ and + ∞ to the set R and write it as

$$R^* = R \cup \{-\infty, +\infty\}.$$

Let us extend the order structure to R* by a relation < as $-\infty < x < +\infty$, for every $x \in R$. Since the symbols – ∞ and + ∞ do not represent any real numbers, you should, therefore, not apply any result stated for real numbers, to the symbols + ∞ and – ∞. The only purpose of using these symbols is that it becomes convenient to extend the notion of (bounded) intervals to unbounded intervals which are as follows:

Let a and b be any two real numbers. Then we adopt the following notations:

$$[a, \infty] = \{X \in R: x \geq a\}$$

$$[a, do] = \{X \in R: x > a\}$$

$$[-\infty, b] = \{x \in R: x \leq b\}$$

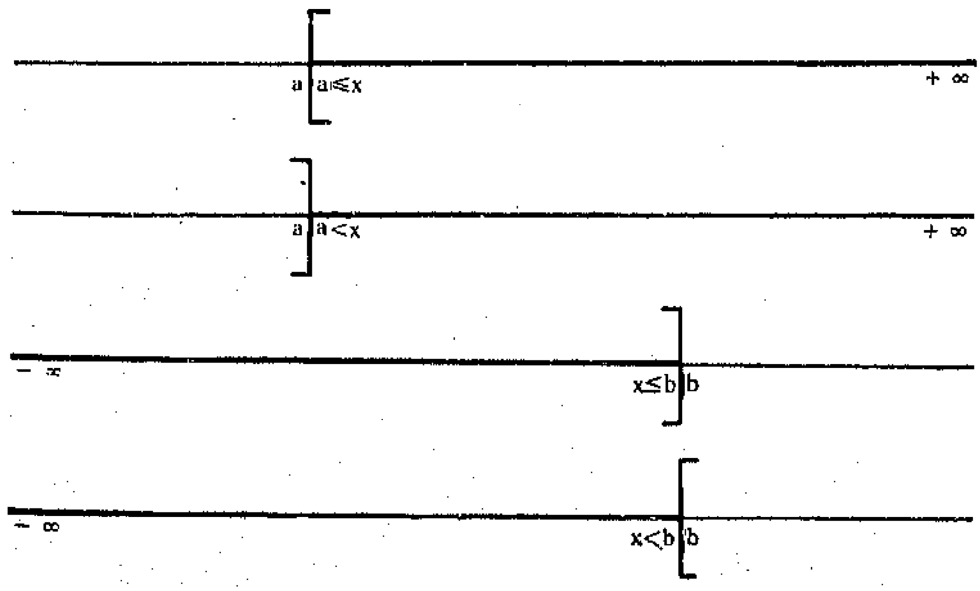
$$[-\infty, b] = \{x \in R: x < b\}$$

$$[-\infty,\infty] = \{X \in R: -\infty < x < \infty\}.$$

Notes

You can see the geometric representation of these intervals in Figure 2.2.

Figure 2.2



All these unbounded intervals are also sometimes called infinite intervals.

You can perform the operations of addition and multiplication involving – ∞ and + ∞ in the following way: For any $x \in R$, we have

$$x + (+\infty) = +\infty,$$

$$x + (+\infty) = -\infty,$$

$$x. (+\infty) = +\infty, \text{ if } x > 0$$

$$x. (+\infty) = -\infty, \text{ if } x < 0$$

$$x. (+\infty) = -\infty, \text{ if } x > 0$$

$$x. (-\infty) = +\infty, \text{ if } x < 0$$

$$\infty + \infty = +\infty, -\infty - \infty = -\infty$$

$$\infty. (-\infty) = -\infty, (-\infty). (-\infty) = +\infty.$$

Note that the operations $\infty - \infty$, $0. \infty$, $\frac{\infty}{\infty}$ are not defined.

2.2 Algebraic Structure

During the 19th Century, a new trend emerged in mathematics to use algebraic structures in order to provide a solid foundation for Calculus and Analysis. In this quest, several methods were used to characterise the red numbers. One of the methods was related to the least upper bound principle used by Richard Dedekind which we discuss in this section.

This leads us to the description of the real numbers as a complete ordered field. In order to define a complete ordered field. We need some definitions and concepts.

You are quite familiar with the operations of addition and multiplication on numbers, union and intersection on the subsets of a universal set. For example, if you add or multiply any two natural numbers, the sum or the product is a natural number. These operations of addition or multiplications on the sets of numbers are examples of a binary operation on a set. In general, we can define a binary operation on a set in the following way:

Definition 2: Binary Operation

Given a non-empty set S, a binary operation on S is a rule which associates with each pair of elements of S, a unique element of S.

We denote this rule by symbols such as ., *, +, etc.

By an Algebraic Structure, we mean a non-empty set together with one or more binary operations defined on it. A field is an algebraic structure which we define, as follows:

Definition 3: Field Structure

A field consists of a non-empty set F together with two binary operations defined on it, denoted by the symbols '+'(addition) and '.' (multiplication) and satisfying the following axioms for any elements x, y, z of the set F.

A_1: $x + y \in F$ (Additive Closure)

A_2: $x + (y + z) = (x + y) + z$ (Addition is Associative)

A_3: $x + y = y + x$ (Addition is Commutative)

A_4: There exists an element in F, denoted by '0' and called the zero or the zero element of F such that $x + 0 = 0 + x = x \ \forall \ x \in F$ (Additive Identity)

A_5: For each $x \in F$, there exists an element $-x \in F$ with the property (Additive Inverse)

$x + (-x) = (-x) + x = 0$

The element $-x$ is called additive inverse of x.

M_1: $x.y \in F$ (Multiplicative Closure)

M_2: $(x.y).z = x.(y.z)$ (Multiplication is Associative)

M_3: $x.y = y.x$ (Multiplication is Commutative)

M_4: There exists an element 1 different from 0 called the unity of F, such that (Multiplicative Identity)

$1.x = x.1 = x \ \forall \ x \in F$

M_5: For each $x \in F$, $x \neq 0$, there exists an element $x^{-1} \in F$ such that (Multiplicative Inverse)

$x.x^{-1} = .x^{-1} x = 1.$

The element x^{-1} is called the multiplicative inverse of x.

D: $x.(y + z) = x.y + x.z$ (Multiplication is distributive over Addition).

$(x + y) z = x.z + y.z.$

Since the unity is not equal to the zero i.e. $1 \neq 0$ in a field, therefore any field must contain at least two elements. Note that the axioms A_1 (closure under addition) and M_1 (closure under multiplication) are unnecessary because the closures are implied in the definition of a binary operation. However, we include them, for the sake of emphasis.

Notes

Now, you can easily verify that all the eleven axioms are satisfied by the set of rational numbers with respect to the ordinary addition and multiplication. Thus, the set Q forms a field under the operations of addition and multiplication, and so does, the set R of all the real numbers.

We state (without proof) some important properties satisfied by a field. They follow from the field axioms. Can you try?

Property 2

For any x, y, z in F,

1. $x + z = y + z \Rightarrow x = y$,
2. $x, 0 = 0 = 0.x$,
3. $(-x). y = - x,y = x. (-y)$,
4. $(-x), (- y) = x.y$,
5. $x.z = y.z, z \neq 0 \Rightarrow x = y$,
6. $x.y = 0 \Rightarrow$ either $x = 0$ or $y = 0$.

Thus by now you know that the sets Q, R and C form fields under the operations of addition and multiplication.

2.2.1 Ordered Field

We defined the order relation $\leq$ in R. It is easy to see that this order relation satisfies the following properties:

Property 3

Let x, y, z be any elements of R. Then

O_1: For any two elements x and y of R, one and only of the following holds:

(i) $x < y$, (ii) $y < x$, (iii) $x = y$,

O_2: $x \leq y, y \leq x \Rightarrow x \leq z$,

O_3: $x \leq y \Rightarrow x \tau z \leq y + z$,

O_4: $x \leq y, 0 < z \Rightarrow x.z \leq y.z$

We express this observation by saying that the field R is an ordered field (i.e. it satisfies the properties $0_1 - 0_4$). It is easy to see that these properties are also satisfied by the field Q of rational numbers. Therefore, Q is also an ordered field. What about the field C of Complex numbers?

2.2.2 Complete Ordered Field

Although R and Q are both ordered fields, yet there is a property associated with the order relation which is satisfied by R but not by Q. This property is known as the Order-Completeness, introduced for the first time by Richard Dedekind. To explain this situation more precisely, we need a few more mathematical concepts which are discussed as follows:

Consider set S = {1, 3, 5, 7). You can see that each element of S is less than or equal to 7. That is $x \leq 7$, for each $x \in S$. Take another set S, where S = {$x \in R : x \leq 17$). Once again, you see that each element of S is less than 18. That is, $x < 18$, for each $x \in S$. In both the examples, the sets have a special property namely that every element of the set is less than or equal to some number.

Notes

This number is called an upper bound of the corresponding set and such a set is said to be bounded above. Thus, we have the following definition:

Definition 4: Upper Bound of a Set

Let $S \subset R$. If there is a number $u \in R$ such that $x \le u$, for every $x \in S$, then S is said to be bounded above. The number u is called an upper bound of S.

Example: Verify whether the following sets are bounded above. Find an upper bound of the set, if it exists.

(i) The set of negative integers

$\{-1, -2, -3, ...\}$.

(ii) The set N of natural numbers.

(iii) The sets Z, Q and R.

Solution:

(i) The set is bounded above with –1 as an upper bound,

(ii) The set N is not bounded above.

(iii) All these sets are not bounded above.

Now consider a set S = {2, 3, 4, 5, 6, 7). You can easily see that this set is bounded above because 7 is an upper bound of S. Again this set is also bounded below because 2 is a lower bound of S. Thus S is both bounded above as well as bounded below. Such a set is called a bounded set. Consider the following sets:

$$S_1 = \{... -3, -2, -1, 0, 1, 2, ...\},$$

$$S_2 = \{0, 1, 2, ...\},$$

$$S_3 = (0, -1, -2, ...\}.$$

You can easily see that S, is neither bounded above nor bounded below. The set S_4 is not bounded above while S, is not bounded below. Such sets are known as Unbounded Sets.

Thus, we can have the following definition.

Definition 5: Bounded Sets

A set S is bounded if it is both bounded above and bounded below.

In other words, S has an upper bound as well as a lower bound. Thus, if S is bounded, then there exist numbers u (an upper bound) and v (a lower bound) such that $v \le x \le u$, for every $x \in S$.

If a set S is not bounded then S is called an unbounded set. Thus S is unbounded if either it is not bounded above or it is not bounded below.

Example:

(i) Any finite set is bounded.

(ii) The set Q of rational numbers is unbounded.

(iii) The set R of real numbers is unbounded.

(iv) The set P = {sin x, sin 2x, sin 3x,......, sin nx, ...} is bounded because $-1 \le \sin nx \le 1$, for every n and x.

You can easily verify that a subset of a bounded set is always bounded since the bounds of the given set will become the bounds of the subset.

Now consider any two bounded sets say S = {1, 2, 5, 7} and T = {2, 3, 4, 6, 7, 8}. Their union and intersection are given by

$$S \cup T = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

and

$$S \cap T = \{2, 7\}.$$

Obviously $S \cup T$ and $S \cap T$ are both bounded sets. You can prove this assertion in general for any two bounded sets.

Task Prove that the union and the intersection of any two bounded sets are bounded.

Now consider the set of negative integers namely

$$S = \{-1, -3, -2, -4, ...\}.$$

You know that –1 is an upper bound of S. Is it the only upper bound of S? Can you think of some other upper bound of S? Yes, certainly, you can. What about 0? The number 0 is also an upper bound of S. Rather, any real number greater than –1 is an upper bound of S. You can find infinitely many upper bounds of S. However, you can not find an upper bound less than –1. Thus –1 is the least upper bound of S.

It is quite obvious that if a set S is bounded above, then it has an infinite number of upper bounds. Choose the least of these upper bounds. This is called the least upper bound of the set S and is known as the Supremum of the set S. (The word 'Supremum' is a Latin word). We formulate the definition of the Supremum of a set in the following way:

Definition 6: The Supremum of a Set

Let S be a set bounded above. The least of all the upper bounds of S is called the least upper bound or the Supremum of S. Thus, if a set S is bounded above, then a real number m is the supremum of S if the following two conditions are satisfied:

(i) m is an upper bound of S,

(ii) if k is another upper bound of S, then $m\ 5 \leq k$.

Task Give an example of an infinite set which is bounded below. Show that it has an infinite number of lower bounds and hence develop the concept of the greatest lower bound of the set.

The greatest lower bound, in Latin terminology, is called the Infimum of a set.

Let us now discuss a few examples of sets having the supremum and the infimum:

Example: Each of the intervals]a, b[, [a, b]]a, b], [a, b[has both the supremum and the infimum. The number a is the infimum and b is the supremum in each case. In case of [a, b] the supremum and the infimum both belong to the set whereas this is not the case for the set]a, b[. In case of the set]a, b], the infimum does not belong to it and the supremum belongs to it. Similarly, the infimum belongs to [a, b[but the supremum does not belong to it.

Very often in our discussion, we have used the expressions 'the supremum', rather than asupremum. What does it mean? It simply means that the supremum of a set, if it exists, is unique i.e. a set can not have more than one supremum. Let us prove it in the form of the following theorem:

Theorem 1: Prove that the supremum of a set, if it exists, is unique.

Proof: If possible, let there be two supremums (Suprema) say m and m′ of a set $.

Since m is the least upper bound of S, therefore by definition, we have

$$m \leq m'$$

Similarly, since m′ the least upper bound of S, therefore, we must have

$$m' \leq m.$$

This shows that m = m′ which proves the theorem.

You can now similarly prove the following result:

> *Task* Prove that the infimum of a set, if it exists, is unique.

In example 3, you have seen that supremum or the infimum of a set may or may not belong to the set. If the supremum of a set belongs to the set, then it is called the greatest member of the set. Similarly, if the infimum of a set belongs to it, then it is called the least member of the set.

Example:

(i) Every finite set has the greatest as well as the least member.

(ii) The set N has the least member but not the greatest. Determine that number.

(iii) The set of negative integers has the greatest member but not the least member. What is that number?

You have seen that whenever a set S is bounded above, then S has the supremum. In fact this is true in general. Thus, we have the following property of R without proof:

Property 4: Completeness Property

Every non-empty subset S of R which is bounded above, has the supremum.

Similarly, we have

Every non-empty subset S of R that is bounded below, has the infimum.

In fact, it can be easily shown that the above two statements are equivalent.

Now, if you consider a non-empty subset S of Q, then S considered as a subset of R must have, by property, a supremum. However, this supremum may not be in Q. This fact is expressed by saying that Q considered as a field in its down right is not Order-Complete. We illustrate this observation as follows:

Construct a subset S of Q consisting of all those positive rational numbers whose squares are less than 2 i.e.

$$S = \{x \in Q: x > 0, x^2 < 2\}.$$

Since the number 1 ES, therefore S is non-empty. Also, 2 is an upper bound of S because every element of S is less than 2. Thus the set S is non-empty and bounded, above. According to the

Notes

Axiom of Completeness of R, the subset S must have the supremum in R. We claim that this supremum does not belong to Q.

Suppose m is the supremum of the set S. If possible, let m belong to Q. Obviously, then $m > 0$. Now either $m^2 < 2$ or $m^2 = 2$ or $m^2 > 2$.

Case (i) When $m^2 < 2$. Then a number y defined as

$$y = \frac{4+3m}{3+2m}$$

is a positive rational number and

$$y - m = \frac{2(2-m^2)}{3+2m}$$

Since $m^2 < 2$, therefore $2 - m^2 > 0$, Hence

$$y - m = \frac{3(2-m^2)}{3+2m} > 0$$

which implies that $y > m$.

Again,

$$y^2 - 2 = \left(\frac{4+3m}{3+2m}\right)^2 - 2$$

$$= \frac{m^2-2}{(3+2m)^2}$$

Since $m^2 < 2$, therefore

$$y^2 - 2 < 0 \text{ i.e. } y^2 < 2.$$

This shows that $y \in S$ and also it is greater than m (the supremum of S). This is absurd. Thus the case $m^2 < 2$ is not possible.

Case (ii) When $m^2 = 2$.

This means there exists a rational number whose square is equal to 2 which is again not possible.

Case (iii) When $m^2 > 2$

In this case consider the positive rational number y defined in case (i). Accordingly, we have

$$y - m = \frac{2(2-m^2)}{3+2m} < 0 \text{ (check yourself)}$$

i.e. $y < m$.

Also $$2 - y^2 = 2 - \left(\frac{4+3m)}{3+2m}\right) 2 = \frac{2-m^2}{(3+2m)^2}$$

i.e. $$2 - y^2 < 0 \text{ or } y^2 > 2,$$

which shows that y is an upper bound of S.

Thus y is an upper bound of S which does not belong to S. At the same time y is less than the supremum of S. This is again absurd. Thus $m^2 > 2$ is also not possible. Hence none of three possibilities is true. This means there is something wrong with our supposition. In other words, our supposition is false and therefore the set, S does not possess the suprernum in Q.

This justifies that the field Q of rational numbers is not order-complete.

Now you can also try a similar exercise.

2.3 Countability

As we recalled the notion of a set and certain related concepts. Subsequently, we discussed certain properties of the sets of numbers N, Z, Q, R and C. A few more important properties and related aspects concerning these sets are yet to be examined. One such significant aspect is the countability of these sets. The concept of countability of sets was introduced by George Cantor which forms a corner stone of Modern Mathematics.

2.3.1 Countable Sets

You can easily count the elements of a finite set. For example, you very frequently use the term 'one hundred rupees' or 'fifty boxes', 'two dozen eggs', etc. These figures pertain to the number of elements of a set. Denote the number of elements in a finite set S by n (S). If S = {a, b, c, d}, then n (S) = 4. Similarly n (S) = 26, if S is the set of the letters of English alphabet. Obviously, then n (ϕ) = 0, where ϕ is the null set.

You can make another interesting observation when you count the number of elements of a finite set. While you are counting these elements, you are indirectly and perhaps unconsciously, using a very important concept of the one-one correspondence between two sets. Recall the concept of one-one correspondence. Here one of the sets is a finite subset of the set of natural numbers and the other set is the set consisting of the articles/objects like rupees, boxes, eggs, etc. Suppose you have a basket of oranges. While counting the oranges, you are associating a natural number to each of the oranges. This, as you know, is a one-one correspondence between the set of oranges and a subset of natural members. Similarly, when you count the fingers of your hands, you are in fact showing a one-one correspondence between the set of the fingers with a subset, say N_{10} = (1, 2, ... 10) of N.

Although, we have an intuitive idea of finite and infinite sets, yet we give a mathematical definition of these sets in the following way:

Definition 7: Finite and Infinite Sets

A set S is said to be finite if it is empty or if there is a positive integer k such that there is one-one correspondence between the elements of the set S and the set N_K = {1, 2, 3 ..... k}. A set is said to be infinite if it is not finite.

The advantage of using the concept of one-one correspondence is that it helps in studying the countability of infinite sets. Let E = {2, 4, 6, ...} be the set of even natural numbers. If we define a mapping f: N $\to$ E as

$$f(n) = 2n \ \forall \ n \in N,$$

then we find that f is a one-one correspondence between N and E.

Consider another examples, Suppose S = {1, 2, ... n} and T = {x_1, X_2, ... x_n}. Define a mapping f: S $\to$ T as

$$f(n) = x_n \ \ n \in S.$$

Notes

Then again f is a one-one correspondence between S and T.

Such sets are known as equivalent sets. We define the equivalent sets in the following way:

Definition 8: Equivalent Sets

Any two sets are equivalent if there is one-one correspondence between them.

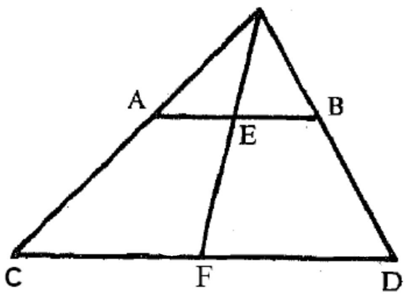
Thus if two sets S and T are equivalent, we write, as S – T.

You can easily show; that S, T and P are any three sets such that S ~ T and T – P, then S – P.

The notion of the equivalent sets is very important because it forms the basis of the 'counting' of the infinite sets.

Now, consider any two line segments AB and CD.

Figure 2.3



Let M denote the set of points on AB and N the set of points on CD. Let us check whether M and N are equivalent.

Join CA and DB to meet in the point P. Let a line through P meet AB in E and CD in F. Define f: M $\rightarrow$ N as f(x) = y where x is any point on AB and y is any point on CD. The construction shows that f is a one-one correspondence. Thus M and N are equivalent sets.

The following are some examples of equivalent sets: Let I be an interval with end points a and b, and J be an interval with end points c and d. Also, we assume that I and J are intervals of the same type. Define f : I $\rightarrow$ J, by

$$f(t) = \ d + \ c, \text{ for } t \in I.$$

Then, it is not difficult to see that f is a one-to-one correspondence between intervals I and J. Hence, all the intervals of same type are equivalent to each other.

Now, we introduce the following definition:

Definition 9: Denumerable and Countable Sets

A set which is equivalent to the set of natural numbers is called a denumerable set. Any set which is either finite or denumerable, is called a Countable set.

Any set which is not countable is said to be an uncountable set.

Example:

(i) A mapping f: Z $\rightarrow$ N defined by

$$f(n) = \begin{cases} -2n, \text{ if n is a negative integer} \\ 2n + 1, \text{ if n is non-negative integer} \end{cases}$$

is a one-to-one correspondence. Hence Z – N. Thus the set of integers is a denumerable set and hence a countable set.

(ii) Let E denote the set of all even natural numbers. Then the mapping f: N → E defined as f(n) = 2n is a one-one correspondence. Hence the set E of even natural numbers is a denumerable set and hence a countable set.

(iii) Let D denote the set of all odd integers and E the set of even integers. Then the, mapping f: E → D, defined as f(n) = n + 1 is a one-one correspondence. Thus E ~ D, But, E – N, therefore D – N. Hence D is a denumerable set and hence a countable set.

We observe that a set S is denumerable if and only if it is of the form {a,, a,, a, ...} for distinct elements a,, a,, a, ... For, in this case the mapping $f(a_n) = n$ is one-one mapping of S onto N i.e. the sets {a,, a,, a_3, ...} and the set N are equivalent.

If we consider the set S_2 = {2, 3, 4, ...}, we find that the mapping f: N → S_2 defined as f(n) = n + 1 is one-one and onto. Thus S_2 is denumerable. Similarly if we consider S_3 = {3, 4,....} or S_k = {k, k + 1, ...}, then we find that all these are denumerable sets and hence are countable sets.

We have seen that the set of integers is countable.

Now we discuss the countability of the rational and real numbers. Here is an interesting theorem.

Theorem 2: Every infinite subset of a denumerable set is denumerable.

Proof: Let S be a denumerable set. Then S can be written as

$$S = \{a,, a,, a_3, ...\}.$$

Let A be an infinite subset of S. We want to show that A is also denumerable.

You can see that the elements of S are designated by subscripts 1, 2, 3, ... Let n, be the smallest subscript for which $a_{n1} \in A$. Then consider the set A – $\{a_{n1}\}$. Again, in this new set, let n_2 be the smallest subscript such that $a_{n2} \in A - \{a,,,\}$.

Let n_k be the smallest subscript such that

$$a_{nk} \in A - \{a_{n1}, a_{n2}, ..., a_{nk-1}\}.$$

Note that such an element a_{nk} always exists for each k ∈ N as A is infinite. For, then

$$A = \left\{a_{n_1}, a_{n_2}, ..., a_{n_k}\right\} \neq \varnothing$$

for each k ∈ N. Thus, we can write

$$A = \left\{a_{n_1}, a_{n_2}, a_{n_3}, ..., a_{n_k}, ...\right\}.$$

Define f: N → A by $f(k) = a_{nk}$. Then it can be verified that f is a one-one correspondence. Hence A is denumerable. This completes the proof of the theorem.

Now consider the sets S = {6, 8, 10, 12, ...} and T = {3, 5, 7, 9, 11, ...}, which are both denumerable. Their union S ∪ T = {3, 5, 6, 7, 8, 9, ...} is an infinite subset of N and hence its denumerable. Again, if S = (–1, 0, 1, 2} and T = {20, 40, 60, 80, ...}, then we see that S ∪ T = (–1, 0, 1, 2, 20, 40, 60, ...} is a denumerable set. Note that in each case S ∩ T = 0. In fact, you can prove a general result in the following exercise.

Thus, it follows that the union of any two countable sets is countable.

Indeed, let S and T be any two countable sets. Then S and T are either finite or denumerable.

If S and T are both finite, then S ∪ T is also a finite set and hence S ∪ T is countable.

Notes

If S is denumerable and T is finite, then also we know that $S \cup T$ is denumerable. Hence $S \cup T$ is countable. Again, if S is finite and T is denumerable, then again $S \cup T$ is denumerable and countable.

Finally, if both S and T are denumerable, then $S \cup T$ is also denumerable and hence countable. In fact, this result can be extended to countably many countable sets. We prove this in the following theorem:

Theorem 3: The union of a countable number of countable sets is countable.

Proof: Let the given sets be A,, A_n, A,,.... Denote the elements of these sets, using double subscripts, as follows:

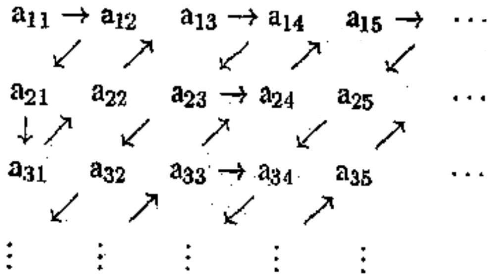
$$A_1 = \{a_{11}, a_{12}, a_{13}, ...\}$$

$$A_2 = \{a_{21}, a_{22}, a_{23}, ...\}$$

$$A_3 = \{a_{31}, a_{32}, a_{33}, ...\},$$

and so on. Note that the double subscripts have been used for the sake of convenience only. Thus a_{ij} is the jth element in the set A. Now, let us try to form a single list of all elements of the union of these given sets.

One method of doing this is by using Cantor's diagonalised counting as indicated by arrows in the following table:



Diagonalised Counting of $\bigcup_{i=1}^{\infty} A_j$.

Enlist the elements as indicated through the arrows. This is a scheme for making a single list of all the elements.

Following the arrows in above table, you can easily arrive at the new single list:

$$a,,, a,,, a,,, a_{31}\ a_{22} > a_{13}, a,,, a,,, ...$$

Note that while doing so, you must omit the duplicates, if any.

Now, if any of the sets A,, A,, ..., are finite, then this will merely shorten the final list. Thus, we have

$$\cup_i A_i = \cup_i \{a,,, a_{i2}, ...\}, i = 1, 2, 3, ...$$

which each element appears only once. This set is countable and, so, complete the proof of the theorem.

We are now in a position to discuss the countability of the sets of rational and real numbers.

2.3.2 Countability of Real Numbers

We have already established that the sets N and Z are countable. Let us, now, consider the case of the set Q of rational numbers. For this we need the following theorems:

Theorem 4: The set of all rational numbers between [0,1] is countable.

Proof: Make a systematic scheme in an order for listing the rational numbers x where * ≤ x ≤ 1, (without duplicates) of the following sets

$$A_1 = \{0, 1\}$$

$$A_2 = \left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\right\}$$

$$A_3 = \left\{\frac{2}{3}, \frac{2}{5}, \frac{2}{7}, \ldots\right\}$$

$$A_4 = \left\{\frac{3}{4}, \frac{3}{5}, \frac{3}{7}, \frac{3}{8}, \ldots\right\}$$

....

....

You can see that each of the above sets is countable. Their union is given by

$$A_i = \left\{0, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \ldots\right\} = [0, 1] \cap Q,$$

which is countable by Theorem 3.

Theorem 5: The set of all positive rational numbers is countable.

Proof: Let Q, denote the set of all positive rational numbers. To prove that Q, is countable, consider the following sets:

$$A_1 = \{1, 2, 3, \ldots\ldots..\}$$

$$A_2 = \left\{\frac{1}{2}, \frac{2}{2}, \frac{5}{2}, \ldots..\right\}$$

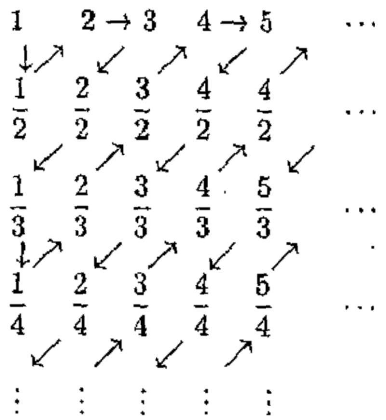
$$A_3 = \left\{\frac{1}{3}, \frac{2}{3}, \frac{4}{3}, \ldots..\right\}$$

$$A_4 = \left\{\frac{1}{4}, \frac{3}{4}, \frac{5}{4}, \ldots..\right\}$$

....

....

Enlist the elements of these sets in a manner as you have done in Theorem 3 or as known below:



Notes

You may follow the method of indicating by arrows for making a single list or you may follow another path as indicated here. Accordingly, write down the elements of Q_+ as they appear in the figure by the arrows, while omitting those numbers which are already listed to avoid the duplicates. We will have the following list:

$$Q_+ = \left\{1, \frac{1}{2}, 2, 3, \frac{1}{3}, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, 4, ...\right\}$$

$$= \bigcup_i A_i \ (i = 1, 2, 3, ...),$$

which is countable by Theorem 3. Thus Q_+ is countable.

Now let Q_- denote the set of all negative rational numbers. But Q_+ and Q_- are equivalent; sets because there is one-one correspondence between Q_+ and Q_-, f: $Q_+ \to Q_-$, given by

$$f(x) = -x, \ \forall \ x \in Q_+.$$

Therefore Q_- is also countable. Further {0} being a finite set is countable. Hence,

$$Q = Q_+ \bigcup_i \{0\} \cup Q_-$$

is a countable set. Thus, in fact, we have proved the following theorem:

Theorem 6: The set Q of all rational numbers is countable.

Proof: You may start thinking that perhaps every finite set is denumerable. This is not true. We have not yet discussed the countability of the set of real numbers or of the set of irrational numbers. To do so, we first discuss the countability of the set of real numbers in an interval with end points 0 and 1, which may be closed or open or semi-closed.

Consider the real numbers in the interval]0, 1[.

Each real number in]0, l[can be expressed in the decimal expansion. This expansion may be non-terminating or may be terminating, e.g.

$$\frac{1}{3} = .333, ...$$

is an example of non-terminating decimal expansion, whereas

$$\frac{1}{4} = .25, \ \frac{1}{2} = .5, ...,$$

are terminating decimal expansions. Even the terminating expansion can also be expressed as non-terminating expansion in the sense that you can write

$$\frac{1}{4} = .25 = .24999 .....$$

Thus, we agree to say that each real number (rational of irrational) in the]0, 1[can be expressed as a non-terminating decimal expansion in terms of the digits from 0 to 9.

Suppose x ∈]0, 1[. Then it can be written as

$$x = .c_1c_2c_3 .....$$

where c_1, c_2,.... take their values from the set {0, l, 2, 3, 4, 5, 6, 7, 8, 9) of ten digits.

Similarly, let y be another, real number in (0, l). Then y can also be expressed as

$$y = .d_1d_2d_3 .....$$

We say that x = y if the digits in their corresponding position in the expansions of x and y are identical. Thus, if there is even a single decimal places, say, 10th place such that $d_{10} \neq c_{10}$, then

$$x \neq y.$$

We now discuss the following result due to George Cantor.

Theorem 7: The set of real numbers in the interval]0, 1[is not countable.

Proof: Since the set of numbers in]0, 1[is an infinite set, therefore, it is enough to how that the set of real numbers in]0, 1[is not denumerable.

If possible, suppose the set of real numbers in]0, 1[is denumerable. Then there is a one-one correspondence between N and the elements of]0, 1[i.e. there is a function f: N →]0, 1[which is one-one and onto. Thus, if

$f(1) = x_1, f(2) = x_2, ..., f(k) = x_k, ..., \text{ then}$

$]0, 1[= \{x_1, x_2, ..., x_k, ...\}.$

We shall show that there is at least one real number]0, 1[which is not an image of any element of N under f. In other words, there is an element of]0, 1[which is not in the list $x_1, x_2, ...$

Let $x_1, x_2, ...$ be written as

$$x_1 = 0, a_{11}\ a_{12}\ a_{13}\ a_{14}\ ....$$

$$x_2 = 0, a_{21}\ a_{22}\ a_{23}\ a_{24}\ ....$$

$$x_3 = 0, a_{31}\ a_{32}\ a_{33}\ a_{34}\ ....$$

$$x_4 = 0, a_{41}\ a_{42}\ a_{43}\ a_{44}\ ....$$

....

....

....

From this we construct a real number

$$z = b_1\ b_2\ b_3\ b_4\ .....,$$

where $b_1, b_2, ...$ can take any digits from {0, 1, 2,........., 9} in such a way that $b_1 \neq a_{11}, a_2 \neq a_{22}, b_3 = a_{33}, ...$ Thus,

$$z = b_1\ b_2\ b_3\ ....$$

As a real number in]0, 1[such that $z \neq x_1$ because $b_1 \neq a_{11}$, $z \neq x_2$ because $b_2 \neq a_{22}$. In general $z \neq x_n$ because $a_{nn} \neq b_n$. Therefore z is not in the list $\{x_1, x_2\ x_3,....\}$.

Hence]0, 1[is not countable.

We have already mentioned that the intervals [0, 1], [0, 1[,]0, 1] and]0, 1[are equivalent sets. Since the set of real numbers in]0, 1[is not countable, therefore none of the intervals is a countable set of real numbers.

Now you can easily conclude that the set of irrational numbers in]0, 1[is not countable. If possible, suppose that the set of irrational numbers in]0, 1[is countable. Also you know that the set of rational numbers in]0, 1[is countable and that the union of two countable sets is countable. Therefore, the union of the set of rational numbers and the set of irrational numbers]0, 1[is countable i.e. the set of all real numbers in]0, 1[is countable which by above theorem is not so. Hence the set of irrational numbers in]0, 1[is not countable.

In fact, every interval]a, b[or [a, b],]a, b], [a, b[is an uncountable set of real numbers.

Notes

Now what about the countability of the set R of real numbers?

Suppose that R is countable. Then an interval]0, 1[, being an infinite subset of R, must be countable. But then, we have already proved that the set]0, 1[is not countable. Hence R cannot be countable.

Thus even by the method of countability of sets, we have established the much desired distinction between Q and R in the sense that Q is countable but R is not countable.

Also, we observe that although R is not countable, yet it contains subsets which are countable. For example R has subsets as Q, Z and N which are countable. At the same time R is an infinite set. We sum up this observation in the form of the following theorem:

Theorem 8: Every infinite set contains a denumerable set.

Proof: Let S be an infinite set. Consider some element of S. Denote it by n_1. Consider the set S – $\{a_1\}$. Now pick up an element from the new set and denote it by a_2.

Consider the set

$$S = \{a_1, a_2\}.$$

Proceeding in this way, having chosen a_{k-1}, you can have the set

$$S = \{a_1, a_2, ... a_{k-1},\}.$$

This set is always non-empty because S is an infinite set. Hence, we can choose an element in this set. Denote the element by a_k. This can be done for each $k \in N$. Thus the set

$$\{a_1, a_2, ..., a_k, ...\}$$

is a denumerable subset of S and hence a countable subset of S. This proves the theorem.

The importance of this theorem is that it leads us to an interesting area of Cardinality of sets by which we can determine and compare the relative sizes of various infinite sets,

This, however, is beyond the scope of this course.

Self Assessment

Fill in the blanks:

1. Let E denote the set of all even natural numbers. Then the mapping f: N → E defined as .... is a one-one correspondence. Hence, the set E of even natural numbers is a denumerable set and hence a countable set.
2. Let D denote the set of all odd integers and E the set of even integers. Then the, mapping f: E → D , defined as .... is a one-one correspondence. Thus E ~ D, But, E – N, therefore D – N. Hence D is a denumerable set and hence a countable set.
3. Every .... of a denumerable set is denumerable.
4. The set of all rational numbers between [0, 1] is ....
5. The set of all .... numbers is countable.

2.4 Summary

- We have discussed the order-relations (inequalities) in the set R of real numbers. Given any two real numbers x and y, either x > y or x = y or x < y.

- This is known as the law of Trichotomy. Then we have stated a few properties with respect to the inequality '≤'. The first property states that the inequality ≤ is antisymmetric. The second states the transitivity of ≤. The third allows us to add or subtract across the inequality, while preserving the inequality. The last property gives the situation in which the inequality is preserved if multiplied by a positive real number, while it is reversed if multiplied by a negative real number.
- We have also defined the bounded and unbounded intervals. The bounded intervals are classified as open intervals, closed intervals, semi-openor semi-closed intervals. The unbounded intervals are introduced with the help of the extended real number system $R \cup \{-\infty, \infty)$ using the symbols $+\infty$ (called plus infinity) and $-\infty$, (called minus infinity).
- There are three important aspects of the real numbers: algebraic, order and the completeness. To describe these aspects, we have specified a number of axioms in each case. In the algebraic aspect, an algebraic structure called the field is used. A field is a non-empty set F having at least two distinct elements 0 and 1 together with two binary operations + (addition) and . (multiplication) defined on F such that both + and . are commutative, associative, 0 is the additive identity, 1 is the multiplicative identity, additive inverse exists for each element of F, multiplicative inverse exists for each element other than 0 and multiplication is distributive over addition. The second aspect is concerned with the Order Structure in which, we use the axioms of the law of trichotomy, the transitivity property, the property that preserve the inequality under addition and the property that preserve the inequality under multiplication by a positive real number.
- In the completeness aspect, we introduce the notions of the supremum (or infimum) of a set and state the axiom of completeness. We find that both Q and R are ordered fields but the axioms of completeness distinguishes Q from R in the sense that Q does not satisfy the axiom of completeness. Thus, we conclude that R is a complete-ordered Field whereas Q is not a complete-ordered field.
- We introduce the notion of the countability of sets. A set is said to be denumerable if it is in one-one correspondence with the set of natural numbers. Any set which is either finite or denumerable is called a countable set. We have shown that the sets N, Z Q are countable sets but the sets 1 (set of irrational numbers) and R are not countable.
- Thus in this unit, we have discussed the algebraic structure, the order structure and the countability of the real numbers.

2.5 Keywords

Countable Set: A set which is equivalent to the set of natural numbers is called a denumerable set. Any set which is either finite or denumerable, is called a Countable set.

Uncountable Set: Any set which is not countable is said to be an uncountable set.

2.6 Review Questions

1. State the properties of order relation in the set R of real numbers with respect to the relation 3 (is greater than or equal to) and illustrate the inequality under multiplication by a negative real number.
2. Give examples to show that the intersection of any two intervals may not be an interval. What happens, if the two intervals are not disjoint? Justify your answer by an example.

Notes

3. Show that the set {0, 1} forms a field under the operations '+' and '.' defined by the following tables:

+	0	1
0	0	1
1	1	0

.	0	1
0	0	0
1	0	1

4. Show that the zero and the unity are unique in a field.

5. Do the sets N (of natural numbers) and Z (set of integers) form fields? Justify your answers. Also verify that the set C of complex numbers is a field.

6. Show that the field C of Complex numbers is not an ordered field.

7. (i) Define a set which is bounded below. Also define a lower bound of a set.

   (ii) Give at least two examples of a set (one of an infinite set) which is bounded below and mention a lower bound in each case.

   (iii) Is the set of negative integers bounded below? Justify your answers.

8. Test which of the following sets are bounded above, bounded below, bounded and unbounded.

   (i) The intervals]a, b], [a, b],]a, b] and [a, b[, where a and b are any two real numbers.

   (ii) The intervals $[2, \infty[$, $]-\infty, 3[$, $]5, \infty[$ and $]-\infty, 4]$.

   (iii) The set {cos e, cos 2 θ, cos 3 e, ...}.

   (iv) $S = \{x \in R : -a \le x \le a\}$ for some $a \in R$.

Answers: Self Assessment

1. f(n) = 2n
2. f(n) = n + 1
3. infinite subset
4. countable
5. positive rational

2.7 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 3: Matric Spaces

Notes

CONTENTS

Objectives
Introduction
3.1 Matric Spaces
3.1.1 Space Properties
3.1.2 Distance between Points and Sets; Hausdorff Distance and Gromov Metric
3.1.3 Product Metric Spaces
3.2 Modulus of Real Number
3.2.1 Properties of the Modulus of Real Number
3.3 Neighbourhoods
3.4 Open Sets
3.5 Limit Point of a Set
3.5.1 Bulzano Weierstrass Theorem
3.6 Closed Sets
3.7 Compact Sets
3.8 Summary
3.9 Keywords
3.10 Review Questions
3.11 Further Readings

Objectives

After studying this unit, you will be able to:

- Define the modulus of a real number
- Describe the notion of a neighbourhood of a point on the line
- Define an open set and give examples
- Discuss the limit points of a set
- Define a closed set and establish its relation with an open set
- Explain the meaning of an open covering of a subset of real numbers

Introduction

You all are quite familiar with an elastic string or a rubber tube or a spring. Suppose you have an elastic string. If you first stretch it and then release the pressure, then the string will come back to its original length. This is a physical phenomenon but in Mathematics, we interpret it differently. According to Geometry, the unstreched string and the stretched string are different since there is a change in the length. But you will be surprised to know that according to another branch of Mathematics, the two positions of the string are identical and there is no change. This branch is known as Topology, one of the most exciting areas of Mathematics.

Notes

The word 'topology' is a combination of the two Greek words 'topos' and 'logos'. The term 'topos' means the top or the surface of an object and 'logos' means the study. Thus 'topology' means the study of surfaces. Since the surfaces are directly related to geometrical objects, therefore there is a close link between Geometry and Topology. In Geometry, we deal with shapes like lines, circles, spheres, cubes, cuboids, etc. and their geometrical properties like lengths, areas, volumes, congruences etc. In Topology, we study the surfaces of these geometrical objects and certain related properties which are called topological properties. What are these topological properties of the surfaces of a geometrical figure? We shall not answer this question at this stage. However, since our discussion is confined to the real line, therefore, we shall discuss this question pertaining to the topological properties of the real line. These properties are related to the points and subsets' of the real line such as neighbourhood of a point, open sets, closed sets, limit points of a set of the real line, etc. We shall, therefore, discuss these notions and concepts in this unit. However, prior to all these, we discuss the modulus of a real number and its relationship with the order relations or inequalities.

3.1 Matric Spaces

Definition

A metric space is an ordered pair (M, d) where M is a set and d is a metric on M, i.e., a function

$$d : M \times M \to \mathbb{R}$$

such that for any $x, y, z \in M$, the following holds:

1. $d(x, y) \geq 0$ (non-negative),
2. $d(x, y) = 0$ iff $x = y$ (identity of indiscernibles),
3. $d(x, y) = d(y, x)$ (symmetry) and
4. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

The first condition follows from the other three, since:

$$2d(x, y) = d(x, y) - d(y, x) \geq d(x, x) = 0$$

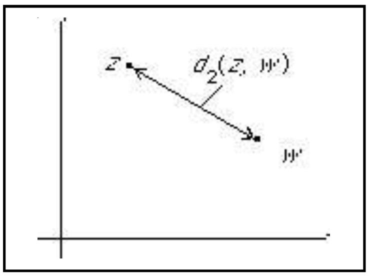
The function d is also called distance function or simply distance. Often, d is omitted and one just writes M for a metric space if it is clear from the context what metric is used.

Example:

1. The prototype: the line R with its usual distance $d(x, y) = |x - y|$.
2. The plane R^2 with the "usual distance" (measured using Pythagoras's theorem):

   $d((x_1, y_1), (x_2, y_2)) = \sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2]}$.

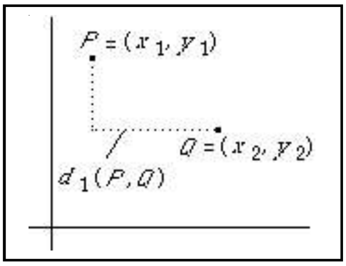
   This is sometimes called the 2-metric d_2.



3. The same picture will give metric on the complex numbers C interpreted as the Argand diagram. In this case the formula for the metric is now: $d(z, w) = |z - w|$ where the | | in the formula represent the modulus of the complex number rather than the absolute value of a real number.

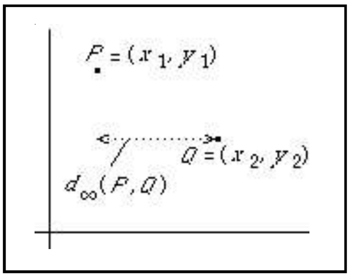
4. The plane with the taxi cab metric $d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|$.

   This is often called the 1-metric d_1.

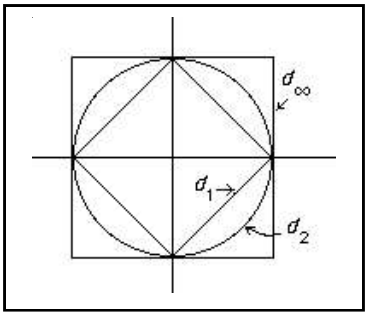


5. The plane with the supremum or maximum metric $d((x_1, y_1), (x_2, y_2)) = \max(|x_1 - x_2|, |y_1 - y_2|)$. It is often called the infinity metric d_∞.

   These last examples turn out to be used a lot. To understand them it helps to look at the unit circles in each metric. That is the sets $\{x \in R^2 \mid d(0, x) = 1\}$. We get the following picture:



6. Take X to be any set. The discrete metric on the X is given by: $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ otherwise. Then this does define a metric, in which no distinct pair of points are "close". The fact that every pair is "spread out" is why this metric is called discrete.



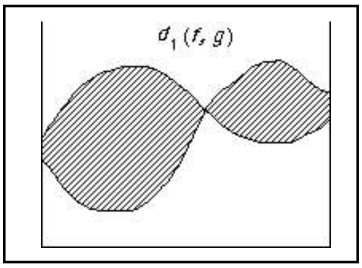
7. Metrics on spaces of functions. These metrics are important for many of the applications in analysis. Let C[0, 1] be the set of all continuous R-valued functions on the interval [0, 1]. We define metrics on by analogy with the above examples by:

   (a) $d_1(f, g) = \int_0^1 |f(x) - g(x)|\, dx$

   So the distance between functions is the area between their graphs.

Notes

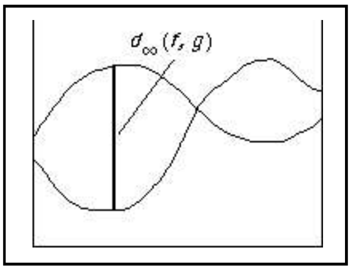
(b) $d_2(f, g) = \sqrt{\left[\int_0^1 f(x) - g(x)^2 dx\right]}$



Although this does not have such case straight forward geometric interpretation as the last example, this case turns out to be the most important in practice. It corresponds to who doing a "least squares approximation".

(c) $d(f, g) = \max \{|f(x) - g(x))| \mid 0 \le x \le 1\}$

Geometrically, this is the largest distance between the graphs.



Remarks:

1. The triangle inequality does hold for these metrics
2. As in the R^2 case one may define d_p for any $p \ge 1$ and get a metric.

3.1.1 Space Properties

Metric spaces are paracompact Hausdorff spaces and hence normal (indeed they are perfectly normal). An important consequence is that every metric space admits partitions of unity and that every continuous real-valued function defined on a closed subset of a metric space can be extended to a continuous map on the whole space (Tietze extension theorem). It is also true that every real-valued Lipschitz-continuous map defined on a subset of a metric space can be extended to a Lipschitz-continuous map on the whole space.

Metric spaces are first countable since one can use balls with rational radius as a neighborhood base.

The metric topology on a metric space M is the coarsest topology on M relative to which the metric d is a continuous map from the product of M with itself to the non-negative real numbers.

3.1.2 Distance between Points and Sets; Hausdorff Distance and Gromov Metric

A simple way to construct a function separating a point from a closed set (as required for a completely regular space) is to consider the distance between the point and the set. If (M, d) is a metric space, S is a subset of M and x is a point of M, we define the distance from x to S as

$$d(x, S) = \inf \{d(x, s) : s \in S\}$$

Then d(x, S) = 0 if and only if x belongs to the closure of S. Furthermore, we have the following generalization of the triangle inequality:

$$d(x, S) \le d(x, y) + d(y, S)$$

which in particular shows that the map is continuous.

Given two subsets S and T of M, we define their Hausdorff distance to be

$$d_H(S, T) = \max \{\sup \{d(s, T) : s \in S\}, \sup \{d(t, S) : t \in T\}\}$$

In general, the Hausdorff distance $d_H(S, T)$ can be infinite. Two sets are close to each other in the Hausdorff distance if every element of either set is close to some element of the other set.

The Hausdorff distance d_H turns the set K(M) of all non-empty compact subsets of M into a metric space. One can show that K(M) is complete if M is complete. (A different notion of convergence of compact subsets is given by the Kuratowski convergence.)

One can then define the Gromov–Hausdorff distance between any two metric spaces by considering the minimal Hausdorff distance of isometrically embedded versions of the two spaces. Using this distance, the set of all (isometry classes of) compact metric spaces becomes a metric space in its own right.

3.1.3 Product Metric Spaces

If (M_1, d_1), ..., (M_n, d_n) are metric spaces, and N is the Euclidean norm on R^n, then (M1 x .. x Mn, N(d1, ..., dn)) is a metric space, where the product metric is defined by

$$N(d1, ..., dn)\ ((x1, ..., xn), (y1, . . ., yn)) = N(d1(x1, y1), . . ., dn(xn, dn)),$$

and the induced topology agrees with the product topology. By the equivalence of norms in finite dimensions, an equivalent metric is obtained if N is the taxicab norm, a p-norm, the max norm, or any other norm which is non-decreasing as the coordinates of a positive n-tuple increase (yielding the triangle inequality).

Similarly, a countable product of metric spaces can be obtained using the following metric

$$d(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{di(x_i, y_i)}{1 + d_i(x_i, y_i)}.$$

An uncountable product of metric spaces need not be metrizable. For example, R^R is not first-countable and thus isn't metrizable.

Continuity of Distance

It is worth noting that in the case of a single space (M, d), the distance map d: $M \times M \to R^+$ (from the definition) is uniformly continuous with respect to any of the above product metrics N(d, d), and in particular is continuous with respect to the product topology of M × M.

Quotient Metric Spaces

If M is a metric space with metric d, and ~ is an equivalence relation on M, then we can endow the quotient set M/~ with the following (pseudo)metric. Given two equivalence classes [x] and [y], we define

$$d'([x], [y]) = \inf \{d(p_1, q_1) + d(p_2, q_2) + . . . + d(p_n, q_n)$$

where the infimum is taken over all finite sequences ($p_1, p_2, \ldots, p_n$) and ($q_1, q_2, \ldots, q_n$) with $[p_1] = [x]$, $[q_n] = [y]$, $[q_i] = [p_i + 1]$, i = 1, 2, . . ., n – 1. In general this will only define a pseudometric, i.e. $d'([x], [y]) = 0$ does not necessarily imply that $[x] = [y]$. However for nice equivalence relations (e.g., those given by gluing together polyhedra along faces), it is a metric. Moreover if M is a compact space, then the induced topology on M/~ is the quotient topology.

The quotient metric d is characterized by the following universal property. If $f : (M, d) \to (X, \delta)$ is a metric map between metric spaces (that is, $\delta(f(x), f(y)) \leq d(x, y)$ for all (x, y) satisfying $f(x) = f(y)$ whenever $x \sim y$, then the induced function $\bar{f} : M/\sim \to X$, given by $\bar{f}([x]) = f(x)$, is a metric map $\bar{f} : (M/\sim, d') \to (X, \delta)$. A topological space is sequential if and only if it is a quotient of a metric space

Generalizations of Metric Spaces

- Every metric space is a uniform space in a natural manner, and every uniform space is naturally a topological space. Uniform and topological spaces can therefore be regarded as generalizations of metric spaces.
- If we consider the first definition of a metric space given above and relax the second requirement, or remove the third or fourth, we arrive at the concepts of a pseudometric space, a quasimetric space, or a semi-metric space.
- If the distance function takes values in the extended real number line $R \cup \{+\infty\}$, but otherwise satisfies all four conditions, then it is called an extended metric and the corresponding space is called an ∞-metric space.
- Approach spaces are a generalization of metric spaces, based on point-to-set distances, instead of point-to-point distances.
- A continuity space is a generalization of metric spaces and posets, that can be used to unify the notions of metric spaces and domains.

Metric Spaces as Enriched Categories

The ordered set ($\mathbb{R}, \geq$) can be seen as a category by requesting exactly one morphism $a \to b$ if $a \geq b$ and none otherwise. By using + as the tensor product and 0 as the identity, it becomes a monoidal category R*. Every metric space (M, d) can now be viewed as a category M* enriched over R*:

- Set $Ob(M^*) := M$
- For each set $X, Y \in M$ set $Hom(X, Y) := d(X, Y) \in Ob(R^*)$.
- The composition morphism $Hom(Y, Z) \otimes Hom(X, Y) \to Hom(X, Z)$ will be the unique morphism in R* given from the triangle inequality $d(y, z) + d(x, y) \geq d(x, z)$.
- The identity morphism $0 \to Hom(X, X)$ will be the unique morphism given from the fact that $0 \geq d(X, X)$.
- Since R* is a strict monoidal category, all diagrams that are required for an enriched category commute automatically.

3.2 Modulus of Real Number

You know that a real number x is said to be positive if x is greater than 0. Equivalently, if 0 represents a unique point 0 on the real line, then a positive real number x lies on the right side of 0. Accordingly, we defined the inequality $x > y$ (in terms of this positivity of real numbers) if

$x - y > 0$. You will recall from Section 2.2 that for the validity of the properties of order relations or the inequalities. Such as the one concerning the multiplication of inequalities, it is essential to specify that some of the numbers involved should be positive. For example, it is necessary that $z > 0$ so that $x > y$ implies $xz > yz$. Again, the fractional power of a number will not be real if the number is negative, for instance $x^{1/2}$ when $x = -4$. Many of the fundamental inequalities, which you may come across in higher Mathematics, will involve such fractional powers of numbers. In this context, the concept of the absolute value or the modulus of a real member is important to which you are already familiar. Nevertheless, in this section, we recall the notion of the modulus of a real number and its related properties which we need for our subsequent discussion.

Defination: Modulus of Real Number

Let x be any real number. The absolute value or the modulus of x denoted by $|x|$ is defined as follows:

$|x| = x$ if $x > 0$

$= -x$ if $x < 0$

$= 0$ if $x = 0$.

You can easily see that

$|x| = |x|, \forall x \in R$.

Not that $|-x|$ is different from $-|x|$.

3.2.1 Properties of the Modulus of Real Number

Since the modulus of a real number is essentially a non-negative real number, therefore the operations of usual addition, subtraction, multiplication and division can be performed on these numbers. The properties of the modulus are mostly related to these operations.

Property 1: For any real number x, $|x|$ = Maximum of (x, – x),

Proof: Since x is any real number, therefore either $x \geq 0$ or $x < 0$. If $x \geq 0$, then by definition, we have

$|x| = X$.

Also, $x > 0$ implies that $-x \leq 0$. Therefore, maximum of $(x, -x\} = x = |x|$

Again $x < 0$, implies that $-x > 0$. Therefore again maximum of $\{x, -x\} = -x = |x|$.

Thus,

Maximum $(x, -x\} = |x|$

Now consider the numbers $|5|^2$, $|-4.5|$, $\left|\frac{4}{5}\right|$. It is easy to see that

$|5|^2 = |5| = 5.5 = 5^2 = |-5|^2$

$|-4.5| = |-20| = 20$ Also $|-4|.\ |5| = 4.5 = 20$

i.e. $|-4.5| = |-4|.\ |5|$

and

$\left|\frac{4}{5}\right| = \frac{4}{5}$ and $\frac{|4|}{|5|} = \frac{4}{5}$ i.e.

Notes

$$\left|\frac{4}{5}\right| = \frac{|4|}{|5|}.$$

All this lead us to the following properties:

Property 2: For any real number x

$|x|^2 = x^2 = |-x|^2$

Proof: We know that $|x| = x$ for x 2'0.

Thus $|x|^2 = |x|\,|x| = x, x = x$, for $x \geq 0$

Again for $x < 0$, we know that $|x| = -x$. Therefore

$|x^2| = |x|\;|x| = -x. -x = x^2$

Therefore, it follows that

$|x|^2 = x^2$ for any $x \in R$.

Now you should try the other part as an exercise.

Property 3: For any two real numbers x and y, prove that $|x.y| = |x|.|y|$.

Proof: Since x and y are any two real numbers, therefore, either both are positive or one is positive and the other is negative or both are negative i.e. either $x \geq 0, y \geq 0$ or $x \geq 0, y \geq 0$ or $x \geq 0, y \geq 0$ or $x \leq 0, y \leq 0$. We discuss the proof for all the four possible cases separately.

Case (i): When $x \geq 0, y \geq 0$.

Since $x \geq 0$, therefore, we have, by definition,

$|x| = x, |y| = y$

Also $x \geq 0, y \geq 0$ simply that $xy \geq 0$ and hence

$|xy| = xy = |x|\,|y|$

which proves the property.

Case (ii): When $x \geq 0, y \leq 0$. Then obviously $xy \leq 0$. Consequently by definition, it, follows that

$|x| = x, |y| = -Y, |xy| = -xy$

Hence

$|xy| = -xy = x(-y) = |x|\,|y|$

which proves the property.

Case (iii): When $x \leq = 0, y \geq 0$.

Interchange x and y in (ii).

***Case (iv)*:** When $x \leq 0, y \leq 0$, then xy = 20. Accordingly, we have

$|x| = -x, |y| = -y, |xy| = xy.$

Hence

$|xy| = xy = (-x)(-y) = |x|\,|y|$

using the field properties stated. This concludes the proof of the property.

Notes

Alternatively, the proof can be given by using property 2 in following way:

$|xy|^2 = (xy)^2 = x^2 y^2 = |x|^2.|y|^2$

$= (|x|.|y|)^2$

Therefore

$|xy| = \pm(|x| |y|)$

Since $|xy|$, $|x|$ and $|y|$ are non-negative, therefore we take the positive sign only and we have

$|xy| = |x| |y|$

which proves the property.

You can use any of the two methods to try the following exercise.

The next property is related to the modulus of the sum of two real members. This is one of the most important properties and is known as Triangular Inequality:

Property 4: Triangular Inequality

For any two real numbers x and y, prove that

$|x + y| \leq |x| + |y|$.

Proof: For any two real numbers x and y the number $x + y \geq 0$ or $x + y < 0$.

If $x + y \geq 0$, then by definition

$|x + y| = x + y$. . . .(1)

Also, we know that

$|x| \geq x \quad \forall x \in R$

$|x| \geq x \quad \forall y \in R$

Therefore

$|x| + |y| \geq x + y$

$x + y \leq |x| + |y|$. . . .(2)

From (1) and (2), it follows that

$|x + y| \leq |x| + |y|$

Now, if $x + y < 0$, then again by definition, we have.

$|x + y| = -(x + y)$

or $|x + y| = (-x) + (-y)$ . . .(3)

Also we know that (see property 1)

$-x \leq |x|$ and $-y \leq |y|$.

Consequently, we get

$(-x) + (-y) \leq |x| + |y|$

or $(-x) + (-y) \leq |x| + |y|$ . . .(4)

From (3) and (4), we get

$|x + y| \leq |x| + |y|$

Notes

This concludes the proof of the property.

You can try the following exercise similar to this property.

Now let us see another interesting relationship between the inequalities and the modulus of a real number.

By definition, $|x|$ is a non-negative real number for any $x \in R$. Therefore, there always exists a non-negative real number u such that

either $(x| < u$ or $|x| > u$ or $|x| = u$.

Suppose $|x| < u$. Let us choose u = 2. Then

$|x| < C \Rightarrow$ Max. $\{-x, x\} < 2$

$\Rightarrow -x < 2, x < 2$

$\Rightarrow x > -2, x < 2$

$\Rightarrow -2 < x, x < 2$

$\Rightarrow -2 < x < 2.$

i.e. $|x| < 2 \Rightarrow -2 < x < 2$

Conversely, we have

$-2 < x < 2 \Rightarrow -2 < x < 2$

$\Rightarrow 2 > -x, x < 2$

$\Rightarrow -x < 2, x < 2$

$\Rightarrow$ Max. $\{-x, x\} < 2$

$\Rightarrow |x| < 2.$

i.e.

$-2 < x < \Rightarrow |x| < 2$

Thus, we have shown that

$|x| < \Leftrightarrow -2 < x < 2.$

This can be generalised as the following property.

Property 5: Let x and u be any two real numbers.

$|x| \leq u \Leftrightarrow -u \leq x \leq u.$

Proof: $|x| \leq u \Leftrightarrow$ Max. $\{-x, x\} \leq u$

$\Leftrightarrow -x \leq u, x \leq u$

$\Leftrightarrow x \geq -u, x \leq u$

$\Leftrightarrow -u \leq x, x \leq u$

$\Leftrightarrow -u \leq x \leq u$

which proves the desired property.

The property 5 can be generalized in the form of the following exercise.

Notes

Task For any real numbers x, a and d,

$|x - a| \leq d \Leftrightarrow a - d \leq x \leq a + d$.

Example: Write the inequality $3 < x < 5$ in the modulus form.

Solution: Suppose that there exists real numbers a and b such that

$a - b = 3, a + b = 5.$

Solving these equations for a and b, we get

$a = 4 > b = 1.$

Accordingly,

$3 < x < 5 \Leftrightarrow 4 - 1 < x < 4 + 1$

$\Leftrightarrow -1 < x - 4 < 1$

$\Leftrightarrow |x - 4| < 1$

Task

1. Write the inequality $2 < x < 7$ in the modulus form.
2. Convert $|x - 2| < 3$ into the corresponding inequality.

3.3 Neighbourhoods

You are quite familiar with the word 'neighbourhood'. You use this word frequently in your daily life. Loosely speaking, a neighbourhood of a given point c on the real line is a set of all those points which are close to c. This is the notion which needs a precise meaning. The term 'close to' is subjective and therefore must be quantified. We should clearly say how much 'close to'. To elaborate this, let us first discuss the notion of a neighbourhood of a point with respect to a (small) positive real number δ.

Let c be any point on the real line and let 6 > 0 be a real number. A set consisting of all those points on the real line which are at a distance of 6 from c is called a neighbourhood of c. This set is given by

$\{x \in R : |x - c| < 6\}$

$= \{x \in R : c - \delta < x < c + 6)$

$=]c - 6, c + \delta[$

Which is an open interval. Since this set depends upon the choice of the positive real number δ, we call it a 6-neighbourhood of the point c.

Thus, a δ-neighboured of a point c on the real line is an open interval]c – 6, c + 6[, δ > 0 while c is the mid point of this neighbourhood. We can give the general definition of neighbourhood of a point in the following way.

Notes

Neighbourhood of a Point

A set P is said to be a Neighbourhood (NBD) of a point V if there exists an open interval which contains c and is contained in P.

This is equivalent to saying that there exists an open interval of the form]c – δ, c + δ[, for some 6 > 0, such that

$$]c - 6, c + \delta[\subset P.$$

Example: (i) Every open interval]a, b[is a NBD of each of its points.

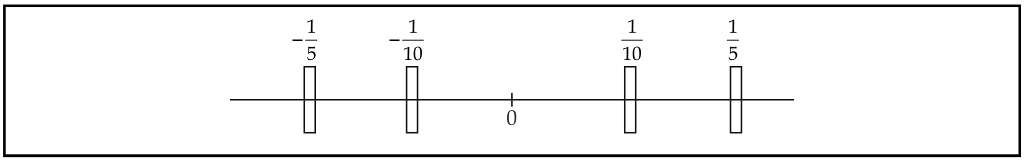
(ii) A closed interval [a, b] is a NBD of each of its points except the end point i.e. [a, b] is not a NBD of the points a and b, because it is not possible to find an open interval containing a or b which is contained in [a, b]. For instance, consider the closed interval [0,1]. It is a NBD of every point in]0, l[. But, it is not a NBD of 0 because for every δ > 0,]-δ, δ[$\not\subset$ [0, 1]. Similarly [0, 1] is not a NBD of 1.

(iii) The null set 0 is a NBD of each of its point in the sense there is no point in 0 of which it is not a NBD.

(iv) The set R of real numbers is a NBD of each real number x because for every 5 > 0, the open interval]x - 6, x + δ[is contained in R.

(v) The set Q of rational numbers is not a NBD of any of its points x because any open interval containing x will also contains an infinite number of irrational numbers and hence the open interval can not be a subset of Q.

Now consider any two neighbourhoods of the point 0 say $]-\frac{1}{10}, \frac{1}{10}[$ and $]-\frac{1}{5}, \frac{1}{5}[$ as shown in the Figure below.



The intersection, of these two neighbourhood is

$$]-\frac{1}{10}, \frac{1}{10}[\cap]-\frac{1}{5}, \frac{1}{5}[=]-\frac{1}{10}, \frac{1}{10}[$$

which is again a NBD of 0. The union of these two neighbourhoods is $]-\frac{1}{5}, \frac{1}{5}[$, which is also a NBD of 0. Let us now examine these results in general.

Example: The intersection of any two neighbourhoods of a point is a neighbourhood of the point.

Solution: Let A and B be any two NBDS of a point c in R. Then there exist open intervals $]c - \delta_1$, $c + \delta_1]$ and $]c - \delta_2, c + \delta_2$ [such that]$c - \delta_1, c + \delta_1] \subset A$, for some $\delta_1 > 0$, and $]c - \delta_2, c + \delta_2 [\subset B$, for some $\delta_2 > 0$.

Let 6 = Min. $\{\delta_1, \delta_2)$ = minimum of δ_1, δ_2.

This implies that]c – 6, c + δ [C A ∩ B which shows that A ∩ B is a NBD of c.

Notes

Example: Show that the superset of a NBD of a point is also a NBD of the point.

Solution: Let A be a NBD of a point c. Then there exists an open interval]c – 6, – c + 6[, for some 6 > 0 such that

]c – 6, c + δ[C A.

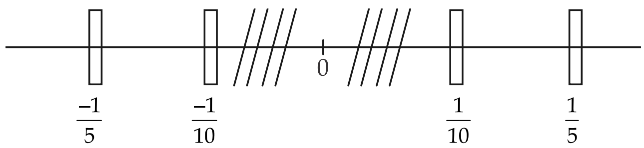
Now let S be a super set which contains A. Then obviously

$A \subset S \Rightarrow [c - \delta, c + \delta[\subset S$

which shows that S is also a NBD of c.

For instance, if] $\frac{1}{10}$, $\frac{1}{10}$ [is a NBD of the point 0.

Then,] – $\frac{1}{5}$ $\frac{1}{5}$ [is also a NBD of 0 as can be seen from Figure below.



Is a subset of a NBD of a point also a NBD of the point? Justify your answer.

Now you can try the following exercise.

Task Prove that the Union of any two NBDS of a point is a NBD of the point.

The conclusion of the Exercise, in fact, can be extended to a finite or an infinite or an arbitrary number of the NBDS of a point.

However, the situation is not the same in the case of intersection of the NBDS. It is true that the intersection of a finite number of NBDS of a point is a NBD of the point. But the intersection of an infinity collection of NBDS of a point may not be a NBD of the point. For example, consider the class of NBDS given by a family of open intervals of the form

$I_1 =]-1, 1[, I_2 =]-\frac{1}{2}, \frac{1}{2}[\ I_3 =]-\frac{1}{3}, \frac{1}{3}[,$.

$I_n =]-\frac{1}{n}, \frac{1}{n}[...$

which are NBDS of the point 0. Then you can easily verify that

$I_1 \cap I_2 \cap I_3 \cap I_4 \cap .... \cap I_n \cap .....$

or $\bigcap_{n=1}^{\infty} I_n = \{0\}$

3.4 Open Sets

You have seen from the previous examples and exercises that a given set may or may not be a NBD of a point. Also, a set may be a NBD of some of its points and not of its other points. A set

Notes

may even be a NBD of each of its points as in the case of the interval]a, b[. Such a set is called an open set.

Definition: X a set S is said to be open if it is a neighbourhood of each of its points.

Thus, a set S is open if for each x in S, there exists an open interval]x – 6, x + δ[, δ > 0 such that

$x \in]\, x - \delta, x + \delta [\cap S.$

It follows at once that a set S is not open if it is not a NBD of even one of its points.

Example: An open interval is an open set

Solution: Let]a, b[be an open interval. Then a < b. Let c∈] a, b[. Then a < c < b and therefore

$c - a > 0$ and $b - c > 0$

Choose

δ = Minimum of {b – c, c – a)

= Min (b – c, c – a).

Note that $b - c > 0$, $c - a > 0$. Therefore $\delta > 0$.

Now $\delta \le c - a \Rightarrow a \le c - \delta$

and $\delta \le b - c \Rightarrow c + \delta < b$.

i.e.

Therefore,]c – 6, c + δ[⊂]a, b[and hence]a, b[is a NBD of c.

Example: (i) The sat R of real numbers is an open set

(ii) The null set f is an open set

(iii) A finite set is not an open set

(iv) The interval]s, b] is not an open set.

Example: Prove that the intersection of any two open sets is an open set.

Solution: Let A and B be any two open sets. Then we have to show that $A \cap B$ is also an open set. If $A \cap B = \phi$, then obviously $A \cap B$ is an open set. Suppose $A \cap B \neq \phi$.

Let x be an arbitrary element of $A \cap B$. Then $x \in A \cap B \Rightarrow x \in A$ and $x \in B$.

Since A and B are open sets, therefore A and B are both NBDS of x. Hence $A \cap B$ is a NBD of x. But $x \in A \cap B$ is chosen arbitrarily. Therefore, $A \cap B$ is a NBD of each of its points and hence $A \cap B$ is an open set. This proves the result. In fact, you can prove that the intersection of a finite number of open sets is an open set. However, the intersection of an infinite number of open sets may not be an open set.

Task

1. Give an example to show that intersection of an infinite number of open sets need not be an open set.
2. Prove that the union of any two open sets is an open set. In fact, you can show that the union of an arbitrary family of open sets is an open set.

Notes

3.5 Limit Point of a Set

You have seen that the concept of an open set is linked with that of a neighbourhood of a point on the real line. Another closely related concept with the notion of neighbourhood is that of a limit point of a set. Before we explain the meaning of limit point of a set, let us study the following situations:

(i) Consider a set S = [1, 2[> Obviously the number 1 belongs to S. In any NBD of the point 1, we can always find points of S which are different from 1. For instance]0-5, l[is a NBD of 1. In this NBD, we can find the point 1.05 which is in S but at the same time we note that $1.05 \neq 1$,

(ii) Consider another set S = $\left\{\frac{1}{n} : n \in N\right)$. The number 0 does not belong to this set.

Take any NBD of 0 say,] –0.1, 0.1 [. The number $\frac{1}{20}$ = 0.05 of S is in this NBD of 0. Note that $0.05 \neq 0$.

(i) Again consider the same set S of (ii) in which the number 1 obviously belongs to S. We can find a NBD of 1, say]0.9, 1.1[in which we can not find a point of S different from 1.

In the light of the three situations, we are in a position to define the following:

Limit Point of a Set

A number p is said to be a limit point of a set S of real numbers if every neighbourhood of p contains at least one point of the set S different from p.

Examples: (i) In the set S = [1, 2[, the number 1 is a limit point of S This limit point belongs to S. The set S = $\left\{\frac{1}{n} : n \in N\right)$ has only one limit point 0. You may note that 0 does not belong to S.

(ii) Every point in Q, (the set of rational numbers), is a limit point of Q, because for every rational number r and $\delta > 0$, i.e.] r – 6, r + δ [has at least one rational number different from r. This is because of the reason that there are infinite rationals between any two real numbers. Now, you can easily see that every irrational number is also a limit point of the set Q for the same reason.

(iii) The set N of natural numbers has no limit point because for every real number a, you can always find $\delta > 0$ such that]a – 6, a + δ[does not contain a point of the set N other than a.

(iv) Every point of the interval]a, b] is its limit point. The end points a and b are also the limit points of]a, b]. But the limit point a does not belong to it whereas the limit point b belongs to it.

(v) Every point of the set [a, ∞[is a limit point of the sets. This is also true for] –∞, a[.

From the foregoing examples and exercises, you can easily observe that

(i) A limit point of set may or may not belong to the set,

(ii) A set may have no limit point,

(iii) A set may have only one limit point.

(iv) A set may have more than one limit point.

Notes

The question, therefore, arises: "How to know whether or not a set has a limit point?" One obvious fact is that a finite set can not have a limit point. Can you give a reason for it? Try it. But then there are examples where even an infinite set may not have a limit point e.g. the sets N and Z do not have a limit point even though they are infinite 'sets. However, it is certainly clear that a set which has a limit point, must necessarily be an infinite set. Thus our question takes the following form:

"What are the conditions for a set to have a limit point?"

This question was first studied by a Czechoslovakian Mathematician, Bemhard Bulzano [1781-1848] in 1817 and he gave some ideas.

Unfortunately, his ideas were so far ahead of their time that the world could not appreciate the full significance of his work. It was only much later that Bulzario's work was extended by Karl Weierstrass [1815-1897], a great German Mathematician, who is known as the "father of analysis". It was in the year 1860 that Weierstrass proved a fundamental result, now known as Bulzano-Weierstrass Theorem for the existence of the limit points of a set. We state and prove this theorem as follows.

3.5.1. Bulzano Weierstrass Theorem

Theorem 1: Every infinite bounded subset of set R has a limit point (in K).

Proof: Let S be an infinite and bounded subset of R. Since A is bounded, therefore A has both a- lower bound as well as an upper bound.

Let m be a lower bound and M be an upper bound of A. Then obviously

$m \le x \le M, \ \forall x \in A.$

Construct a set S in the following way:

S = (x∈R: x exceeds at most finite number of the elements of A}. Now, let us examine the following two questions:

(i) Is S a non-empty set?

(ii) Is S also a bounded set?

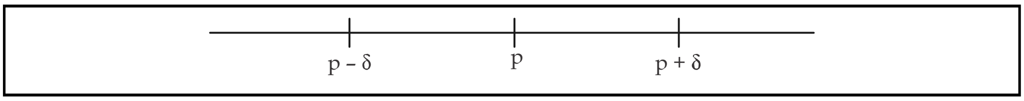
Indeed, S is non-empty because m ≤ x, ≤ M, ∀ x∈A, implies that m∈S. Also M is an upper bound of S because no number greater than or equal to M can belong to S. Note that M cannot belong to S because it exceeds an infinite number of elements of A.

Since the set S is non-empty and bounded above, therefore, by the axiom of completeness, S has its supremum in R. Let p be the supremum of S. We claim that p is a limit point of the set A.

In order to show that p is a limit point of A, we must establish that every NBD of p has at least one point of the set A other than p. In other words, we have to show that every NBP of p has an infinite number of elements of A. For this, it is enough to show that any open interval $]p - \delta, p + \delta[$, for $\delta > 0$, contains an infinite number of members of set A. For this, we proceed as follows.

Since p is the supremum of S, therefore, by the definition of the Supremum of a set, there is at least one element y in S such that $y > p - \delta$, for $\delta > 0$. Also y is a member of S, therefore, y exceeds at the most a finite number of the elements of A. In other words, if you visualise it on the line as shown in the Figure below, the number of elements of A lying on the left of $p - \delta$ is finite at the most. But certainly, the number of elements of A lying on the right side of the point $p - \delta$ is infinite.

Notes



Again since p is the supremum of S, therefore, by definition p + δ can not belong to S. In other Words, p + δ exceeds an infinite number of elements of A. This means that there lie an infinite number of elements of A on the left side of the point p + δ.

Thus we have shown that there lies, an infinite number of elements of A on the right side of p – δ and also there ia an infinite number of elements of A on the left side of p + δ. What do you conclude from this? In other words, what is the number of elements of A in between (i.e., within) the interval]p – δ, p + δ[. Indeed, this number is infinite i.e., there is an infinite number of elements of A in the open interval]p + δ[. Hence the interval] p – δ, p + δ [contains an infinite number of elements of A for some δ > 0. Since δ > 0 is chosen arbitrarily, therefore every interval]p – δ, p + δ [has an infinite number of elements of A. Thus, every NBD-of p contains an infinite number of elements of A. Hence p is a limit point of the set A.

This completes the proof of the theorem.

Example: (i) The intervals [0, 1],]0,1[,] 0, 1], [0,1[are all infinite and bounded sets. Therefore each of these intervals has a limit point. In fact, each of these intervals has an infinite number of limit points because every point in each interval is a limit point of the interval.

(ii) The set [a, ∞[is infinite and unbounded set but has every point as a limit point. This shows that the condition of boundedness of an infinite set is only sufficient in the theorem.

From the previous examples and exercises, it is clear that it is not necessary for an infinite set to be bounded to possess a limit point. In other words, a set may be unbounded and still may have a limit point. However, for a set to have a limit point, it is necessary that it is infinite.

Another obvious fact is that a limit point of a set may or may not belong to the set and a set may have more than one limit point. We shall further study how sets can be characterized in terms of their limit points.

3.6 Closed Sets

You have seen that a limit point of a set may or may not belong to the set. For example, consider the set S = {x∈R : 0 ≤ 5 x < 1}. In this set, 1 is a limit point of S but it does not belong to S. But if you take S = {x: 0 ≤ x ≤ 1}, then all the limit points of S belong to S. Such a set is called a closed set. We define a closed set as follows:

Definition

A set is said to be closed if it contains all its limit points.

Example: (i) Every closed and bounded interval such as [a, b] and [0, 1] is a closed set.

(ii) An open interval is not a closed set. Check Why?

(iii) The set R is a closed set because every real number is a limit point of R and it belongs to R.

(iv) The null set ϕ is a closed set.

(v) The set S = $\left\{\frac{1}{n} : n \varepsilon N\right\}$ is not a closed set. Why?

(vi) The set]a, ∞[is not a closed set, but] –∞, a] is a closed set.

Notes

You may be thinking that the word open and closed should be having some link. If you are guessing some relation between the two terms, then you are hundred per cent correct. Indeed, there is a fundamental connection between open and closed sets. What exactly is the relation between the two? Can you try to find out? Consider, the following subsets of R:

(i) $]0, 4[$

(ii) $[-2, 5]$

(iii) $]0, = \infty[$

(iv) $]-\infty > 6]$.

The sets (i) and (iii) are open while (ii) and (iv) are closed. If you consider their complements, then the complements of the open sets are closed while those of the closed sets are open. In fact, we have the following concrete situation in the form of following theorem.

Theorem 2: A set is closed if and only if its complement is open.

Proof: We assume that S is a closed set. Then we prove that its complement S^c is open.

To show that S^c is open, we have to prove that S^c is a NBD of each of its points. Let $x \in S^C$. Then, $x \in S^c \Rightarrow x \notin S$. This means x is not a limit point of S because S is given to be a closed set. Therefore, there exists a $\delta > 0$ such that $]x - \delta, x + \delta[$ contains no points of S. This means that $]x - \delta, x + \delta[$ is contained in S^c. This further implies that S^c is a NBD of x. In other words, S^c is an open set, which proves the assertion.

Conversely, let a set S be such that its complement S^c is open. Then we prove that S is closed.

To show that S is closed, we have to prove that every limit point x of S belongs to S. Suppose $x \notin S$, Then $x \notin S^C$.

This implies that S^C is a NBD of x because S^c is open. This means that there exists an open interval $]x - \delta, x + \delta[$, for some $6 > 0$, such that

$$]x - \delta, x + \delta[\in S^c$$

In other words, $]x - \delta, x + \delta[$ contains no point of S. Thus x is not a limit point of S, which is a contradiction. Thus our supposition is wrong and hence, $x \notin S$ is not possible. In other words, the (limit) point x belongs to S and thus S is a closed set.

Note that the notions of open and closed sets are not mutually exclusive. In other words, if a set is open, then it is not necessary that it can not be closed. Similarly, if a set is closed, then it does not exclude the possibility of its being open. In fact, there are sets which are both open and closed and there are sets which are neither open nor closed as you must have noticed in the various examples we have given in our discussion. For example the set R of all the real numbers is both an open sets as well as a closed set. Can you give another example? What about the null set. Again Q, the set of rational numbers is neither open nor closed.

Task Give examples of two sets which are neither closed nor open.

We have discussed the behaviour of the union and intersection of open sets. Since closed sets are closely connected with open sets, therefore, it is quite natural that we should say something about the union and intersection of closed sets. In fact, we have the following results:

Example: Prove that the union of two closed sets is a closed set.

Solution: Let A and B be any two closed sets. Let $S = A \cup B$, we have to show that S is a closed set. For this, it is enough to prove that the complement S^c is open

Now

$$S^c = (A \cup B)^c = B^c \cap A^c = A^c \cap Bc$$

Since A and B are closed sets, therefore A^c and B^c are open sets. Also, we have proved in the intersection of any two open sets is open. Therefore $A^c \cap B^c$ is an open set and hence S is open.

This result can be extended to a finite number of closed sets. You can easily verify that the union of a finite number of closed sets is a closed set. But, note that the union of an arbitrary family of closed sets may not be closed.

For example, consider the family of closed sets given as

$$S_1 = [l, 2], S_2 = [\frac{1}{2}, 2], S_3 = [\frac{1}{3}, 2], ...$$

and in general

$$S_n = [\frac{1}{n}, 2] ..... \text{ for } n = 1, 2, 3, ...$$

Then,

$$\bigcup_{n=1}^{\infty} S_n = S_1 \cup S_2 \cup S_3 ..... \cup S_n \cup .....$$

$$= I0, 2]$$

which is not a closed set.

Definition: Derived Set

The set of all limit points of a given set S is called the derived set and is denoted by S'.

Example: (i) Let S be a finite set. Then $' = ϕ

(ii) $S = (\frac{1}{n} : n \in N)$, the derived set S' = {0}

(iii) The derived set of R is given by R' = R

(iv) The derived set of Q is given by Q' = R

We define another set connected with the notion of the limit point of a set. This is called the closure of a set.

Definition: Closure of a Set

Let S be any set of real numbers ($S \in R$). The closure of S is defined as the union of the set S and its derived set S. It is denoted by $\overline{S}$, Thus

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

In other words, the closure of a set is obtained by the combination of the elements of a given set S and its derived set S'.

Notes

For example, $\bar{S}$ of $S = \{\frac{1}{n} : n \in N\}$ is given by.

$$\bar{S} = \{\frac{1}{n}, n \in N\} \cup \{0\} = \{0, 1, \frac{1}{2}, \frac{1}{3}, ...\}$$

Similarly, you can verify that

$$\vec{Q} = Q \cup Q' = Q \cup R = R$$

$$\vec{R} = R \cup R' = R \cup R = R$$

3.7 Compact Sets

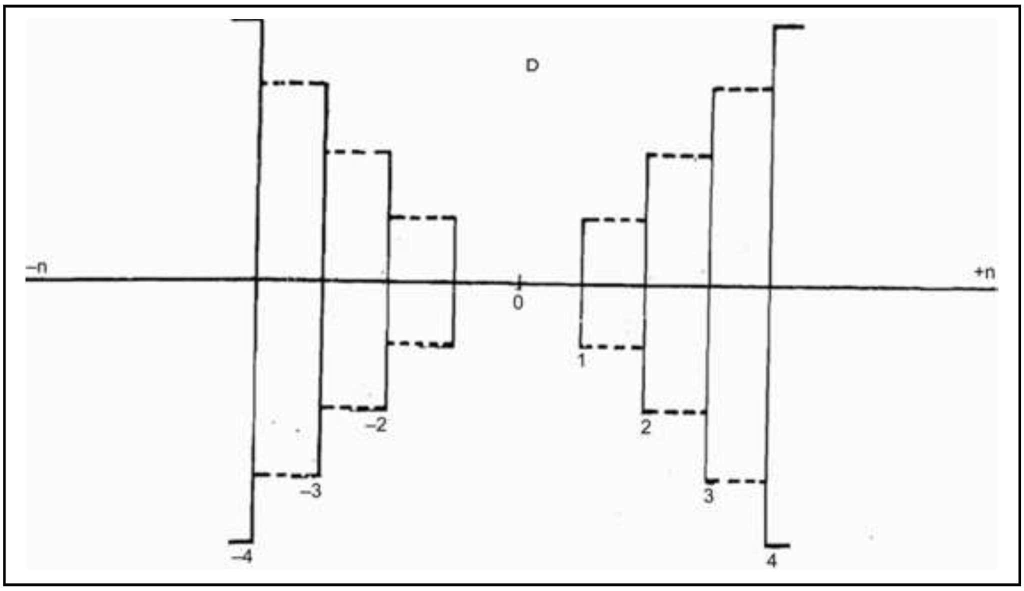
We discuss yet another concept of the so called compactness of a set. The concept of compactness is formulated in terms of the notion of an open cover of a set.

Definition: Open Cover of a Set

Let S be a set and $\{G_\alpha\}$ be a collection of some open subsets of R such that $S \subset \cup G_\alpha$. Then {G} is called an open cover of S.

Example: Verify that the collection $G_n = \{G\}^{\infty}_{n=\infty}$, where $G_n =]-n, n[$ is an open cover of the set R.

Solution:



As shown in the Figure above, we see that every real number belongs to some G_n.

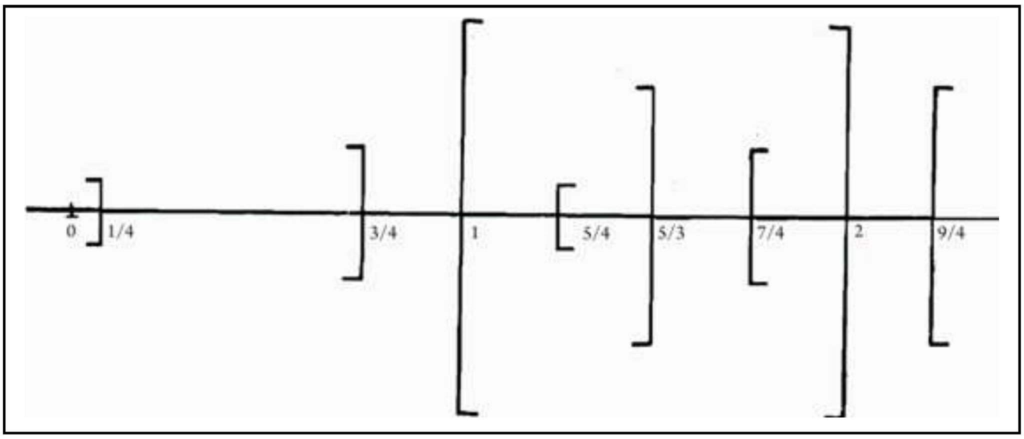
Hence,

$$R = \overset{\infty}{\underset{n=1}{\cup}} G_n$$

Example: Examine whether or not the following collections are open covers of the interval [1, 2].

Notes

Solution: (i) Plot the subsets of G, on the real line as shown in the Figure.



From above figure, it follows that every element of the set S = [1, 2] – {x : 1 ≤ x ≤ 2) belongs to at least one of the subsets of G. Since each of the subsets in G, is an open set, therefore G, is an open cover of S.

(ii) Again plot the subsets of G_j on the real line as done in the case of (i).

You will find that none of the points in the interval $[\frac{5}{4}, \frac{3}{2},]$ belongs to any of the subsets of G_2.

Therefore G_2 is not an open cover of S.

Now consider the set [0, 1] and two classes of open covers of this set namely G, and G given as

$$G_1 = \{] - \frac{1}{n}, 1 + \frac{1}{n} [\}_{n=1}^{\infty}, G_2 = \{] - 1 - \frac{1}{2n}, 1 + \frac{1}{2n} [\}_{n=1}^{\infty}.$$

You can see that $G_2 \subset G$,. In this case, we say that G, is a subcover of G. In general, we have the following definition.

Definition: Subcover and finite subcover of a set

Let G be an open cover of a set S. A subcollection E of G is called a subcover of S if E too is a cover of S. Further, if there are only a finite number of sets in E, then we say that E is a finite subcover of the open cover G of S. Thus, if G is an open cover of a set S, then a collection E is a finite subcover of the open cover G of S provided the following three conditions hold.

(i) E is contained in G.

(ii) E is a finite collection.

(iii) E is itself a cover of S.

From the forgoing example and exercise, it follows that an open cover of a set may or may not admit of a finite subcover. Also, there may be a set whose every open cover contains a finite subcover. Such a set is called a compact set. We define a compact set in the following way.

Definition: Compact set

A set is said to be compact if every open cover of it admits of a finite subcover of the set.

For example, consoder the finite set S = {1, 2, 3} and an open cover $\{G_\alpha\}$ of S. Let G^1, G^2, G^3, be the sets in G which contain 1, 2, 3 respectively. Then $\{G^1, G^2, G^3\}$ is a finite subcover of S. Thus S is a compact set. In fact, you can show that every finite set in R is a compact set.

Notes

The collection G = {] – n, n[: n ∈N} is an open cover of R but does not admit of a finite subcover of R. Therefore the set R is not a compact set.

Thus you have seen that every finite set is always compact. But an infinite set may or may not be a compact set. The question, therefore, arises, "What is the criteria to determine when a given set is compact?" This question has been settled by a beautiful theorem known as Heine-Borel Theorem named in the honour of the German Mathematician H.E. Heine [1821-1881] and the French Mathematician F.E.E. Borel [1871-1956], both of whom were pioneers in the development of Mathematical Analysis.

We state this theorem without proof.

Theorem: Heine-Borel Theorem

Every closed and bounded subset of R is compact.

The immediate consequence of this theorem is that every bounded and closed interval is compact.

Self Assessment

Fill in the blanks:

1. A number p is said to be a .... of real numbers if every neighbourhood of p contains at least one point of the set S different from p.
2. Let S be an infinite and bounded subset of R. Since A is bounded, therefore A has both a lower bound as well as an ....
3. A set is said to be closed if it contains all its ....
4. A set is closed if and only if its .... is open.
5. Let S be a set and $\{G_\alpha\}$ be a collection of some open subsets of R such that .... Then {G} is called an open cover of S.
6. A set is said to be compact if every open cover of it admits of a .... of the set.

3.8 Summary

- We have defined the absolute value or the modulus of a real number and discussed certain related properties. The modulus of real number x is defined as

  $|x| = x \quad \text{if} \quad x \geq 0$

  $\quad = -x \quad \text{if} \quad x < 0.$

  Also, we have shown that

  $|x - a| < d \Leftrightarrow a - d < x < a + d$

- We have discussed the fundamental notion of NBD of a point on the real line i.e. first we have defined it as a δ – neighbourhood and then, in general, as a set containing, an open interval with the point in it.
- With the help of NBD of a point we have defined, an open set in the sense that a set is open if it is a NBD of each of its points.
- We have introduced the notion of the limit point of a set. A point p is said to be a limit point of a set S if every NBD of p contains a point of S different from p. This is equivalent to saying that a point p is a limit point of S if every NBD of p contains an infinite number of the members of S. Also, we have discussed Bulzano-Weiresstrass theorem which gives

a sufficient condition for a set to possess a limit point. It states that an infinite and bounded set must have a limit point. This condition is not necessary in the sense that an unbounded set may have a limit point.

- The limit points of a set may or may not belong to the set. However, if a set is such that every limit point of the set belongs to it, then the set is said to be a closed set. The concept of a closed set has been discussed. Here, we have also shown a relationship between a closed set and an open set in the sense that a set is closed if and only if its complement is open. Further, we have also defined the Derived set of a set S as the set which consists of all the limit points of the set S. The Union of a given set and its Derived set is called the closure of the set. Note the distinction between a closed set and the closure of a set S.
- Finally, we have introduced another topological notion. It is about the open cover of a given set. Given a set S, a collection of open sets such that their Union contains the set S is said to an open cover of S. A set S is said to be compact if every open cover of S admits of a finite subcover. The criteria to determine whether a given set is compact or not, is given by a theorem named Heine-Borel Theorem which states that every closed and bounded subset of R is compact. An immediate consequence of this theorem is that every bounded and closed interval is compact.

3.9 Keywords

Bulzano Weierstrass Theorem: Every infinite bounded subset of set R has a limit point (in K).

Compact Set: A set is said to be compact if every open cover of it admits of a finite subcover of the set.

Heine-Borel Theorem: Every closed and bounded subset of R is compact.

3.10 Review Questions

1. Prove that $-|x| = \text{Min.}\ \{x, -x\}$ for any $x \in R$. Deduce that $-|x| \leq x$, for every $|x| \in R$. Illustrate it with an example.
2. Prove that $|x|^2 = x^2$, for my $x \in R$.
3. For any two real numbers x and y ($y \neq \theta$), prove that

$$\left|\frac{x}{y}\right| = \frac{|x|}{|y|}.$$

4. Prove that $|x - y| \geq ||x| - |y||$ for any real numbers x and y.
5. Test which of the following are open sets:
   (i) An interval [a, b] far $a \in R$, $b \in R$, $a < b$
   (ii) The intervals [0, l [; and] 0, 1[
   (iii) The set Q of rational numbers
   (iv) The set N of natural numbers and the set Z of integers.
   (v) The set $\left\{\frac{1}{n} : n \in N\right\}$
   (vi) The intervals]a, ∞[and [a, ∞[for $a \in R$.

Notes

Answers: Self Assessment

1. limit point of a set S
2. upper bound
3. limit points
4. complement
5. $S \subset \cup G_\alpha$
6. finite subcover

3.11 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 4: Compactness

Notes

CONTENTS

Objectives

Introduction

4.1 Compactness

4.2 Compactness of Subsets

4.3 Intersections of Closed Sets

4.4 Compactness of Products

4.5 Compactness and Continuity

4.6 Compact Sets in $\mathbb{R}^n$

4.7 Sequential Compactness

4.8 Summary

4.9 Keywords

4.10 Review Questions

4.11 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss the compactness of a set
- Explain intersection of closed set
- Discuss compactness and continuity
- Describe sequential compactness

Introduction

In last unit you have studied about matric spaces. You all go through concept of open sets, limit points of sets in last unit. This unit will provide you explanations of compactness of a set.

4.1 Compactness

Definition:

- A cover of A is a collection U of sets whose union contains A.
- A subcover is a subcollection of U which still covers A.
- A subcover is open if its members are all open.

Definition: Topological space T is compact if every open cover has finite subcover.

Theorem: (Heine-Borel). Any closed bounded interval $[a, b] \subset \mathbb{R}$ is compact.

Proof: Let U be open cover of [a, b]. Let

$$A = \{x \in [a, b] : [a, x] \text{ covered by finite subfamily of } U\}$$

Notes

Then a $\in$A so A $\neq \theta$, bounded above by b. Let c = sup A. a $\leq$ c $\leq$ b so c $\in U$ for some $U \in \mathcal{U}$. U open so $\exists\ \delta > 0$ s.t. (c – δ, c + δ)$\cdot \subset U$.

c = sup A so $\exists$ x $\in$ A s.t. x > c – δ. [a, c + δ) $\subseteq$ [a, x] $\cup$ (c – δ, c + δ) can be covered by finite subfamily of U so (c, c + δ) $\cap$ [a, b] = θ (since any point in here is in A but > c sup A). So c = b.

4.2 Compactness of Subsets

Proposition: Any closed subset C of compact space compact.

Proof: Let U be cover of C by sets open in T. Adding open T\C get open cover of T. Finite subcover of this cover contains finite subcover of C of sets from U.

Proposition: Compact subspace C of Hausdorff T is closed in T.

Proof: a $\in$T\C. $\forall$ x $\in$ C $\exists$ disjoint $U_x \ni x$, $V_x \ni a$ open in T since T Hausdorff. U_x open cover of C so has finite subcover $U_{x1},\ldots, U_{xn}$. Then $V = \bigcap_{i=1}^{n} V_{xi}$ open, a $\in$ V and disjoint from C. Hence a $\in$ (T\C)° and T\C open.

Proposition: Compact subspace C of metric space M is bounded.

Proof: Let a $\in$ M. Balls B(a, r) (r > 0) are open and cover C, so $\exists\ r_1,\ldots, r_n$ s.t. $C \subset \bigcap_{i=1}^{n} B(a, r_i) =$ B (a, max $\{r_1, \ldots, r_n\}$).

4.3 Intersections of Closed Sets

Theorem: Let F be collection of non-empty closed subsets of compact T s.t. every finite subcollection of F has non-empty intersection. Then intersection of all sets from T non-empty.

Proof: Assume intersection of all sets empty. Let U be collection of complements. U covers T by DeMorgan. U open cover so exists finite subcover $U_1,\ldots, U_n$. Then $F_i := T\backslash U_i \in F$ and empty intersection by DeMorgan. This contradicts the assumption of the theorem.

Corollary: Let $F_1 \supset F_2 \supset \ldots$ sequence of non-empty closed subsets of compact T. Then $\bigcap_{k=1}^{\infty} F_k \neq \theta$.

Corollary: Let $F_1 \supset F_2 \supset \ldots$ sequence of non-empty compact subsets of Hausdorff T. Then $\bigcap_{k=1}^{\infty} F_k \neq \theta$.

Proof: By proposition 4.4 compact subsets of Hausdorff space are closed.

4.4 Compactness of Products

Lemma: T, S compact, U open cover of T × S. If s $\in$ S there exists open V $\subset$ S, s $\in$ V s.t. T × V can be covered by finite subfamily of U.

Proof: $\forall$ x $\in$ T find $W_x \in U$ s.t. (x, s) $\in W_x$. Exists open $U_x \subset T$, $V_x \subset S$ s.t. (x, s) $\in U_x \times V_x \subset W_x$. $\{U_x : x \in T\}$ open cover of T so $\exists\ U_{x1},\ldots, U_{xn}$ which cover T. Let $V = \bigcap_{i=1}^{n} V_{xi}$. V $\subset$ S open and

$$T \times V \subset \bigcup_{i=1}^{n} U_{xi} \times V_{xi} \bigcup_{i=1}^{n} W_{xi}$$

Theorem: (Tychonov). S,T compact $\Rightarrow$ T × S compact.

Proof: By lemma 3.8 $\forall$ y $\in$ S $\exists$ $V_y \subset S$ open s.t. T × V_y can be covered by finite subfamily of U. S compact, $\{V_y : y \in S\}$ form open cover so $\exists\ V_{yi},\ldots, V_{ym}$ which cover S.

Notes

$T \times S = \bigcup_{j=1}^{m} T \times V_{yi}$. Finite union, each $T \times V_{yj}$ can be covered by finite subfamily of *U*, so T × S can be covered by finite subfamily of *U*.

4.5 Compactness and Continuity

Proposition: Cts image of compact space compact.

Proof: f: T → S cts, T compact. *U* open cover of *f*(*T*). $f^{-1}(U)$ open $\forall\, U \in U$.

Cover T since $\forall\, x \in T\, f(x)$ in some U ∈ U. Hence $\exists\, f^{-1}(U_1),\dots, f^{-1}(U_n)$ subcover of T. $\forall\, y \in f(T)$ have y = f(x) where $x \in T$ so $x \in f^{-1}(U_i)$ for some *i* so $y \in U_i$. Hence $U_1,\dots,U_n$.

Theorem: Cts bijection of compact T onto Hausdorff S is homeomorphism.

Proof: U open in T, T\U closed so compact.

Therefore $(f^{-1})^{-1}(U) = f(U) = S\backslash f(T\backslash U)$ open, so f^{-1} cts.

Corollary: Let T be compact. Cts $f : T \to \mathbb{R}$ is bdd and attains max and min.

Proof: f(T) compact so closed.

Then $\sup f(T) \in \overline{f(T)} = f(T)$.

Alternative proof: Let $c = \sup_{x \in T} f(x)$. If f not attain c then $f(x) < c\ \forall\, x$ so $\{x : f(x) < r\} = f^{-1}(-\infty, a)$ where r < c s.t. $T \subset \bigcup_{i=1}^{n}\{x : f(x) < r_i\}$. Then $f(x) < \max\{r_1,\dots, r_n\}\ \forall\, x$ so $c = \sup_{x\in T} f(x) \le \max\{r_1,\dots, r_n\} < c$ Contradiction.

Definition: Given cover *U* of metric M, δ > 0 called Lebesgue number of *U* if $\forall\, x \in M\ \exists\, U \in U$ s.t. $B(x, \delta) \subset U$.

Proposition: Every open cover U of compact metric space has a Lebesgue number.

Proof: $\forall\, x \in M$ pick r(x) > 0 s.t. B(x, r(x)) contained in some set of *U*. Then $M \bigcup_{x\in M} B\left(x, \frac{r(x)}{2}\right)$ so $\exists\, x_1,\dots, x_j$ s.t. $M \subset \bigcup_{i=1}^{j} B\left(x_i, \frac{r(x_i)}{2}\right)$. Let $\delta = \dfrac{\min\{r(x_1),\dots,r(x_j)\}}{2}$. Then $\forall\, x \in M$ pick *i* s.t. $x \in B\left(x_i, \frac{r(x_i)}{2}\right)$ and $B(x, \delta) \subset B(x_i, r(x_i))$ subset of some set from *U*.

Theorem: Cts map of compact metric M to metric N is uniformly cts.

Proof: Let ∈ > 0. Then sets $U_z = f^{-1}(B_N(f(z), \frac{\varepsilon}{2}))$ z ∈ M open cover of M. Let be Lebesgue number. If x, y ∈ M, $d_M(x, y) < \delta \Rightarrow y \in B(x, \delta) \subset U_z$ some z so $d_N(f(x), f(y)) \le d_N(f(x), z) + d_N(f(y), z) < \varepsilon$.

4.6 Compact Sets in $\mathbb{R}^n$

Theorem: (Heine-Borel). $A \subset \mathbb{R}^n$ compact if f closed and bdd.

Proof: (⇒) Metric spaces are Hausdorff, so A closed.

(⇐) $C \subset \mathbb{R}^n$ bdd $\Rightarrow \exists\, [a, b] \subset \mathbb{R}^n$ s.t. $C \subset [a, b] \times \dots \times [a, b]$. This compact by Tychanov. If C closed then closed subset of compact space so compact.

Notes

4.7 Sequential Compactness

Theorem: Metric M is compact if f every sequence in M has convergent subsequence.

Lemma: A_k sequence of subsets of metric M. Then $\forall x \in \cap_{j=1}^{\infty}\overline{A_j}$ $\exists x_k \in A_k$ s.t. $x_k \to x$.

Proof: Take $x_k \in A_k \cap B\left(x, \frac{1}{k}\right) \neq \theta$.

Corollary: $x_k \in M$ and $\cap_{j=1}^{\infty}\overline{\{x_j, x_{j+1}...\}}. \neq \theta$ then x_k have convergent

Proof: Let $x \in \cap_{j=1}^{\infty}\overline{\{x_j, x_{j+1}...\}}$. As $\exists k_j \geq j$ s.t. $x_{kj} \to x$. $k_j \to \infty$ so can choose subsequence k_{ji} s.t. $k_{ji+1} > k_{ji}$ (as k_js not necessarily in order). Then x_{kji} subsequence converging to x.

Proof of ($\Rightarrow$) of theorem 3.16. Let $x_k \in M$, $F_j = \overline{\{x_j, x_{j+1}...\}}$. F_j form decreasing sequence of non-empty closed subsets of M.

By corollary 3.6 $\cap_{j=1}^{\infty}F_j \neq \theta$ so x_k have convergent subsequence by corollary 3.18.

Notation

U open cover of M. $\forall x \in M$

$$r(x) = \sup \{r \leq 1 : \exists U \in \mathcal{U} \text{ s.t. } B(x, r) \subset U\}$$

Lemma: If $y_k \to x$ $\exists$ K s.t. $y_{k+1} \in B\left(y_k, \frac{r(x)}{2}\right)$ for $k \geq K$.

Proof: Let $U \in \mathcal{U}$ be s.t. $B\left(x, \frac{r(x)}{2}\right) \subset U$. Take K s.t. $d(y_k, x) < \frac{r(x)}{16}$ for $k \geq K$. Then $k \geq K \Rightarrow$ $B\left(y_k, \frac{r(x)}{2} - d(x, y_k)\right) \subset B\left(x, \frac{r(x)}{2}\right) \subset \mathcal{U}$, so $r(y_k) \geq \frac{r(x)}{2} - d(x, y_k) \geq \frac{r(x)}{4}$, so

$$d(y_{k+1}, y_k) \leq d(y_{k+1}, x) + d(y_k, x) < \frac{r(x)}{8} \leq \frac{r(y_k)}{2}$$

$M_1 := M$, $s_1 := \sup_{x\in M1} r(x)$. Find $x_1 \in M_1$ s.t. $r(x_1) > \frac{s_1}{2}$, choose $\mathcal{U}_1 \in \mathcal{U}$ s.t. $B\left(x_1, \frac{r(x_1)}{2}\right) \subset \mathcal{U}_1$.

If $x_1, ..., x_j$ have been defined,

$$M_{j+1} := M \setminus B\left(x_j, \frac{r(x_j)}{2}\right) = M \setminus \bigcup_{i=1}^{j} B\left(x_j, \frac{r(x_j)}{2}\right)$$

If $M_{j+1} = \theta$ then $M \subset \cup_{i=1}^{j} B\left(x_i, \frac{r(x_i)}{2}\right) \subset \cup_{i=1}^{j} U_i$ has finite subcover.

If $M_{j+1} \neq \theta$ let $s_{j+1} = \sup_{x\in MJ+1}\{r(x)\}$, find x_{j+1} s.t. $r(x_{j+1}) > \frac{s_{j+1}}{2}$, choose $\mathcal{U}_{j+1} \in \mathcal{U}$ s.t. $B\left(x_{j+1}, \frac{r(x_{j+1})}{2}\right) \subset U_{j+1}$.

If procedure stops we have finite subcover. If it runs forever we have infinite sequence x_j s.t. $x_i \notin$

Notes

$B\left(x_j, \frac{r(x_j)}{2}\right)$ for i > j. This has convergent subsequence x_{jk} by assumption, so ∃ k s.t. $B\left(x_{jk}, \frac{r(x_{jk})}{2}\right)$.

This is a contradiction, so the procedure stops.

Self Assessment

Fill in the blanks:

1. T, S compact, U open cover of .... If s ∈ S there exists open V ⊂ S, s ∈ V s.t. T × V can be covered by finite subfamily of *U*.
2. Let T be compact. Cts .... is bdd and attains max and min.
3. Given cover *U* of metric M, δ > 0 called .... of *U* if ∀ x ∈ M ∃ *U* ∈ *U* s.t. B(x, δ) ⊂ U.

4.8 Summary

- Let $F_1 \supset F_2 \supset \ldots$ sequence of non-empty closed subsets of compact T. Then $\bigcap_{k=1}^{\infty} F_k \neq \theta$. Let $F_1 \supset F_2 \supset \ldots$ sequence of non-empty compact subsets of Hausdorff T. Then $\bigcap_{k=1}^{\infty} F_k \neq \theta$.
- ∀ x ∈ T find $W_x \in U$ s.t. (x, s) ∈ W_x. Exists open $U_x \subset T$, $V_x \subset S$ s.t. (x, s) ∈ $U_x \times V_x \subset W_x$. $\{U_x : x \in T\}$ open cover of T so ∃ $U_{x1}, \ldots, U_{xn}$ which cover T. Let $V = \bigcap_{i=1}^{n} V_{xi}$. V ⊂ S open and

  $T \times V \subset \bigcup_{i=1}^{n} U_{xi} \times V_{xi} \bigcup_{i=1}^{n} W_{xi}$
- Let $x \in \bigcap_{j=1}^{\infty} \overline{\{x_j, x_{j+1} \ldots\}}$. ∃ $k_j \geq j$ s.t. $x_{kj} \to x$. $k_j \to \infty$ so can choose subsequence k_{ji} s.t. $k_{ji+1} > k_{ji}$ (as k_js not necessarily in order). Then x_{kji} subsequence converging to x.
- Let $x_k \in M$, $F_j = \overline{\{x_j, x_{j+1} \ldots\}}$. F_j form decreasing sequence of non-empty closed subsets of M. $\bigcap_{j=1}^{\infty} F_j \neq \theta$ so x_k have convergent subsequence.

4.9 Keywords

Space Compact: Cts image of compact space compact.

Homeomorphism: Cts bijection of compact T onto Hausdorff S is homeomorphism.

Lebesgue Number: Every open cover U of compact metric space has a Lebesgue number.

Convergent Subsequence: Metric M is compact iff every sequence in M has convergent subsequence.

4.10 Review Questions

1. Discuss the compactness of a set.
2. Explain intersection of closed set.
3. Discuss compactness and Continuity.
4. Describe sequential compactness.

Notes

Answers: Self Assessment

1. $T \times S$
2. $f : T \to \mathbb{R}$
3. Lebesgue number

4.11 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 5: Connectedness

Notes

CONTENTS

Objectives
Introduction
5.1 Connected, Separated
5.2 Connectedness in Metric Spaces
5.3 Connected Spaces from Others
5.4 Connected Components
5.5 Path Connectedness
5.6 Open Sets in $\mathbb{R}^n$
5.7 Summary
5.8 Keywords
5.9 Review Questions
5.10 Further Readings

Objectives

After studying this unit, you will be able to:

- Define Connectedness
- Discuss the Connectedness in metric spaces
- Explain connected spaces from others
- Describe connected components and Path connected

Introduction

In last unit you have studied about the compactness of the set. As you all come to know about the compactness and continuity. After understanding the concept of compactness let us go through the explanation of connectedness.

5.1 Connected, Separated

Definition: Topological T connected if for every decomposition $T = A \cup B$ into disjoint open A, B either A or B is empty.

Definition: $T \subset S$ separated by sets U, $V \subset S$ if $T \subset U \cup V$, $U \cap V \cap T = \theta$, $U \cap T \neq \theta$, $V \cap T \neq \theta$.

Proposition: $T \subset C$ S disconnected if T is separated by some U, $V \subset S$.

Proof: ($\Rightarrow$) If disconnected $\exists$ A, $B \subset T$, A, $B \neq \theta$ s.t. $T = A \cup B$ and $A \cap B = \theta$. $T \subset S$ so $\exists$U, V open in S s.t. $A = U \cap T$, $B = V \cap T$. Then U, V separate T.

($\Leftarrow$) If U, V separate T let $A = U \cap T$, $B = V \cap T$ then T not connected.

Proposition: TFAE:

1. T disconnected

Notes

2. T has subset which is open, closed, different from θ, T

3. T admits non-constant cts function to two point discrete space.

Proof: (1. $\Rightarrow$ 2.) $\exists$ decomposition T = A $\cup$ B with A, B open, non-empty. Hence A = T\B is open and closed, different from θ, = T.

(2. $\Rightarrow$ 3.) θ, T $\neq$ A $\subset$ T open, closed. Define f : T $\rightarrow$ {0,1} by f(x) = $\begin{cases} 0 & x \in A \\ 1 & x \notin A \end{cases}$

This cts as pre-images open

(3. $\Rightarrow$ 1.) f: T $\rightarrow$ {0, 1} non-constant and cts. Define A = $F^{-1}(0)$, B = f^{-1}

5.2 Connectedness in Metric Spaces

Theorem: T $\subset$ M (M metric) disconnected iff $\exists$ disjoint open U,V $\subset$ M s.t. T $\cap$ U $\neq \theta \neq$ T $\rightarrow$ V and T $\subset$ U $\cup$ V.

Proof: ($\Leftarrow$) Clear

($\Rightarrow$) T = A U B. Let

$$U = \{x \in M : d(x, A) < d(x, B)\}$$

$$V = \{x \in M : d(x, A) > d(x, B)\}$$

U, V disjoint, open.

Going to prove A $\subset$ U: Let x $\in$ A. A open in T so $\exists$ $\delta > 0$ s.t. B(x, δ) $\cap$ T $\subset$ A. B $\subset$ T disjoint from A so B(x, δ) $\cap$ B = θ, so d(x, B) $\geq \delta > 0$. Since d(x, A) = 0 we have x $\in$ U. Similarly B $\subset$ V.

Lemma: I $\subset \mathbb{R}$ is an interval iff $\forall$ x, y$\in$ I, $\forall$ z $\in \mathbb{R}$,

$$x < z < y \Rightarrow z \in I$$

Proof: Intervals clearly have this property. Conversely suppose I has above property, non-empty, not single point. Let a = inf I, b = sup I.

Show (a, b) $\subset$ I: If z $\in$ (a, b) $\exists$ x, y $\in$ with x < z < y so z $\in$ I. Hence (a, b) $\subset$ I $\subset$ (a, b) $\cup$ {a, b}.

Theorem: T $\subset \mathbb{R}$ connected iff it is an interval.

Proof: ($\Rightarrow$) Suppose I not interval. Then by lemma 4.4 $\exists$ x, y $\in$ I, z $\in \mathbb{R}$ s.t. x < z < y and z $\notin$ I. Let A = $(-\infty, z) \cap$ I, B = $(z, \infty) \cap$ I. A, B disjoint, non-empty, open and I = A $\cup$ B.

($\Leftarrow$) Suppose I not connected. Then 3 cts non-constant f : I $\rightarrow$ {0, 1} where {0, 1} has discrete contradicting IVT.

($\Leftarrow$) I partitioned into non-empty A, B open. Choose a $\in$ A, b $\in$ B, a < b. A, B open cover of [a, b]. Let δ be its Lebesgue number. Then $\left[a, a+\frac{\delta}{2}\right] \subset$ A, $\left[a+\frac{\delta}{2}, a+\frac{2\delta}{2}\right] \subset$ A, ... until we get to an interval containing b. So b $\in$ A and A, B not disjoint.

5.3 Connected Spaces from Others

Proposition: Cts image of connected space connected.

Proof: Suppose f : T $\rightarrow$ S cts, T connected. If f(T) disconnected $\exists$ U, V $\subset$ S open separating f(T). Then $f^{-1}(U)$, $f^{-1}(V)$ open, disjoint, cover T. Contradiction as T connected.

Notes

Proposition: If C, C_j ($j \in J$) connected subspaces of topological T and if $c_j \cap \bar{C} \neq \theta \;\forall\; j \in J$ then

$$K = C \cup \bigcup_{j \in J} C_j$$

is connected.

Proof: Suppose K disconnected. Hence $\exists$ U, V $\subset$ T open that separate K.

C connected so cannot be separated by U,V, so does not meet one of them. Suppose w.l.o.g $C \cap V = \theta$. Then $C \subset U$. Since V open $\bar{C} \cap V = \theta$, so $K \cap \bar{C} \subset U$. Then $C_j \cap U \neq \theta \;\forall j$.

C_j connected so $C_j \subset U$ or $C_j \subset V$. $C_j \cap U \neq \theta$ so $C_j \subset U$.

Then $K \subset U$ contradicting $V \cap K \neq \theta$.

Corollary: $C \subset T$ connected and $C \subset K \subset \bar{C}$. Then K connected.

Proof: $K = C \bigcup_{x \in K} \{x\}$ and $\{x\} \cap \bar{C} \neq \theta \;\forall x$.

Proposition: Product of connected spaces is connected.

Proof: Let T, S connected, so $\in$ S. Define $C = T \times \{s_0\}$ and $C_t = \{t\} \times S$ (for some $t \in T$). Then C, C_t homeomorphic to T and S are connected. $C_t \cap C \neq \theta$ and $T \times S = C \cup \bigcup_{t \in T} C_t$ connected.

Example: $\text{Sin}\left(\frac{1}{t}\right) \cup \underbrace{\{(0, t) \in \mathbb{R}^2 : \in (-1, 1)\}}_{I}$ is connected.

Proof:

$$C = \left\{\left(t, \sin\left(\frac{1}{t}\right)\right) : t > 0\right\}$$

$$D = \left\{\left(t, \sin\left(\frac{1}{t}\right)\right) : t < 0\right\}$$

C, D, I cts images of intervals so connected.

$(0, 0) \in I$ is in $\bar{C}$ as $(t_k) \sin\left(\frac{1}{t_k}\right) \to (0, 0)$ when $t_k = \frac{1}{k\pi}$. Then $I \cup C$ connected. Similarly $I \cup D$.

5.4 Connected Components

Definition: x ~ y if x, y belong to a common connected subspace of T. Equivalence classes are connected components of T.

Are maximal connected subsets of T. Number of connected components is topological invariant.

Property T\{x} connected $\forall x \in T$ topological invariant.

5.5 Path Connectedness

Definition: a, b $\in$ T. $\varphi : [0, 1] \to T$ cts with $\varphi(0) = a$, $\varphi(1) = b$ called a path from a to b.

Definition: T path connected if any two points can be joined by a path.

Notes

Proposition: Path connected $\Rightarrow$ connected.

Proof: a $\in$ T. $\forall$ x $\in$ T image C_x of path a to x is connected, and all C_x contain a. Then T = $\bigcup_{x \in T} C_x$ connected by 4.7.

5.6 Open Sets in $\mathbb{R}^n$

Theorem: Any U C $\mathbb{R}^n$ open, connected is path connected.

Proof: Let a $\in$ U, V = {x $\in$ U : $\exists$ path from a to x}.

Let z $\in$ U $\cap$ $\bar{V}$. Find $\delta > 0$ s.t. B(z, δ) $\subset$ U. z $\bar{V}$ so $\exists$y $\in$ V $\cap$ B(x, δ).

Then B(z, δ) $\subset$ V since join path from a to y to path from y to z.

Theorem: All components of open U $\subset$ $\mathbb{R}^n$ open.

Proof: C component of U, x $\in$ C. Find $\delta > 0$ with B(x, δ) $\subset$ U. B(x, δ) connected and C union of all connected subsets of U containing x so B(x, δ) $\subset$ C, so C open.

Theorem: U $\subset$ $\mathbb{R}$ open iff disjoint union of countably many open intervals.

Proof: ($\Leftarrow$) Any union of open sets open.

($\Rightarrow$) U $\subset$ $\mathbb{R}$ open, C_j (j $\in$ J) its components. C_j connected and open so are open intervals. For each j $\exists$ rational $r_j \in C_j$. C^S_j mutually disjoint so j $\to$ r_j injection into $\mathbb{Q}$, so can order J into a sequence.

Self Assessment

Fill in the blanks:

1. Topological T connected if for every decomposition .... into disjoint open A, B either A or B is empty.
2. T $\subset$ M (M metric) disconnected iff $\exists$ .... U,V $\subset$ M s.t. T $\cap$ U $\neq \theta \neq$ T $\to$ V and T $\subset$ U $\cup$ V.
3. .... is an interval iff $\forall$ x, y$\in$ I, $\forall$ z $\in$ $\mathbb{R}$, x < z < y $\Rightarrow$ z $\in$ I
4. Suppose f : T $\to$ S cts, T connected. If f(T) disconnected $\exists$ U, V $\subset$ S open separating f(T). Then f^{-1}(U), f^{-1}(V) open, disjoint, cover T. Contradiction as ....
5. C $\subset$ T connected and C $\subset$ K $\subset$ $\bar{C}$. Then ....
6. x ~ y if x, y belong to a common connected subspace of T. .... are connected components of T.

5.7 Summary

- Topological T connected if for every decomposition T = A$\cup$B into disjoint open A, B either A or B is empty.
- T $\subset$ S separated by sets U, V $\subset$ S if T $\subset$ U$\cup$V, U $\cap$ V $\cap$ T = θ, U $\cap$ T $\neq \theta$, V $\cap$ T $\neq \theta$.
- T $\subset$ C S disconnected iff T is separated by some U, V $\subset$ S.
- Proof. ($\Rightarrow$) If disconnected $\exists$ A, B $\subset$ T, A, B $\neq \theta$ s.t. T = A $\cup$ B and A $\cap$ B = θ. T $\subset$ S so $\exists$U, V open in S s.t. A = U $\cap$ T, B = V $\cap$ T. Then U, V separate T.
- Suppose K disconnected. Hence $\exists$ U, V $\subset$ T open that separate K.

- C connected so cannot be separated by U,V, so does not meet one of them.
- C component of U, $x \in C$. Find $\delta > 0$ with $B(x, \delta) \subset U$. $B(x, \delta)$ connected and C union of all connected subsets of U containing x so $B(x, \delta) \subset C$, so C open.

5.8 Keywords

Path Connected: T path connected if any two points can be joined by a path.

Topological Invariant: Number of connected components is topological invariant.

5.9 Review Questions

1. Define Connectedness.
2. Discuss the Connectedness in metric spaces.
3. Explain connected spaces from others.
4. Describe connected components and Path connected.

Answers: Self Assessment

1.	$T = A \cup B$	2.	disjoint open
3.	$I \subset \mathbb{R}$	4.	T connected
5.	K connected	6.	Equivalence classes

5.10 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Notes

Unit 6: Completeness

CONTENTS

Objectives

Introduction

6.1 Completeness

6.2 Proving Cauchy

6.3 Completion

6.4 Contraction Mapping Theorem

6.5 Total Boundedness

6.6 Summary

6.7 Keywords

6.8 Review Questions

6.9 Further Readings

Objectives

After studying this unit, you will be able to:

- Define Completeness
- Discuss the Cauchy
- Explain contraction mapping theorem
- Describe total boundness

Introduction

In earlier unit you have studied about the compactness and connectedness of the set. As you all come to know about the connected components and Path connectedness. After understanding the concept of compactness and connectedness let us go through the explanation of completeness.

6.1 Completeness

This is a concept that makes sense in metric spaces only.

Definition: Metric M is complete if every Cauchy sequence in M converges (to a point of M).

Remark: This is not a topological invariant: (0, 1) – incomplete and $\mathbb{R}$ complete are homeomorphic.

Proposition: Cvgt $\Rightarrow$ Cauchy.

Proof: $\forall \varepsilon > 0 \ \exists N$ s.t. $d(x_n, x) < \frac{\varepsilon}{2}$ for $n \geq N$. If $m, n \geq N$ then

$$d(x_m, x_n) < d(x_m, x) + d(x_n, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Proposition:

1. Complete subspace S of metric M is closed.
2. Closed subset S of complete M is complete.

Proof:

1. Let $x_n \in S$, $x_n \to x \in M$. (x_n) Cauchy in S so cvgs in S to $y \in S$. $S \leq M$ so $x_n \to y$ in M. By uniqueness of limits $x = y \in S$.
2. Let $(x_n) \subset S$ Cauchy. Cauchy in M so cvgs to point of M which in S as S closed.

Proposition: $\forall$ S B(S) of bdd functions $S \to \mathbb{R}$ with sup norm is complete.

Proof: Let (f_n) Cauchy, $\varepsilon > 0$. 3 N s.t. $||f_m - f_n|| < \varepsilon$ for $n, m \geq N$. Hence for fixed x $(f_n(x))$ Cauchy in $\mathbb{R}$, so cvgs to $f(x) \in \mathbb{R}$.

For $n \geq N \, |f_m(x) - f_n(x)| \leq \varepsilon \quad \forall m \geq N$. Let $m \to \infty$ then

$|f(x) - f_n(x)| \leq \varepsilon \quad \forall x \, \varepsilon \, S, n \geq N$

Then f bdd and $f_n \to f$.

6.2 Proving Cauchy

Proposition: A sequence $(x_n) \subset M$ is Cauchy iff $\exists$ sequence $\varepsilon_n \geq 0$ s.t. $\varepsilon_n \xrightarrow[n\to\infty]{} 0$ and $d(x_m, x_n) \leq \varepsilon_n$ for m > n.

Proof: $(\Rightarrow)$ Suppose (x_n) Cauchy. Then let $\varepsilon_n = \Sigma_{m>n} d(x_m, x_n) \xrightarrow[n\to\infty]{} 0$.

$(\Leftarrow)$ Given $\varepsilon > 0$ find k s.t. $\varepsilon_n < \varepsilon$ for $n \geq k$. Then $d(x_n, x_n) \leq \varepsilon_n < \varepsilon$ for $m > n \geq k$. Exchanging m, n gives $d(x_m, x_n) < \varepsilon \;\; \forall n, m \geq k$.

Proposition: $(x_n) \subset M$ sequence s.t. $\exists \, \tau_n \geq 0$ with $\Sigma_{n=1}^{\infty} \tau_n < \infty$ and $d(x_n, x_{n+1}) \leq \varepsilon_n \;\; \forall n$. Then (x_n) is Cauchy.

Proof: Follows from 6.4 with $\varepsilon_n = \Sigma_{n=1}^{\infty} \tau_n$. Then

$$d(x_m, x_n) \underset{\Delta \text{ineq}}{\leq} \sum_{k=n}^{m-1} d(x_k, x_{k+1}) \leq \sum_{k=n}^{m-1} \tau_n \leq \varepsilon_n$$

Example: If K compact topological space then space C(K) with sup norm is complete.

Proof: Each f bdd, attains max. Suffices to show C(K) closed in B(K).

Suppose $f_n \in C(K)$ cvg to $f \in B(K)$. Then $\forall \, \varepsilon > 0 \; \exists$ N s.t.

$$\sup_{x \in K} |f(x) - f_n(x)| < \varepsilon \quad \forall n \geq N$$

$$\forall \, a \in \mathbb{R} \; \{x : f(x) > a\} = \bigcup_{\varepsilon > 0} \{x : f_N(x) > a + \varepsilon\}$$

RHS are pre-images of open sets so open. Hence LHS is open. Similarly $\{x : f(x) < a\}$ open. $(-\infty, a)$, (a, ∞) from sub-basis for $\mathbb{R}$ so f cts.

Example: C[0, 1] with norm $||f||_1 = f^1{}_0 |f(x)| dx$ is incomplete.

Notes

Proof:

$$f_n(x) = \begin{cases} \min\left\{\sqrt{n}, \dfrac{1}{\sqrt{x}}\right\} & x > 0 \\ \sqrt{n} & x = 0 \end{cases}$$

so $(f_n) \subset C[0, 1]$.

$$\int_0^1 |f_m(x) - f_n(x)|\,dx = \int_0^{\frac{1}{m}} \left(\sqrt{m} - \sqrt{n}\right)dx + \int_{\frac{1}{m}}^{\frac{1}{n}} \left(\frac{1}{\sqrt{x}} - \sqrt{n}\right)dx$$

$$\leq \frac{1}{\sqrt{m}} + \frac{2}{\sqrt{n}}$$

$$\leq \frac{3}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0$$

so (f_n) Cauchy.

Let $f \in C[0, 1]$. Find $k \in \mathbb{N}$ s.t. $|f| \leq \sqrt{k}$. Then for $n > k$

$$\int_0^1 |f_m(x) - f_n(x)|\,dx = \int_0^{\frac{1}{m}} \left(\sqrt{m} - \sqrt{n}\right)dx + \int_{\frac{1}{m}}^{\frac{1}{n}} \left(\frac{1}{\sqrt{x}} - \sqrt{n}\right)dx$$

$$\geq 2\left(\frac{1}{\sqrt{k}} - \frac{1}{\sqrt{n}}\right) - \frac{1}{\sqrt{k}}$$

$$= \frac{1}{\sqrt{k}} - \frac{2}{\sqrt{n}} \xrightarrow[n\to\infty]{} > 0$$

6.3 Completion

Definition: $S \subset M$ is dense in M if $\overline{S} = M$.

Definition: A completion of metric space M is:

- Complete metric space N s.t. M dense subset of N.
- Complete metric space N and isometry $i: M \to A \subseteq N$ s.t $i(M)$ is dense in N.

Theorem: Any metric M can be isometrically embedded into complete metric space.

Proof: Find isometry of M onto subset of B(M), complete. Fix $a \in M$, define $F : M \to B(M)$ by $F(x)(z) = d(z,x) - d(z, a)$. $|F(x)(z)| \leq d(x, a)$ so $F(x) \in B(M)$.

$$|F(x)(z) - F(y)(z)| = |d(z, x) - d(z, y)|$$

$$\leq d(x, y)$$

Equality occurs when $z = y$. Then $||F(x) - F(y)|| = d(x, y)$ so F isometry.

Corollary: Any metric space M has a completion.

Proof: Embet M into complete N. Then $\overline{M}$ (closure taken in N) is complete by 6.2, M dense in $\overline{M}$. Then $\overline{M}$ completion of M.

6.4 Contraction Mapping Theorem

Definition: $f : M \to M$ contraction if $\exists\, k < 1$ s.t.

$$d(f(x), f(y)) \le kd(x, y) \quad \forall\, x, y \in M$$

Theorem: Banach

If f contraction on complete metric M then f has unique fixed point.

Proof: **Uniqueness:** If $f(x) = x$, $f(y) = y$ then

$$d(x, y) = d(f(x), f(y)) \le kd(x, y) \Rightarrow d(x, y) = 0$$

Existence: Pick $x_0 \in M$, $x_{n+1} = f(f_n)$. By repeated application of the contraction property we get that $d(x_j, x_{j+1}) \le k^j d(x_0, x_1)$. $\sum_{j=1}^{\infty} k^j\, d(x_0, x_1) \le \infty$ so (x_n) Cauchy.

M complete so $x_n \to x \in M$, so $f(x_n) \to f(x)$. But also $f(x_n) = x_{n+1} \to x$ so $f(x) = x$.

6.5 Total Boundedness

Definition: Metric M totally bounded if $\forall\, \varepsilon > 0\ \exists$ finite set $F \subset M$ s.t. $M \subset \bigcup_{x \in F} B(x, \varepsilon)$.

Proposition: Subspace M of metric N is totally bounded iff $\forall\, \varepsilon > 0\ \exists$ finite $H \subset N$ s.t. $M \subset \bigcup_{z \in H} B(z, \varepsilon)$.

Proof: ($\Rightarrow$) Obvious.

($\Leftarrow$) Given $\varepsilon > 0$ let $H \subset N$ be finite set s.t. $M \subset \bigcup_{z \in H} B\left(z, \frac{\varepsilon}{2}\right)$. From each non-empty $M \cap B\left(z, \frac{\varepsilon}{2}\right)$ pick one point. Let F be set of these points.

$F \subset M$ finite.

If $y \in M$ then y in one of $B\left(z, \frac{\varepsilon}{2}\right)$ so $M \cap B\left(z, \frac{\varepsilon}{2}\right) \ne \theta$ so $\exists\, x \in M \cap B\left(z, \frac{\varepsilon}{2}\right)$. Hence $y \in B(x, \varepsilon)$ and $M \subset \bigcup_{z \in F} B(z, \varepsilon)$.

Corollary: Subspace of totally bounded metric space is totally bounded.

Theorem: Metric M totally bounded iff every sequence in M has Cauchy subsequence.

Proof: ($\Rightarrow$) Let $x_n \in M$. M covered by finitely many balls radius 1/2 so $\exists\, B_1$ s.t. $N_1 = \{n \in \mathbb{N} : x_n \in B_1\}$ has $|N_1| = \infty$.

Suppose inductively have defined infinite $N_{k-1} \subset \mathbb{N}$. Since M covered by finitely many balls of radius $\frac{1}{2k}$ $\exists$ one ball B_k s.t. $N_k = \{n \in N_{k-1} : x_n \in B_k\}$ is infinite.

Let n(1) be least element of N_1, n(k) least element of N_k s.t. $n(k) > n(k - 1)$.

Then $\left(x_n(k)\right)_{n=1}^{\infty} \subset \left(x_n\right)_{n=1}^{\infty}$ s.t. $\forall\, k\ x_{n(i)} \in B_k$ for $i \ge k$. Hence $d(x_{n(i)}, x_{n(j)}) < \frac{1}{k}\ \forall i, j \ge k$ so $(x_{n(k)})$ Cauchy.

($\Leftarrow$) Suppose M not totally bounded. Then for some $\varepsilon > 0\ \nexists$ finite F with all points of M within ε of it. Choose $x_1 \in M$, inductively x_k s.t. $d(x_k, x_i) \ge \varepsilon\ \forall\, i < k$. x_k exists by assumption M not totally bounded.

This gives sequence $(x_k)_{k=1}^{\infty}$ s.t. $d(x_i, x_j) \ge \varepsilon\ \forall\, i \ne j$. Then no subsequence of (x_k) Cauchy.

Notes

Self Assessment

Fill in the blanks:

1. A completion of metric space M is .... N s.t. M dense subset of N.
2. Any metric M can be .... into complete metric space.
3. Metric M totally .... if $\forall \varepsilon > 0 \ \exists$ finite set $F \subset M$ s.t. $M \subset \bigcup_{x \in F} B(x, \varepsilon)$.
4. Subspace M of .... is totally bounded iff $\forall \varepsilon > 0 \ \exists$ finite $H \subset N$ s.t. $M \subset \bigcup_{z \in H} B(z, \varepsilon)$.

6.6 Summary

- Complete subspace S of metric M is closed.
- Closed subset S of complete M is complete.
- $\forall$ S B(S) of bdd functions $S \to \mathbb{R}$ with sup norm is complete.
- A sequence $(x_n) \subset M$ is Cauchy iff $\exists$ sequence $\varepsilon_n \geq 0$ s.t. $\varepsilon_n \xrightarrow[n\to\infty]{} 0$ and $d(x_m, x_n) \leq \varepsilon_n$ for m > n.
- $(x_n) \subset M$ sequence s.t. $\exists\ \tau_n \geq 0$ with $\sum_{n=1}^{\infty} \tau_n < \infty$ and $d(x_n, x_{n+1}) \leq \varepsilon_n \ \forall$ n. Then (x_n) is Cauchy.

6.7 Keywords

Cauchy Sequence: Metric M is complete if every Cauchy sequence in M converges (to a point of M).

Cauchy: A sequence $(x_n) \subset M$ is Cauchy iff $\exists$ sequence $\varepsilon_n \geq 0$ s.t. $\varepsilon_n \xrightarrow[n\to\infty]{} 0$ and $d(x_m, x_n) \leq \varepsilon_n$ for m > n.

Completion: Any metric space M has a completion.

6.8 Review Questions

1. Define Completeness.
2. Discuss the Cauchy.
3. Explain contraction mapping theorem.
4. Describe total boundness.

Answers: Self Assessment

1. Complete metric space
2. isometrically embedded
3. bounded
4. metric N

Notes

6.9 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Notes

Unit 7: Convergent Sequence

CONTENTS

Objectives

Introduction

7.1 Convergent Sequence

7.2 Properties of Convergent Sequences

7.2.1 Subsequences

7.3 Subsequences and Compact Metric Spaces

7.4 Subsequences Limits

7.5 Cauchy Sequence

7.6 Cauchy Sequence and Closed Sets

7.7 Cauchy Sequences and Convergent Sequences

7.7.1 Complete Spaces

7.8 Increasing/Decreasing Sequences

7.9 Summary

7.10 Keywords

7.11 Review Questions

7.12 Further Readings

Objectives

After studying this unit, you will be able to:

- Define convergent sequence
- Discuss the properties of convergent sequence
- Explain subsequences and compact metric spaces
- Describe subsequence limits
- Explain the Cauchy sequences and convergent sequences

Introduction

In earlier unit you have studied about the compactness and connectedness of the set. After understanding the concept of compactness and connectedness let us go through the explanation of convergent sequence.

7.1 Convergent Sequence

Definition: A sequence $\{p_n\}$ in a metric space (X, d) is said to converge if there is a point $p \in X$ with the following property:

$$(\forall \in > 0)(\exists N)(\forall n > N)\ d(p_n, p) < \in$$

In this case we also say that $\{p_n\}$ converges to p or that p is the limit of $\{p_n\}$ and we write $pn \to p$ or lim $\lim_{n\to\infty} p_n = p$

If $\{p_n\}$ does not converge we say it diverges

If there is any ambiguity we say $\{p_n\}$ converges/diverges in X

The set of all p_n is said to be the range of $\{p_n\}$ (which may be infinite or finite). We say $\{p_n\}$ is bounded if the range is bounded.

Example: Notice that our definition of convergent depends not only on $\{p_n\}$ but also on X.

For example $\{1/1 : n \in \mathbb{N}\}$ converges in $\mathbb{R}^1$ and diverges in $(0, \infty)$. Consider the following sequence of complex number (i.e. $X = \mathbb{R}^2$)

(a) If $S_n = 1/n$ then $\lim_{n\to\infty} S_n = 0$; the range is infinite, and the sequence is bounded.

(b) If $S_n = n^2$ then the sequence $\{S_n\}$ is divergent; the range is infinite, and the sequence is unbounded.

(c) If $S_n = 1 + [(-1)^n/n]$ then the sequence $\{S_n\}$ converges to 1, is bounded, and has infinite range.

(d) If $S_n = i^n$ the sequence $\{S_n\}$ is divergent, is bounded and has finite range.

(e) If $S_n = 1$ (n = 1, 2, 3, ...) then $\{S_n\}$ converges to 1 is bounded.

7.2 Properties of Convergent Sequences

Theorem:

(a) $\{p_n\}$ converges to $p \in X$ if and only if every neighbourhood of p contains p_n for all but finitely many n.

(b) If $p, p' \in X$ and if $\{p_n\}$ converges to p and to p′ then $p = p'$

(c) If $\{p_n\}$ converges then $\{p_n\}$ is bounded.

(d) If $E \subseteq X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = \lim_{n\to\infty} p_n$

Theorem: Suppose $\{S_n\}$, $\{t_n\}$ are complex sequence with $\lim_{n\to\infty} S_n = S$ and $\lim_{n\to\infty} t_n = t$. Then

(a) $\lim_{n\to\infty} (S_n + t_n) = S + t$

(b) $\lim_{n\to\infty} C \cdot S_n = C \cdot S$ and $\lim_{n\to\infty} C + S_n = C + S$ for any number C.

(c) $\lim_{n\to\infty} S_n t_n = S_t$

(d) $\lim_{n\to\infty} \frac{1}{S_n} = \frac{1}{S}$

Theorem:

(a) Suppose $x_n \in \mathbb{R}^k$ $(n \in \mathbb{N})$ and $x_n = (\alpha_1, n, \ldots \alpha_k, n)$. Then $\{x_n\}$ converges to $x = (\alpha_1, \ldots, \alpha_k)$ if and only if

$$\lim_{n\to\infty} \alpha_{j,n} = \alpha_j (1 \le j \le k)$$

Notes

(b) Suppose $\{x_n\}$, (y_n) are sequence in $\mathbb{R}^k$, $\{\beta_n\}$ is a sequence of real numbers, and $x_n \to x$, $y_n \to y$, $\beta_n \to \beta$. Then

$$\lim_{n\to\infty} (x_n + y_n) = x + y$$
$$\lim_{n\to\infty} (x_n \cdot y_n) = x \cdot y$$
$$\lim_{n\to\infty} \beta_n x_n = \beta x$$

7.2.1 Subsequences

Definition:

Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers such that $n_1 < n_2 < n_3 \ldots$ Then the sequence $\{p_{ni}\}$ is called a subsequence of $\{p_n\}$. If $\{p_{ni}\}$ converges its limit is called a subsequential limit of $\{p_n\}$.

It is clear that $\{p_n\}$ converges to p if and only if every subsequence of $\{p_n\}$ converges to p.

7.3 Subsequences and Compact Metric Spaces

Theorem:

(a) If $\{p_n\}$ is a sequence in a compact metric space X, then some subsequence of $\{p_n\}$ converges to a point of X.

(b) Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.

7.4 Subsequences Limits

Theorem:

The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subset of X.

7.5 Cauchy Sequence

A sequence $\{p_n\}$ in a metric space (X, d) is said to be a Cauchy sequence if for every $\in > 0$ there is an integer N such that $d(p_n, p_m) < \in$ for all $n, m \geq N$.

Definition:

Let E be a non-empty subset of a metric space (X, d), and let $S = \{d(p, q) : p, q \in E\}$. The diameter of E is sup S.

If $\{p_n\}$ is a sequence in X and if E_n consists of the points $p_N, p_{N+1}, \ldots$, it is clear that $\{p_n\}$ is a Cauchy sequence if and only if

$$\lim_{N\to\infty} \text{diam } E_N = 0$$

7.6 Cauchy Sequence and Closed Sets

Theorem:

(a) If $\overline{E}$ is the closure of a set E in a metric space X, then

$$\text{diam } \overline{E} = \text{diam } E$$

(b) If K_n is a sequence of compact sets in X such that $K_n \subset K_{n+1}$ $(n \in \mathbb{N})$ and if $\lim_{n\to\infty} \text{diam } K_n = 0$ then $\bigcap_1^\infty K_n$ consists of exactly one point.

7.7 Cauchy Sequences and Convergent Sequences

Theorem:

(a) In any metric space X, every convergent sequence is a Cauchy sequence.

(b) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X then $\{p_n\}$ converges to some point of X.

(c) In $\mathbb{R}^k$ every Cauchy sequence converges.

7.7.1 Complete Spaces

Definition:

A metric space is said to be complete if every Cauchy sequence converges.

Notice that all compact metric spaces are complete but there are metric spaces (like $\mathbb{R}^k$) which are complete but not compact.

Lemma

Every closed subset of a complete metric space is complete.

7.8 Increasing/Decreasing Sequences

Definition:

A sequence {Sn} of real numbers is said to be

(a) Monotonically increasing if $S_n \leq S_{n+1}$ for all $n \in \mathbb{N}$

(b) Monotonically decreasing if $S_n \geq S_{n+1}$ for all $n \in \mathbb{N}$

(c) Monotonic if it is monotonically increasing or monotonically decreasing.

Theorem: Suppose $\{S_n\}$ is monotonic. Then $\{S_n\}$ converges if and only if $\{S_n\}$ is bounded.

Self Assessment

Fill in the blanks:

1. If there is any ambiguity we say $\{p_n\}$ .... in X.

2. The set of all p_n is said to be the range of $\{p_n\}$ (which may be infinite or finite). We say $\{p_n\}$ is bounded if the range is ....

3. If $\{p_n\}$ is a sequence in a .... space X, then some subsequence of $\{p_n\}$ converges to a point of X.

4. Every bounded sequence in $\mathbb{R}^k$ contains a ....

5. A metric space is said to be complete if every ....

Notes

7.9 Summary

- A sequence $\{p_n\}$ in a metric space (X, d) is said to converge if there is a point $p \in X$ with the following property:

  $(\forall \in > 0)(\exists N) \qquad (\forall n > N)\ d(p_n, p) < \in$

- In this case we also say that $\{p_n\}$ converges to p or that p is the limit of $\{p_n\}$ and we write $p_n \to p$ or lim $\lim_{n\to\infty} p_n = p$.
- If $\{p_n\}$ does not converge we say it diverges.
- If there is any ambiguity we say $\{p_n\}$ converges/diverges in X.
- The set of all p_n is said to be the range of $\{p_n\}$ (which may be infinite or finite). We say $\{p_n\}$ is bounded if the range is bounded.
- Properties of convergent sequences
  - (a) $\{p_n\}$ converges to $p \in X$ if and only if every neighbourhood of p contains p_n for all but finitely many n.
  - (b) If $p, p' \in X$ and if $\{p_n\}$ converges to p and to p′ then $p = p'$
  - (c) If $\{p_n\}$ converges then $\{p_n\}$ is bounded.
  - (d) If $E \subseteq X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = \lim_{n\to\infty} p_n$
- Theorem of couchy sequences and convergent sequences
  - (a) In any metric space X, every convergent sequence is a Cauchy sequence.
  - (b) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X then $\{p_n\}$ converges to some point of X.
  - (c) In $\mathbb{R}^k$ every Cauchy sequence converges.

7.10 Keywords

Subsequential: Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers such that $n_1 < n_2 < n_3 \ldots$ Then the sequence $\{p_{ni}\}$ is called a subsequence of $\{p_n\}$. If $\{p_{ni}\}$ converges its limit is called a subsequential limit of $\{p_n\}$.

Subsequential Limits: The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subset of X.

Cauchy Sequency: A sequences $\{p_n\}$ in a metric space (X, d) is said to be a Cauchy sequency if for every $\in > 0$ there is an integer N such that $d(p_n, p_m) < \in$ for all $n, m \geq N$.

7.11 Review Questions

1. Define convergent sequence.
2. Discuss the properties of convergent sequence.
3. Explain subsequences and compact metric spaces.
4. Describe subsequence limits.
5. Explain the Cauchy sequences and convergent sequences.

Answers: Self Assessment

1.	converges/diverges	2.	bounded
3.	compact metric	4.	convergent subsequence
5.	Cauchy sequence converges		

7.12 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Notes

Unit 8: Completeness and Compactness

CONTENTS

Objectives

Introduction

8.1 Completeness and Compactness

8.2 Cantor's Theorem

8.3 Perfect Set

8.3.1 Perfect Sets are Uncountable

8.4 Cantor Middle Third Set

8.5 Baire Category Theorem

8.6 Compactness and Cantor Set

8.7 Summary

8.8 Keywords

8.9 Review Questions

8.10 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss Completeness and Compactness
- Describe the Cantor's theorem
- Explain Baire category theorem
- Describe Compactness and Cantor set

Introduction

In earlier unit you have studied about the compactness and connectedness of the set. As you all come to know about the Contraction Mapping Theorem. After understanding the concept of Total boundedness let us go through the explanation of completeness and connectedness.

8.1 Completeness and Compactness

Theorem: Subspace C of complete metric M compact iff closed and totally bounded.

Proof: ($\Rightarrow$) C closed, totally bounded since $\forall\, \varepsilon > 0$ open cover B(x, ε) ($x \in C$) has finite subcover.

($\Leftarrow$) Every sequence in C has Cauchy subsequence, converges to point of M since M complete. C closed so limit in C.

Lemma: If M subspace of N totally bounded so is M.

Proof: Fix ε > 0. Let $F \subset cM$ be finite s.t. $M \subset \cup_{x \in F} B(x, \frac{\varepsilon}{2})$. Then

$$\overline{M} \subset \bigcup_{x \in F} \overline{B\left(x, \frac{\varepsilon}{2}\right)} \subset B(x, \varepsilon)$$

Notes

Theorem: Subspace S of complete metric M totally bounded iff $\bar{S}$ compact.

Proof: $(\Rightarrow)$ $\bar{S}$ totally bounded and so compact.

$(\Leftarrow)$ $\bar{S}$ totally bounded so is $S \subset \bar{S}$.

8.2 Cantor's Theorem

Definition: Diameter of $0 \neq S \subset M$ defined by

$$\text{diam}(S) = \sup_{x, y \in S} d(x, y)$$

Theorem: Cantor

Let F_n decreasing sequence of non-empty closed subsets of metric M s.t. diam $(F_n) \xrightarrow[n\to\infty]{} 0$. Then $\bigcap_{n=1}^{\infty} F_n \neq \theta$.

Proof: Pick $x_n \in F_n$. Then $\forall\, i \geq n$, $x_i \in F_i \subset F_n$.

Hence, for i, j ≥ n, $d(x_i, x_j) \leq \text{diam}(F_n)$. Hence (x_n) Cauchy. Converges to some x as M complete.

F_n closed so $x \in F_n$. Hence $x \in \bigcap_{n=1}^{\infty} F_n$.

8.3 Perfect Set

A set S is perfect if it is closed and every point of S is an accumulation point of S.

Example: Find a perfect set. Find a closed set that is not perfect. Find a compact set that is not perfect. Find an unbounded closed set that is not perfect. Find a closed set that is neither compact nor perfect.

Solution:

- A perfect set needs to be closed, such as the closed interval [a, b]. In fact, every point in that interval [a, b] is an accumulation point, so that the set [a, b] is a perfect set.
- The simplest closed set is a singleton {b}. The element b in then set {b} is not an accumulation point, so the set {b} is closed but not perfect.
- The set {b} from above is also compact, being closed an bounded. Hence, it is compact but not perfect.
- The set $\{-1\} \cup [0, \infty)$ is closed, unbounded, but not perfect, because the element –1 is not an accumulation point of the set.
- The set $\{-1\} \cup [0, \infty)$ from above is closed, not perfect, and also not compact, because it is unbounded.

Example: Is the set {1, 1/2, 1/3, ...} perfect? How about the set $\{1, 1/2, 1/3, ...\} \cup \{0\}$?

Solution: The first set is not closed. Hence it is not perfect.

The second set is closed, and {0} is an accumulation point. However, every point different from 0 is isolated, and can therefore not be an accumulation point. Therefore, this set is not perfect either.

Notes

As an application of the above result, we will see that perfect sets are closed sets that contain lots of points:

8.3.1 Perfect Sets are Uncountable

Every non-empty perfect set must be uncountable.

Proof: If S is perfect, it consists of accumulation points, and therefore can not be finite. Therefore it is either countable or uncountable. Suppose S was countable and could be written as

$$S = \{x_1, x_2, x_3, ...\}$$

The interval $U_1 = (x_1 - 1, x_1 + 1)$ is a neighbourhood of x_1. Since x_1 must be an accumulation point of S, there are infinitely many elements of S contained in U_1.

Take one of those elements, say x_2 and take a neighbourhood U_2 of x_2 such that closure (U_2) is contained in U_1 and x_1 is not contained in closure (U_2). Again, x_2 is an accumulation point of S, so that the neighbourhood U_2 contains infinitely many elements of S.

Select an element, say x_3, and take a neighbourhood U_3 of x_3 such that closure (U_3) is contained in U_2 but x_1 and x_2 are not contained in closure (U_3).

Continue in that fashion: we can find sets U_n and points x_n such that:

- closure $(\cup_{n+1})$ $\cup_n$
- x_j is not contained in $\cup_n$ for all $0 < j < n$
- x_n is contained in $\cup_n$

Now consider the set

- $V = \cap$ (closure $(\cup_n) \cap S$)

Then each set closure $(\cup_n) \cap S$) is closed and bounded, hence compact. Also, by construction, (closure $(\cup_{n+1}) \cap S$) (closure $(\cup_n) \cap S$). Therefore, by the above result, V is not empty. But which element of S should be contained in V? It can not be x_1, because x_1 is not contained in closure (U_2). It can not be x_2 because x_2 is not in closure $(\cup_3)$, and so forth.

Hence, none of the elements $\{x_1, x_2, x_3, ...\}$ can be contained in V. But V is non-empty, so that it must contain an element not in this list. That means, however, that S is not countable.

8.4 Cantor Middle Third Set

Start with the unit interval

$S_0 = [0, 1]$

Remove from that set the middle third and set

$S_1 = S_0 \backslash (1/3, 2/3)$

Remove from that set the two middle thirds and set

$S_2 = S_1 \backslash \{(1/9, 2/9)\ (7/9, 8/9)\}$

Continue in this fashion, where

$S_{n+1} = S_n \backslash \{\text{middle thirds of subintervals of } S_n\}$

Then the Cantor set C is defined as

$C = \cap S_n$

The Cantor set gives an indication of the complicated structure of closed sets in the real line. It has the following properties:

Example: The Cantor set is compact.

Solution: The definition of the Cantor set is as follows: let

$A_0 = [0, 1]$

and define, for each n, the sets A_n recursively as

$$A_n = A_{n-1} / \bigcup_{k=0}^{\infty} \left(\frac{1+3k}{3^n}, \frac{2+3k}{3^n} \right)$$

Then the Cantor set is given as:

$C = \cap A_n$

Each set $\bigcup_{k=0}^{\infty} \left(\frac{1+3k}{3^n}, \frac{2+3k}{3^n} \right)$ is open. Since A_0 is closed, the sets A_n are all closed as well, which can be shown by induction. Also, each set A_n is a subset of A_0, so that all sets A_n are bounded.

Example: The Cantor set is perfect and hence uncountable.

The definition of the Cantor set is as follows: let

$A_0 = [0, 1]$

and define, for each n, the sets A_n recursively as

$$A_n = A_{n-1} \setminus \bigcup_{k=0}^{\infty} \left(\frac{1+3k}{3^n}, \frac{2+3k}{3^n} \right)$$

Then the Cantor set is given as:

$C = \cap A_n$

From this representation it is clear that C is closed. Next, we need to show that every point in the Cantor set is a limit point.

One way to do this is to note that each of the sets A_n can be written as a finite union of 2^n closed intervals, each of which has a length of $1/3^n$, as follows:

$A_0 = [0, 1]$

$A_1 = [0, 1/3] \cup [2/3, 1]$

$A_2 = [0, 1/9] \cup [2/9, 3/9] \cup [6/9, 7/9] \cup [8/9, 1]$

...

Note that all endpoints of every subinterval will be contained in the Cantor set. Now take any $x\ C = \cap A_n$. Then x is in A_n for all n. If x is in A_n, then x must be contained in one of the 2^n intervals that comprise the set A_n. Define x_n to be the left endpoint of that subinterval (if x is equal to that endpoint, then let x_n be equal to the right endpoint of that subinterval). Since each subinterval has length $1/3^n$, we have:

$$|x - x_n| < 1/3^n$$

Hence, the sequence $\{x_n\}$ converges to x, and since all endpoints of the subintervals are contained in the Cantor set, we have found a sequence of numbers contained in C that converges to x.

Notes

Therefore, x is a limit point of C. But since x was arbitrary, every point of C is a limit point. Since C is also closed, it is then perfect.

Note that this proof is not yet complete. One still has to prove the assertion that each set A_n is indeed comprised of 2^n closed subintervals, with all endpoints being part of the Cantor set. But that is left as an exercise.

Since every perfect set is uncountable, so is the Cantor.

Hence, C is the intersection of closed, bounded sets, and therefore C is also closed and bounded. But then C is compact.

Example: The Cantor set has length zero, but contains uncountably many points.

Solution: The definition of the Cantor set is as follows: let

$A_0 = [0, 1]$

and define, for each n, the sets A_n recursively as

$$A_n = A_{n-1} \setminus \bigcup_{k=0}^{\infty} \left(\frac{1+3k}{3^n}, \frac{2+3k}{3^n} \right)$$

Then the Cantor set is given as:

$C = \cap A_n$

To be more specific, we have:

$A_0 = [0, 1]$

$A_1 = [0, 1] \setminus (1/3, 2/3)$

$A_2 = A_1 \setminus [(1/9, 2/9) \cup (7/9, 8/9)] =$

$[0, 1] \setminus (1/3, 2/3)) \setminus (1/9, 2/9) \setminus (7/9, 8/9)$

...

That is, at the n-th stage ($n > 0$) we remove 2^{n-1} intervals from each previous set, each having length $1/3^n$. Therefore, we will remove a total length from the unit interval [0, 1]. Since we remove a set of total length 1 from the unit interval, the length of the remaining Cantor set must be 0.

The Cantor set contains uncountably many points because it is a perfect set.

Example: The Cantor set does not contain any open set

The definition of the Cantor set is as follows: let

$A_0 = [0, 1]$

and define, for each n, the sets A_n recursively as

$$A_n = A_{n-1} \setminus \bigcup_{k=0}^{\infty} \left(\frac{1+3k}{3^n}, \frac{2+3k}{3^n} \right)$$

Then the Cantor set is given as:

$C = \cap A_n$

Notes

Another way to write the Cantor set is to note that each of the sets A_n can be written as a finite union of 2^n closed intervals, each of which has a length of $1/3^n$, as follows:

$A_0 = [0, 1]$

$A_1 = [0, 1/3] \cup [2/3, 1]$

$A_2 = [0, 1/9] \cup [2/9, 3/9] \cup [6/9, 7/9] \cup [8/9, 1]$

...

Now suppose that there is an open set U contained in C. Then there must be an open interval (a, b) contained in C. Now pick an integer N such that

$$1/3^N < b - a$$

Then the interval (a, b) can not be contained in the set A_N, because that set is comprised of intervals of length $1/3^N$. But if that interval is not contained in A_N it can not be contained in C. Hence, no open set can be contained in the Cantor set C.

8.5 Baire Category Theorem

Definition: $S \subset M$ is

- Dense in M if $\bar{S} = M$.
- Nowhere dense in M if $M \setminus \bar{S}$ is dense in M.
- Meagre in M if it is the union of a sequence of nowhere dense sets.

Proposition: $S \subset M$ nowhere dense in M iff $\bar{S}^{\circ} = \emptyset$

Proof: $\bar{S}^{\circ} = \emptyset = M \setminus \overline{(M \setminus \bar{S})}$ so if RHS $= \emptyset$ then $M \setminus \bar{S}$ is dense in M so S is nowhere dense.

Conversely if S is nowhere dense in M then $M \setminus \bar{S} = M$ so RHS $= \emptyset$.

Theorem: Baire Category

A complete metric space is not meagre in itself.

I.e. if S_n are the nowhere dense subsets of non-empty complete M then

$$M \setminus \bigcup_{n=1}^{\infty} Sn \neq \emptyset$$

Proof: IDEA: Find decreasing sequence of dense sets with non-empty intersection of their closures by Cantor. Any point in this intersection cannot be in any nowhere dense set.

$G_k := M \setminus \overline{S_k}$ dense in M, open.

Then $G_1 \neq \emptyset$. Choose $x_1 \in G_1$ and $\delta_i > 0$ s.t. $B(x_i, \delta_i) \subset G_1$.

Continue inductively: Having defined x_{k-i}, δ_{k-i} use fact that G_k dense to find $x_k \in G_k \cap B\left(x_{k-1}, \frac{\delta_{k-1}}{2}\right)$. Find $0 < \delta_k < \frac{\delta_{k-1}}{2}$ s.t. $B(x_k, \delta_k) \subset G_k$.

$\delta_k \xrightarrow[k\to\infty]{} 0$ and $\forall$ k, $\overline{B(x_k, \delta_k)} \subset B(x_{k-i}, \delta_{k-i})$.

Notes

Then by Cantor (5.15) $\bigcap_{k=1}^{\infty} \overline{B(x_k, \delta_k)} \neq \emptyset$. Let x be in this intersection. Then $x \in B(x_k, \delta_k) \subset G_k$ $\forall$ k so $x \notin S_k$ $\forall$ k. Hence, there is a point x that is not in the union of all nowhere dense subsets of M, so M cannot be meagre.

Proposition: The Cantor set $\mathfrak{C}$ is uncountable.

Proof: $\forall x \in \mathfrak{C}$ there are points $y \in \mathfrak{C}$, $y \neq x$ arbitrarily close to x. In other words, $\mathfrak{C}\backslash\{x\}$ is dense in $\mathfrak{C}$. Therefore $\{x\}$ is nowhere dense in $\mathfrak{C}$ as it is closed.

If $\mathfrak{C}$ were countable would have $\mathfrak{C} = \bigcup_j^{\infty}\{x_j\}$ showing $\mathfrak{C}$ meagre in itself. This contradicts Baire's theorem.

Lemma: Let f: $[1, \infty) \to \mathbb{R}$ be cts s.t. for some $a \in \mathbb{R}$ $\exists$ arbitrarily large x with f(x) < a. Then $\forall k \in \mathbb{N}$: $S = \bigcup_{n=k}^{\infty}\{x \in [1,\infty) : f(nx) \geq a\}$ is nowhere x with dense.

Proof: f cts so S closed. Let $1 \leq \alpha < \beta < \infty$. RTP $(\alpha, \beta)\backslash S \neq \emptyset$

For large n, $\frac{n+1}{n} < \frac{\beta}{\alpha}$ so $(n + 1)\alpha < n\beta$. Then $\bigcup_{n=k}^{\infty}(n\alpha, n\beta)$ contains some (r, ∞) and so a point y s.t. f(y) < a.

Find n s.t. $y \in (n\alpha, n\beta)$. Then $x = \frac{y}{n} \in (\alpha, \beta)$ and f(nx) < a so $x \notin S$.

Proposition: Let f: $[1, \infty) \to \mathbb{R}$ be cts s.t. $\forall x \geq 1$, $\lim_{n\to\infty} f(nx)$ exists. Then $\lim_{x\to\infty} f(x)$ exists.

Proof: If $\lim_{x\to\infty} f(x)$ not exist then $\exists$ a, b; a < b s.t. $\exists$ arbitrarily large x,y with f(x) < a, f(y) > b.

Then by previous lemma:

$$\bigcup_{k=1}^{\infty}\bigcap_{n=1}^{\infty}\{x \in [1, \infty) : f(nx) \geq a\} \cup \bigcup_{k=1}^{\infty}\bigcap_{n=1}^{\infty}\{x \in [1, \infty) : f(nx) \leq b\}$$

is meagre. By Baire $\exists x \notin T$.

x not in first union so $\forall k \,\exists n \geq k$ s.t. f(nx) < a. x not in second union so $\forall k \,\exists m \geq k$ s.t. f(mx) > b. Hence f(nx) not converge.

Theorem: $\exists f \in C[0, 1]$ not differentiable at any point.

Proof: IDEA: C[0, 1] is complete. Functions with derivative at at least one point form a meagre subset. Result by Baire.

Define S_n:

$$S_n = \{f \in C[0, 1] : (\exists x \in [0, 1])(\forall y \in [0, 1]) \ |f(y) - f(x)| \leq n|y - x|\}$$

8.6 Compactness and Cantor Set

Theorem: Every compact metric M is continuous image of Cantor set $\mathfrak{C}$.

Proof: Let $A_k \subset M$ be finite s.t. $\forall x \in M \quad d(A_k, x) \leq 2^{-k}$.

By induction construct sequence of cts functions $f_k : \mathfrak{C} \to M$ s.t. $f_k(\mathfrak{C}) = A_k$, $d(f_k(x), f_{k+1}(x)) \leq 2^k$ $\forall x \in C$.

Notes

(f_k) Cauchy in $C(\mathfrak{C}, M)$ so converge to cts $f : \mathfrak{C} \to M$. $f(\mathfrak{C})$ dense in M. Also compact, so closed, hence $f(\mathfrak{C}) = M$.

Corollary: $\exists$ continuous surjective map $f : [0, 1] \to [0, 1]$.

Proof: Extend surjective cts $f : (\mathfrak{C} \to [0, 1]^2$ linearly to each interval removed during construction of $\mathfrak{C}$.

Self Assessment

Fill in the blanks:

1. A complete .... is not meagre in itself.
2. The Cantor set $\mathfrak{C}$ is ....
3. Let f: $[1, \infty) \to \mathbb{R}$ be cts s.t. for some $a \in \mathbb{R}$ $\exists$ .... x with $f(x) < a$. Then $\forall k \in \mathbb{N} : S = \bigcup_{n=k}^{\infty} \{x \in [1,\infty) : f(nx) \geq a\}$ is nowhere x with dense.
4. Subspace S of complete metric M totally .... compact.

8.7 Summary

- S_n closed.
- S_n nowhere dense as has dense complement and closed.
- If $f'(x)$ exists for some x then $f \in S_n$ for some n.
- Let $f_k \in S_n$, $f_k \to f$. Find $x_k \in [0, 1]$ s.t. $\forall y \in [0, 1]$,

  $$|f_k(y) - f_k(x_k)| \leq n|y - x_k|$$

  x_k has convergent subsequence so assume $x_k \to x$. By uniform convergence

  $$|f(y) - f(x)| \leq n|y - x|$$

  Therefore $f \in S_n$, so S_n closed.
- Let $g \in C[0, 1]$, $\in > 0$. g uniformly cts so $\exists\ \delta > 0$ s.t.

  $$|x - y|\delta \Rightarrow |g(x) - g(y)| < \frac{\varepsilon}{4} \qquad \text{. . . (1)}$$

  Let $x_i\ \frac{i}{k} = \varphi(x) = k\varepsilon \min_{0 \leq i \leq k} |x - x_i|$. Then $0 \leq \varphi \leq \frac{\varepsilon}{2}$ show suffices to show $f = \varphi + g \notin S_n$.

  Suppose $f \in S_n$ and find x "responsible for it".

  Choose $1 \leq j \leq k$ s.t. $x \in [x_{j-1}, x_j]$. Let $y = \frac{x_{j-1} + x_j}{2}$

  $$\begin{aligned}\frac{\varepsilon}{2} &= |\varphi(y) - \varphi(x_j)| \\ &\leq |f(y) - f(x_j)| + |g(y) - g(x_j)| \\ &\underset{(1)}{\leq} |f(y) - f(x)| + |f(x_j) - f(x)| + \frac{\varepsilon}{4}\end{aligned}$$

Notes

$$\leq |n|y-x| + n|x_j - x| + \frac{\varepsilon}{4}$$

$$\leq \frac{2n}{k} + \frac{\varepsilon}{4}$$

$$\leq \frac{\varepsilon}{2}$$

This is a contradiction. So $f \notin S_n$.

- If f′ (x) exists find δ > 0 s.t. $\forall\, 0 < |y - x| < \delta$,

$$\left|\frac{f(y)-f(x)}{y-x} - f'(x)\right| < \left|\frac{f(y)-f(x)}{y-x}\right| < 1 + |f'(x)|$$

Function $y \mapsto \frac{f(y)-f(x)}{y-x}$ is continuous on $[0, 1]\backslash(x - \delta, x + \delta)$ which is compact. Hence the function is bounded, so $\exists\, n \in \mathbb{N}$ s.t.

$$y \in [0, 1]\backslash(x-\delta, x+\delta) \Rightarrow \left|\frac{f(y)-f(x)}{y-x}\right| \leq n$$

May take $n > 1 + |f'(x)|$ so get inequality holding $\forall\, y \in [0, 1]\backslash\{x\}$.

Then $|f(y) - f(x)| \leq n|y - x|$ $\forall\, y \in [0, 1]$. (This clearly holds for y = x and holds by the above for y ≠ x.) So if $\exists\, f \in C[0, 1]$ s.t. f′(x) exists for some x then $f \in S_n$.

- These three parts together complete the proof, since by Baire (5.17) C[0, 1] is not meagre, so there must be a function which is not differentiable at any point, as any that are differentiable at at least one point are in a nowhere dense subset.

8.8 Keywords

Complete Metric: Subspace C of complete metric M compact iff closed and totally bounded.

Cantor: Let F_n decreasing sequence of non-empty closed subsets of metric M s.t. diam $(F_n) \xrightarrow[n\to\infty]{} 0$. Then $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$.

Continue Inductively: Having defined x_{k-i}, δ_{k-i} use fact that G_k dense to find $x_k \in G_k \cap B\left(x_{k-1}, \frac{\delta_{k-1}}{2}\right)$. Find $0 < \delta_k < \frac{\delta_{k-1}}{2}$ s.t. $B(x_k, \delta_k) \subset G_k$.

$\delta_k \xrightarrow[k\to\infty]{} 0$ and $\forall\, k$, $\overline{B(x_k, \delta_k)} \subset B(x_{k-i}, \delta_{k-i})$.

8.9 Review Questions

1. Discuss Completeness and Compactness.
2. Describe the Cantor's theorem.
3. Explain Baire category theorem.
4. Describe Compactness and Cantor set.

Answers: Self Assessment

Notes

1. metric space
2. uncountable
3. arbitrarily large
4. bounded iff $\overline{S}$

8.10 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Notes

Unit 9: Functions

CONTENTS

Objectives
Introduction
9.1 Algebraic Functions
9.1.1 Algebraic Combinations of Functions
9.1.2 Notion of an Algebraic Function
9.2 Transcendental Functions
9.2.1 Trigonometric Functions
9.3 Inverse Trigonometric Functions
9.3.1 Monotonic Functions
9.3.2 Logarithmic Function
9.3.3 Exponential Function
9.4 Some Special Functions
9.4.1 Identity Function
9.4.2 Periodic Function
9.4.3 Modulus Function
9.4.4 Signum Function
9.4.5 Greatest Integer Function
9.4.6 Even and Odd Functions
9.4.7 Bounded Functions
9.5 Summary
9.6 Keywords
9.7 Review Questions
9.8 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss the different types of algebraic functions
- Explain the trigonometric and the inverse trigonometric functions
- Describe the exponential and logarithmic functions
- Explain some special functions including thus bounded and monotonic functions

Introduction

Real Analysis is often referred to as the Theory of Real Functions. The word 'function' was first introduced in 1694 by L.G. Leibniz [1646-1716], a famous German mathematician, who is also

Notes

credited along with Isacc Newton for the invention of Calculus, Leibniz used the term function to denote a quantity connected with a curve. A Swiss mathematician, L. Euler [1707-1783] treated function as an expression made up of a variable and some constants. Euler's idea of a function was later generalized by an eminent French mathematician J. Fourier [1768-1830]. Another German mathematician, L. Dirichlet (1805-1859) defined function as a relationship between a variable (called an independent variable) and another variable (called the dependent variable). This is the definition which, you know, is now used in Calculus.

The concept of a function has undergone many refinements. With the advent of Set Theory in 1895, this concept was modified as a correspondence between any two non-empty sets. Given any two non-empty sets S and T, a function f from S into T, denoted as f: $S \rightarrow T$, defines a rule which assigns to each $x \in S$, a unique element Leonard Euler $y \in T$. This is expressed by writing as y = f (x). This definition, as you will recall, was given in Section 1.2. A function f $S \rightarrow T$ is said to be a

1. Complex-valued function of a complex variable if both S and T are sets of complex numbers;
2. Complex-valued function of a real variable if S is a set of real numbers and T is a set of complex numbers;
3. Real-valued function of a complex variable if S is a set of complex numbers and T is a set of real numbers;
4. Real-valued function of a real variable if both S and T are some sets of real numbers.

Since we are dealing with the course on Real Analysis, we shall confine our discussion to those functions whose domains as well as co-domains are some subsets of the set of real numbers. We shall call such functions as Real Functions.

In this unit, we shall deal with the algebraic and transcendental functions. Among the transcendental functions, we shall define the trigonometric functions, the exponential and logarithmic functions. Also, we shall talk about some special real functions including the bounded and monotonic functions. We shall frequently use these functions to illustrate various concepts in Blocks 3 and 4.

9.1 Algebraic Functions

In Unit 1, we identified the set of natural numbers and built up various sets of numbers with the help of the algebraic operations of addition, subtraction, multiplication, division etc. In the same way, let us construct new functions from the real functions which we have chosen for our discussion. Before we do so, let us review the algebraic combinations of the functions under the operations of addition, subtraction, multiplication and division on the real-functions.

9.1.1 Algebraic Combinations of Functions

Let f and g be any two real functions with the same domain S C R and their co-domain as the set R of real numbers. Then we have the following definitions:

Definition 1: Sum and Difference of Two Functions

1. The Sum of f and g, denoted as f + g, is a function defined from S into R such that

   $(f + g)(x) = f(x) + g(x)$, If $x \in S$.

2. The Difference of f and g, denoted as f – g, is a function defined from S to R such that

   $(f - g)(x) = f(x) - g(x), \ \forall x \in S$.

Notes

Note that both f(x) and g(x) are elements of R. Hence each of their sum and difference is again a unique member of R.

Definition 2: Product of Two Functions

Let f: S →R and g: S → R be any two functions. The product of f and g, denoted as f, g, is defined as a function f. g: S → R by

$$(f . g) (x) = f(x) . g(x), \ \forall x \in S.$$

Definition 3: Scalar Multiple of a Function

Let f: S → R be a function and k be same fixed real number. Then the scalar multiple of 'f' is a function k f S → R defined by

$$(kf (x) = k. f(x), \ \forall x \in S.$$

This is also called the scalar multiplication.

Definition 4: Quotient of Two Functions

Let f: S → R and g: S → R be any two functions such that g(x) # 0 for each x in S. Then s function $\frac{f}{g}$: S → R defined by

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}, \forall x \in S$$

is called the quotient of the two functions.

Exercise 1: Let f, g, h be any three functions, defined on S and taking values in R, as f (x) = ax^2, g(x) = bx for every x in S, where a, b, are fixed real numbers. Find f + g, f – g, f, g, f/g and kf, when k is a constant.

9.1.2 Notion of an Algebraic Function

You are quite familiar with the equations ax + b = 0 and $ax^2 + bx + c = 0$, where a, b, c ∈ R, a ≠ 0. These equations, as you know are, called linear (or first degree) and quadratic (or second degree) equations, respectively. The expressions ax + b and $ax^2 + bx + c$ are, respectively, called the first and second degree polynomials in x. In the same way an expression of the form ax′ + bx^2 + cx + d (a # 0, a, b, c, d ER) is called a third degree polynomial (cubic polynomial) in x. In general, an expression of the form $a_o x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n$ where $a_0 \neq 0$, a ER, i = 0, 1, 2, ..., n, is called an nth degree polynomial in x.

A function which is expressed in the form of such a polynomial is called a polynomial function. Thus, we have the following definition:

Definition 5: Polynomial Function

Let a_1 (i = 0, 1, ..., n) be fixed real numbers where n is some fixed non-negative integer. Let S be a subset of R. A function f: S → R defined by

$$f(x) = a_0 x^n + a_1 x^{a-1} + a_2 x^{n-2} + ... + a_n, \ \forall x \in S, a_0 \neq 0.$$

is called a polynomial function of degree n.

Let us consider some particular cases of a polynomial function on R:

Suppose f : S → R is such that

(i) f(x) = k, ∀ x ∈ S (k is a fixed real number). This is a polynomial function. This is generally called a constant function on S.

Notes

Example:

$f(x) = 2, f(x) = -3, f(x) = \pi, \forall\ x \in R$, are all constant functions.

(ii) One special case of a constant function is, obtained by taking

$k = 0$ i.e. when

$f(x) = 0, \forall x \in S$.

This is called the zero function on S.

Let $f: S \to R$ be such that

(iii) $f(x) = a_0 x + a_1, \forall x \in S, a, \# 0$.

This is a polynomial function and is called a linear function on S. For example,

$f(x) = 2x + 3, f(x) = -2x + 3,$

$f(x) = 2x - 3, f(x) = -2x - 3, f(x) = 2x$ for every

$x \in S$ are all linear functions

(iv) The function $f: S \to R$ defined by

$f(x) = x, \forall x \in S$

s called the identity function on S,

(v) $f: S \to R$ given as .

$f(x),=a_0.x^2 + a_1 x + a_2, \forall x \in R, a_o \# O.$

is a polynomial function of degree two and is called a quadratic function on S.

Example: $f(x) = 2x^2 + 3x - 4, f(x) = x^2 + 3, f(x) = x^2 + 2x,$

$f(x) = -3x^2,$

for every $x \in S$ are all quadratic functions.

Definition 6: Rational Function

A function which can be expressed as the quotient of two polynomial functions is called a rational function.

Thus a function $f: S \to R$ defined by

$$f(x) = \frac{a_0x^n + a_1x^{n-1} + ... + a_n}{b_0x^m + b_1 x^{m-1} + ... + b_m},\ \forall x \in S$$

is called a rational function.

Here $a_0 \# b_0 \# 0, a_i, b_j \in R$ where i, j are some fixed real numbers and the polynomial function in the denominator is never zero.

Example: The following are all rational functions on R.

$$\frac{2x+3}{x^2+1}, \frac{4x^2+3x+1}{3x-4}(x \neq \frac{4}{3}) \text{ and } \frac{3x+5}{x-4}(x \neq 4).$$

Notes

The functions which are not rational are known as irrational functions. A typical example of an irrational function is the square root function which we define as follows:

Definition 7: Square Root Function

Let S be the set of non-negative real numbers. A function f: S → R defined by

$$f(x) = \sqrt{x}, \quad \forall x \in S$$

is called the square root function.

You may recall that $\sqrt{x}$ is the non-negative real number whose square is x. Also it is defined for all $x \geq 0$.

Polynomial functions, rational functions and the square root function are some of the examples of what are known as algebraic functions. An algebraic function, in general, is defined as follows

Definition 8: Algebraic Function

An algebraic function f : S → R is a function defined by y = f(x) if it satisfies identically an equation of the form

$$p_0(x)y^n + p_1(X)y^{n-1} + \ldots + p_{n-1}(x)y + p_n(x) = 0$$

where $p(x), p_1(x), \ldots\ p_{n-1}(x), p_n(x)$ are Polynomials in x for all x in S and n is a positive integer.

Example: Show that f: R → R defined by

$$f(x) = \sqrt{\frac{x^2 - 3x + 2}{4x - 1}}$$

is an algebraic function.

Solution:

$$\text{Let} \quad y = f(x) = \sqrt{\frac{x^2 - 3x + 2}{4x - 1}}$$

Then $(4x - 1)y^2 - (x^2 - 3x + 2) = 0$

Hence f(x) is an algebraic function.

In fact, any function constructed by a finite number of algebraic operations (addition, subtraction, multiplication, division and root extraction) on the identity function and the constant function, is an algebraic function.

Example: The functions f : R → R defined by

(i) $$f(x) = \frac{(x^2 + 2)\sqrt{x - 1}}{x^2 + 4}$$

or $$f(x) = \frac{x^2 + 2x}{\sqrt{x.(3x^2 + 5)}}$$

are algebraic functions.

Notes

Example: Prove that every rational function is an algebraic function.

Solution: Let f : R → R be given as

$$f(x) = \frac{p(x)}{q(x)}, \ \forall x \in R,$$

where p(x) and q(x) are some polynomial functions such that $q(x) \neq 0$ for any $x \in R$.

Then we have

$$y = f(x) = \frac{p(x)}{q(x)}$$

$$q(x)\, y - p(x) = 0$$

which shows that y = f(x) can be obtained by solving the equation

$$q(x)\, y - p(x) = 0.$$

Hence f(x) is an algebraic function.

A function which is not algebraic is called a Transcendental Function. Examples of elementary transcendental functions are the trigonometric functions, the exponential functions and the logarithmic functions, which we discuss in the next section.

9.2 Transcendental Functions

In earlier unit, we gave a brief introduction to the algebraic and transcendental numbers. Recall that a number is said to be an algebraic if it is a root of an equation of the form

$$a_0 x^n + a_1 x^{n-1} + ... x + a_{n-1} x + a_n = 0$$

with integral coefficients and $a_0 \neq 0$, where n is a positive integer. A number which is not algebraic is called a transcendental number. For example the numbers e and IT are transcendental numbers. In fact, the set of transcendental numbers is uncountable. Based on the same analogy, we have the transcendental functions. We have discussed algebraic functions. The functions that are non-algebraic are called transcendental functions. In this section, we discuss some of these functions.

9.2.1 Trigonometric Functions

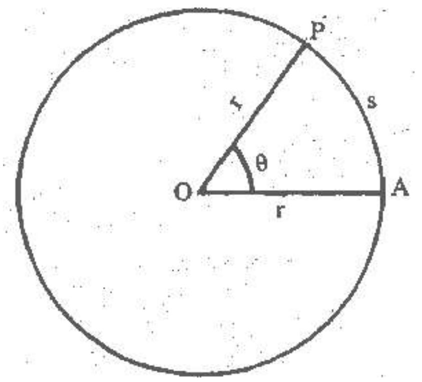
You are quite familiar with the trigonometric functions from the study of Geometry and Trigonometry. The study of Trigonometry is concerned with the measurement of the angles and the ratio of the measures of the sides of a triangle. In Calculus, the trigonometric functions have an importance much greater than simply their use in relating sides and angles of a triangle. Let us review the definitions of the trigonometric functions sin x, cos x and some of their properties. These functions form an important class of real functions.

Consider a circle $x^2 + y^2 = r^2$ with radius r and centre at O. Let P be a point on the circumference of this circle. If θ is the radian measure of a central angle at the centre of the circle as shown in the Figure 9.1 then you know that the lengths of the arc AP is given by

$$s = \theta r.$$

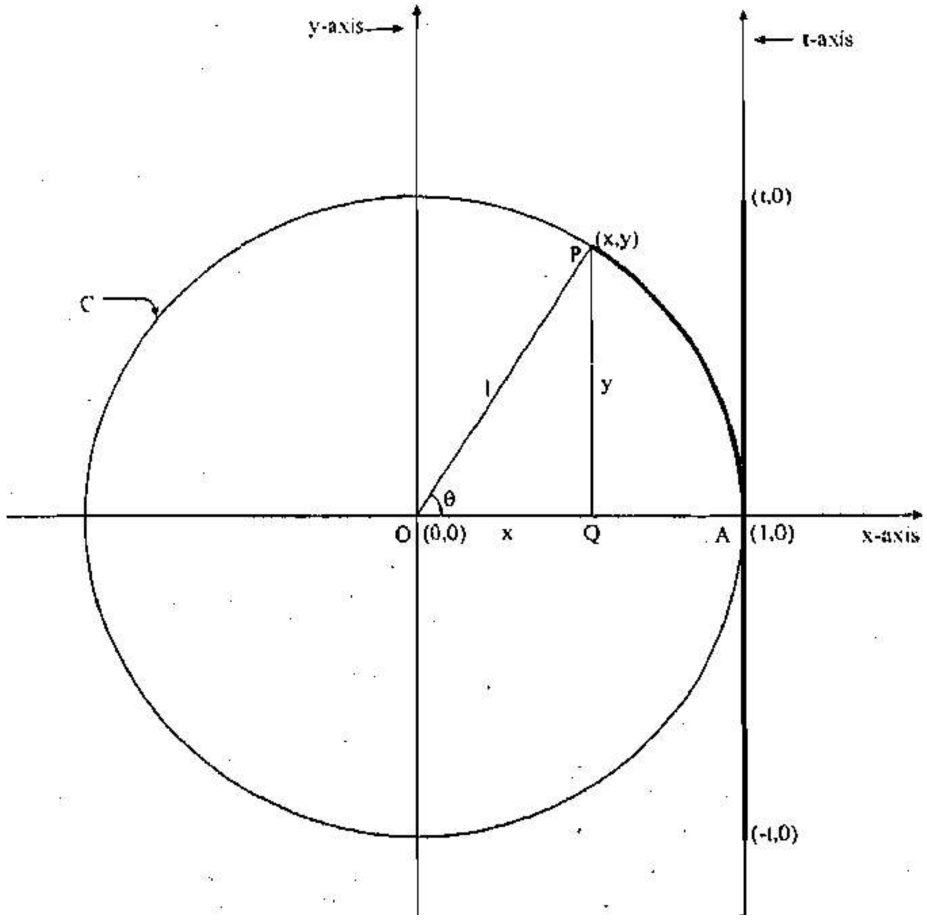
Notes

Figure 9.1



You already know how the trigonometric ratios sin θ, cos θ, etc., are defined for an angle θ measured in degrees or radians. We now define sin x, cos x, etc., for x ∈ R.

Figure 9.2



If we put r = 1 in above relation, then we get θ = s. Also the equation of circle becomes $x^2 + y^2 = 1$. This, as you know, is the Unit Circle. Let C represents this circle with centre O and radius 1. Suppose the circle meets the x-axis at a point A as shown in the Figure 9.2.

Through the point A = (1, 0); we draw a vertical line labeled as t-axis with origin at A and positive direction upwards. Now, let t be any real number and we will think of this as a point on this verticle number line i.e., t-axis.

Imagine this t-axis as a line of thread that can be wrapped around the circle C. Let p(t) = (x, y) be the point where 't' ends up when this wrapping takes place. In other words, the line segment from A to point (t, 0) becomes the arc from A to P, positive or negative i.e., counterclockwise or clockwise, depending on whether $t > 0$ or $t < 0$. Ofcourse, when t = 0, P = A. Then, the trigonometric functions 'sine' and 'cosine', for arbitrary $t \in R$, are defined by

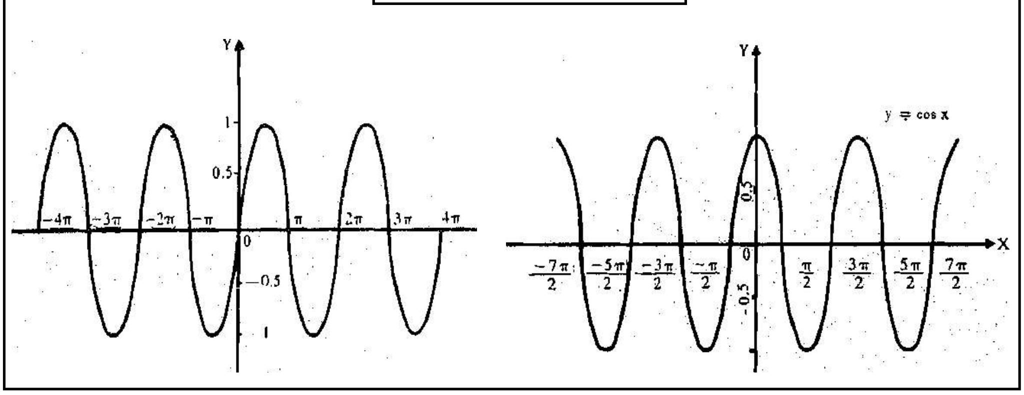
$$\sin t = \sin \theta = y, \text{ and } \cos t = \cos \theta = x,$$

where 'θ' is the radian measure of the angle subtended by the arc AP at the centre of the circle C. More generally, if t is any real number, we may take $(0 < \theta < 2\pi)$ to be the angle (rotation) whose radian measure is t. It is then clear that

$$\sin (t + 2\pi) = \sin t \text{ and } \cos (t + 2\pi) = \cos t.$$

You can easily see that as θ increases from 'θ' to $\pi/2$, PQ increases from 0 to 1 and OQ decreases from 1 to 0. Further, as θ increases from $\frac{\pi}{2}$ to θ, PQ decreases from 1 to 0 and OQ decreases from 0 to –1. Again as θ increases from π to $\frac{3\pi}{2}$, PQ decreases from 0 to –1 and OQ increases from –1 to 0. As θ increases from –1 to 0. The graphs of these functions take the shapes as shown in Figure 9.3.

Figure 9.3



Thus, we define sin x and cos x as follows:

Definition 9: Sine Function

A function $f : R \to R$ defined by

$$f(x) = \sin x, \quad \forall x \in R$$

is called the sine of x. We often write y = sin x.

Definition 10: Cosine Function

A function $f : R \to R$ defined by

$$f(x) = \cos x, \quad \forall x \in R$$

is called the cosine of x and we write y = cos x.

Notes

Note that the range of each of the sine and cosine, is [–1, 1]. In terms of the real functions sine and cosine, the other four trigonometric functions can be defined as follows:

(i) A function $f : S \to R$ defined by

$$f(x) = \tan x = \frac{\sin x}{\cos x}, \cos x \neq 0, \ \forall x \in S = R - \{(2n + 1) \frac{\pi}{2}\}$$

is called the cos x Tangent Function. The range of the tangent function is $]-\infty, +\infty[= R$ and the domain is $S = R - \{(2n + 1) \frac{\pi}{2}\}$, where n is a non-negative integer.

(ii) A function $f : S \to R$ defined by

$$f(x) = \cot x = \frac{\cos x}{\sin x}, \sin x \neq 0, \ \forall x \in S = S - \{n \pi\},$$

is said to be the Cotangent Function. Its range is also same as its co-domain i.e. range $=]-\infty, \infty[= R$ and the domain is $S = R - \{n\pi\}$ where n is a non-negative integer.

(iii) A function $f : S \to R$ defined by

$$f(x) = \sec x = \frac{1}{\cos x}, \cos x \neq 0, \ \forall x \in S = S - \{2n + 1) \frac{\pi}{2}\},$$

is called the Secant Function. Its range is the set

$$S =]-\infty, -1] \cup [1, \infty[\text{ and domain is } S = R - \{2n + 1) \frac{\pi}{2}\}.$$

(iv) A function $f : S \to R$ defined by

$$f(x) = \operatorname{cosec} x = \frac{1}{\sin x}, \sin x \neq 0, x \in S = R - \{n\pi\},$$

is called the Cosecant function. Its range is also the set $S =]-\infty, -1] \cup [1, \infty[$ and domain is $S = R - \{n \pi),$

The graphs of these functions are shown in the Figure 9.4.

Example: Let $S = [-\frac{\pi}{2}, \frac{\pi}{2}]$. Show that the function $f : S \to R$ defined by

$$f(x) = \sin x, \ \forall x \in S$$

is one-one. When is f only onto? Under what conditions f is both one-one and onto?

Solution: Recall from Unit 1 that a function f is one-one if

$$f(X_1) = f(X_2) \Rightarrow X_1 = X_2$$

for every x_1, X_2 in the domain of f.

Therefore, here we have for any $X_1, X_2 \in S$,

$$f(x_1) = f(x_2) \Rightarrow \sin x_1 = \sin x_2$$

$$\Rightarrow \sin x_1 - \sin X_2 = 0$$

$$\Rightarrow 2 \sin \frac{x_1 - x_2}{2} \cos \frac{x_1 + x_2}{2} = 0$$

Notes

$\Rightarrow$ Either sin $\frac{x_1 - x_2}{2}$ = 0, or cos $\frac{x_1 + x_2}{2}$ = 0.

If sin $\frac{x_1 - x_2}{2}$ = 0, then $\frac{x_1 - x_2}{2}$ = 0, ± π, + 2π, ...

If cos $\frac{x_1 + x_2}{2}$ = 0, then $\frac{x_1 + x_2}{2}$ = 0, ± $\frac{\pi}{2}$, ± $\frac{3\pi}{2}$, ...

Since x_1, $x_2 \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Therefore we can only have

$$-\frac{\pi}{2} \leq \frac{x_1 - x_2}{2} \leq \frac{\pi}{2}$$

and $$-\frac{\pi}{2} \leq \frac{x_1 - x_2}{2} \leq \frac{\pi}{2}$$

Thus, $\frac{x_1 - x_2}{2}$ = 0 i.e., $x_1 = x_2$. Also, if $\frac{x_1 + x_2}{2} = \pm \frac{\pi}{2}$

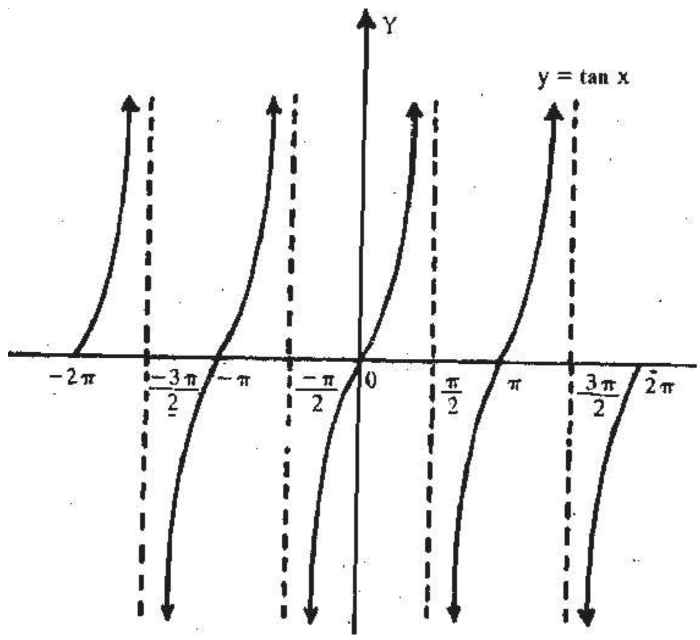
i.e. then $x_1 + x_2 = \pm x$.

Since $x_{,}, x, \in [-\frac{x}{2}, \frac{x}{2}]$,

therefore, $x_1 = x_2 = \frac{\pi}{2}$ or $x, = x_1 = -\frac{\pi}{2}$

Hence $(x_1) = f(x_2) \Rightarrow x, = x_{,}$, which proves that f is one-one. Then function f(x) = sin x defined as such, is not onto because you know that the range of sin x is [–1, 1] ≠ R.

Figure 9.4



Notes

Figure 9.5

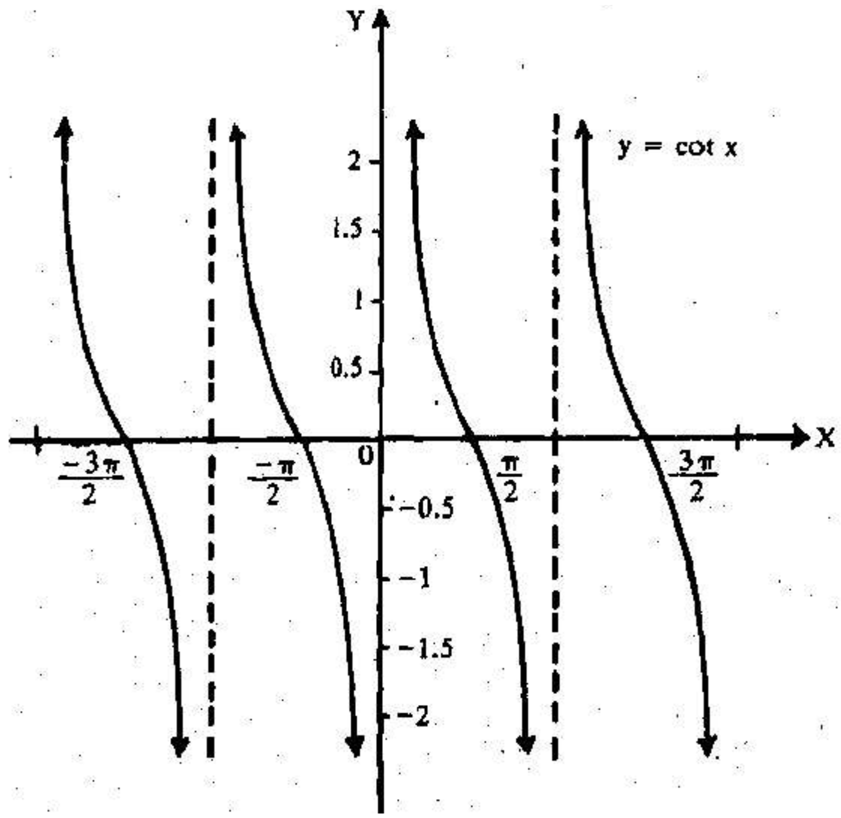
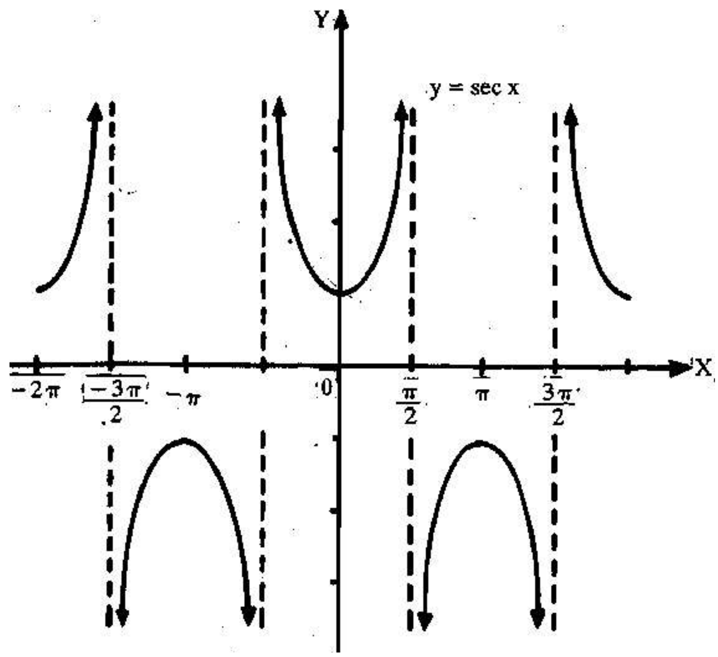
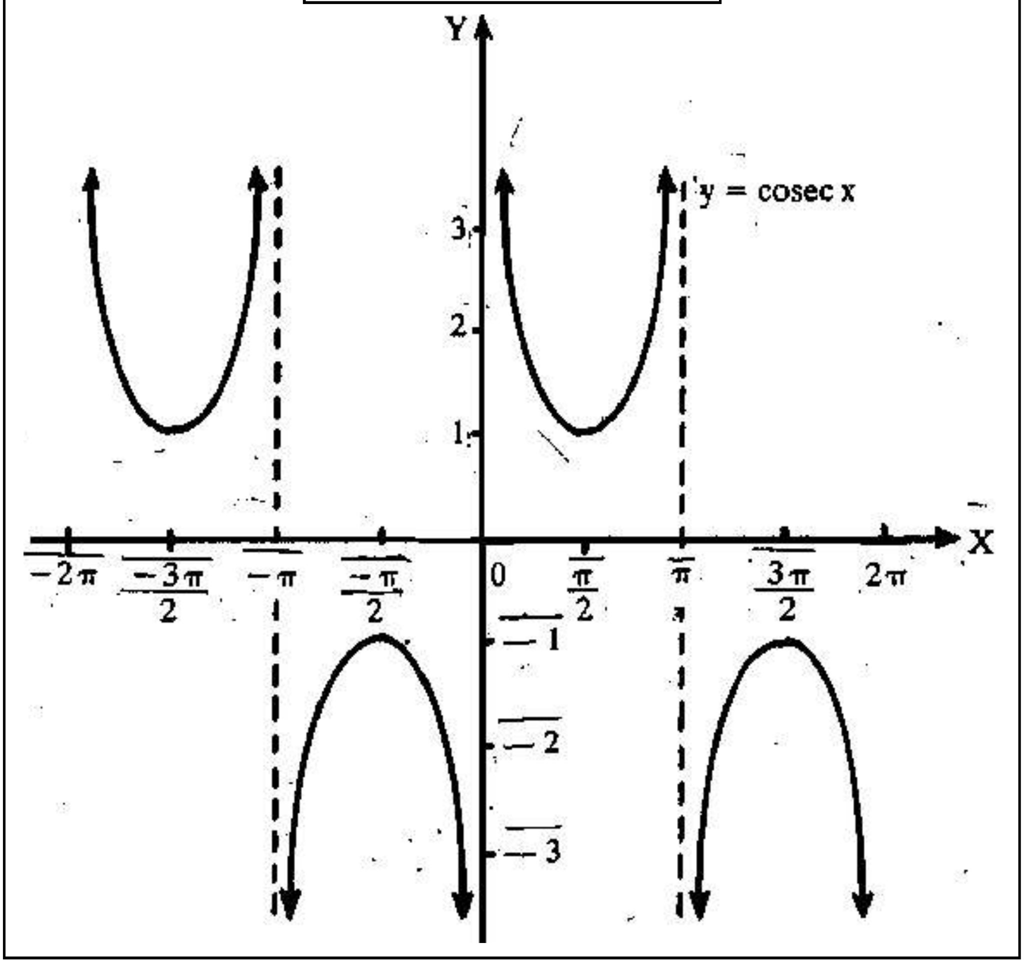


Figure 9.6



Notes

Figure 9.7



If you define f : R → [–1, 1] as

$f(x) = \sin x, \ \forall x \in R$

Then f is certainly onto. But then it is not one-one. However the function.

$f : \frac{\pi}{2} [–, \frac{\pi}{2}] \to [–1, 1]$ defined by

$f(x) = \sin x, \ \forall x \in R$

is both one-one and onto.

Exercise 2: Two functions g and h are defined as follows:

(i) g : S → R defined by

$g(x) = \cos x, x \in S = [0, \pi]$

(ii) h : S → R defined by

$h(x) = \tan x, x \in S =]-\frac{\pi}{2}, \frac{\pi}{2}[$

Show that the functions are one-one. Under what conditions the function are one-one and onto?

Notes

9.3 Inverse Trigonometric Functions

Here we discussed inverse functions. You know that if a function is one-one and onto, then it will have an inverse. If a function is not one-one and onto, then sometimes it is possible lo restrict its domain in some suitable manner such that the restricted function is one-one and onto. Let us use these ideas to define the inverse trigonometric functions. We begin with the inverse of the sine function.

Refer to the graph of f(x) = sin x in Figure 9.8. The x-axis cuts the curve y = sin x at the points x = 0, x = π, x = 2π. This shows that function f(x) = sin x is not one-one. If we restrict the domain of f(x) = sinx to the interval [–π/2, 71/21], then the function

$$f : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1] \text{ defined by}$$

$$f(x) = \sin x, -\frac{71}{21} \leq x \leq \frac{71}{21}$$

is one-to-one as well as onto. Hence it will have the inverse. The inverse function is called the inverse sine of x and is denoted as sin x. In other words,

$$y = \sin\text{-'}x \Leftrightarrow x = \sin y,$$

where $-\frac{71}{21} \leq y \ \frac{\pi}{2}$ and $-1 \leq x \leq 1$.

Thus, we have the following definition:

Definition 11: Inverse Sine Function

A function $g : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ defined by

$$g(x) = \sin^{-x} x, \quad \forall\, x \text{ s } [-1, 1]$$

is called the inverse sine function.

Again refer back to the graph of f(x) = cos x in Figure. You can easily see that cosine function is also not one-one. However, if you restrict the domain of f(x) = cos x to the interval [0, π], then the function f : [0, π] → [–1, 1] defined by

$$f(x) = \cos 0, \mid x \leq \pi,$$

is one-one and onto. Hence it will have the inverse. The inverse function is called the inverse cosine of x and is denoted by $\cos^{-1} x$ (or by arc cos x). In other words,

$$y = \cos^{-1} x \Leftrightarrow x = \cos y,$$

where 0 I y ≤ π and –1 | x ≤ I.

Thus, we have the following definition:

Definition 12:

A function g : [–1, 1] → {0, π] defined by

$$g(x) = \cos^{-1} x, \quad \forall\, x \in [-1, 1],$$

is called the inverse cosine function.

You can easily see from Figure that the tangent function, in general, is not one-one. However, again if we restrict the domain of f(x) = tan x to the interval]–π/2, π/2[, then the function.

$f :]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow R$ defined by

$f(x) = \tan x, -\frac{\pi}{2}, < x < \frac{\pi}{2}$

is one-one and onto. Hence it has an inverse. The inverse function is called the inverse tangent of x and is denoted by tan –1 x (or by arctan x). In other words,

$y = \tan^{-1} x \Leftrightarrow x = \tan y,$

where $-\frac{\pi}{2} < y < \frac{\pi}{2}$ and $-\infty < x < +\infty$.

Thus, we have the following definition:

Definition 13: Inverse Tangent Function

A function $g : R \rightarrow], -\frac{\pi}{2}, \frac{\pi}{2} [$ defined by

$g(x) = \tan^{-1} x, \ \forall x \in R$

is called the inverse tangent function.

Task Define the inverse cotangent, inverse secant and inverse cosecant function. Specify their domain and range.

Now, before we proceed to define the logarithmic and exponential functions, we need the concept of the monotonic functions. We discuss these functions as follows:

9.3.1 Monotonic Functions

Consider the following functions:

(i) $f(x) = x, \ \forall x \in R.$

(ii) $f(x) = \sin x, \ \forall x \in [-\pi/2, \pi/2].$

(iii) $f(x) = -x^2, \ \forall x \in [0, \infty[,$

(iv) $f(x) = \cos x, \ \forall x \in [0, \pi].$

Out of these functions, (i) and (ii) are such that for any $x_,, x_2$ in their domains,

$x_, < x_2 \Rightarrow f(X_1) \leq f(x_2),$

whereas (iii) and (iv) are such that for any $x_,, x_2$ in their domains,

$x_, < x_2 \Rightarrow f(X_1) \geq f(x_2).$

The functions given in (i) and (ii) are called monotonically increasing while those of (iii) and (iv) are called monotonically decreasing. We define these functions as follows:

Let $f : S \rightarrow R \ (S \subset R)$ be a function

(i) It is said to be a monotonically increasing function on S if

$x_1 < x_, \Rightarrow f(X_1) < f(X_1)$ for any $x_,, x_2 \in S$

Notes

(ii) It is said to be a monotonically decreasing function on S if

$x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2)$ for any $x_{,}, x_2 \in S$.

(iii) The function f is said to be a monotonic function on S if it is either monotonically increasing or monotonically decreasing.

(iv) The function f is said to be strictly increasing on S if

$x < x_2 \Rightarrow f(X_1) < f(x_2)$, for $x_{,}, x_2 \in S$,

(v) It is said to be strictly decreasing on S if

$x_1 < x_2 \Rightarrow f(x_1) > f(X_2)$, for $x_{,}, x_2 \in S$.

You can notice immediately that if f is monotonically increasing then –f i.e. $-f: R \to R$ defined by $(-f)(x) = -f(x), \ \forall x \in R$

is monotonically decreasing and vice-versa.

Example: Test the monotonic character of the function $f: R \to R$ defined as

$$f(x) = \begin{cases} x^2, x \leq 0 \\ -x^2, x > 0 \end{cases}$$

Solution: For any $X_1, x_2 \in R, X_1 \leq 0; X_2 \leq 0$

$X_1 < X_2 \Rightarrow X^2_1 > X^2_2 \Rightarrow f(x_1) > f(X_2)$

which shows that f is strictly decreasing.

Again if $X_1 > 0, X_2 > 0$, then

$X_1 < X_2 \Rightarrow X^2_1 < x^2_2 \Rightarrow -X^2_1 > -X^2_2 \Rightarrow f(X_1) > f(X_2)$

which shows that i is strictly decreasing for $x > 0$. Thus f is strictly decreasing for every $x \in R$.

Now, we discuss an interesting property of a strictly increasing function in the form of the following theorem:

Theorem 1: Prove that a strictly increasing function is always one-one.

Proof: Let $f : S \to T$ be a strictly increasing function. Since f is strictly increasing, therefore,

$X_1 < X_2 \Rightarrow f(x_1) < f(x_2)$ for any $X_1, x_2 \in S$.

Now to show that $f : S \to T$ is one-one, it is enough to show that

$f(x_1) = f(x_2) \Rightarrow X_1 = X_2$.

Equivalently, it is enough to show that distinct elements in S have distinct images in T

i.e. $X_1 \neq X_2 \Rightarrow f(x_1) \neq f(x_2)$, for $X_1, X_2 \in S$.

Indeed,

$x_1 \neq x_2 \Rightarrow x_1 < x_1$ or $x_1 > x_2$

$\Rightarrow f(x_1) < f(x_2)$ or $f(x_1) > f(x_2)$

$\Rightarrow f(x_1) \neq f(x_2)$

which proves the theorem.

Notes

Example: Let $f : S \to T$ be a strictly increasing function such that $f(S) = T$. Then prove that f is invertible and $f^1 : T \to S$ is also strictly increasing.

Solution:

Since $f : S \to T$ is strictly increasing, therefore, f is one-one. Further, since f (S) = T, therefore f is onto. Thus f is one-one and onto. Hence f is invertible. In other words, $f^{-1} : T \to S$ exists.

Now, for any $y_1, y_2 \in T$, we have $y_1 = f(x_1)$, $y_2 - f(x_2)$. If $y_1 < y_2$) then we claim $x_1 < x_2$.

Indeed if $x_1 \geq x_2$, then $f(x_1) \geq f(x_2)$ (why?).

This implies that $y_1 \geq y_2$ which contradicts that $y_1 < y_2$.

Hence $y_1 < y_2 \Rightarrow x_1 < x_2 \Rightarrow f^{-1}(y_1) < f^{-1}(y_2)$

which shows that f^{-1} is strictly increasing.

You can similarly solve the following exercise for a strictly decreasing function:

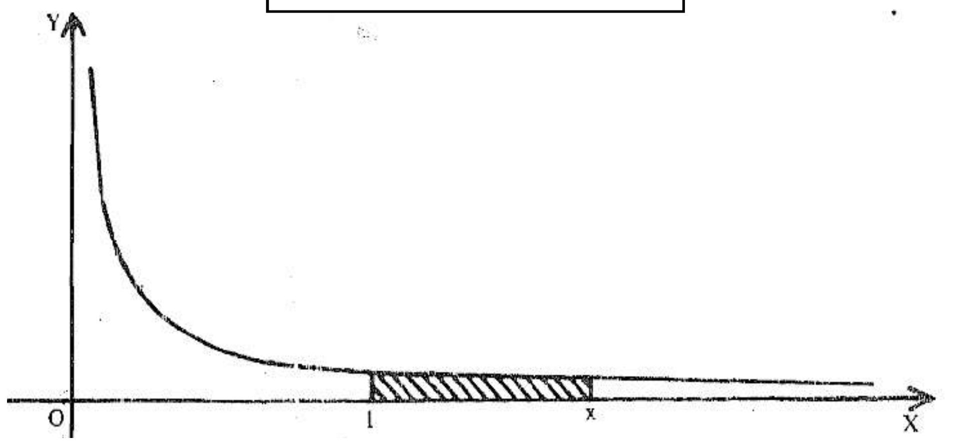
Exercise 3: Let $f : S \to T$ be a strictly decreasing function such that $f(S) = T$. Show that f is invertible and $f^{-1} : T \to S$ is also strictly decreasing.

9.3.2 Logarithmic Function

You know that a definite integral of a function represents the area enclosed between the curve of the function, X-axis and the two Ordinates. You will now see that this can be used to define logarithmic function and then the exponential function.

We consider the function $f(x) = \frac{1}{x}$ for $x > 0$, We find that the graph of the function is as shown in the figure 9.8.

Figure 9.8



Definition 14: Logarithmic Function

For $x \geq 1$, we define thus natural logarithmic function log x as

$$\log x = \int_1^x \frac{1}{t} dt$$

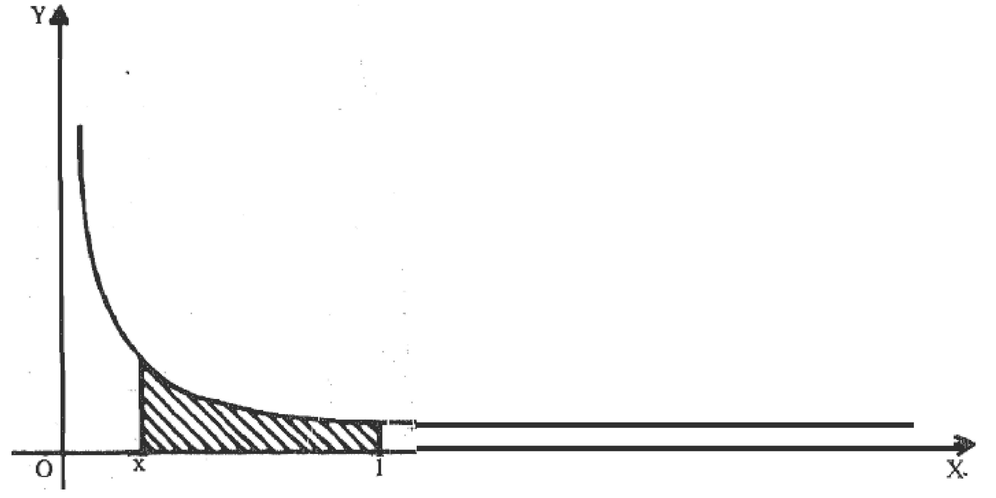
Notes

In the Figure 9.9, log x represents the area between the curve $f(t) = \frac{1}{t}$, x-Axis and the two ordinates at 1 and at x. For 0 < x < 1, we define

$$\log x = \int_x^1 \frac{1}{t} dt$$

This means that for 0 < x < 1. log x is the negative of the area under the graph of $f(t) = \frac{1}{t}$, X-Axis and the two ordinates at x and at 1.

Figure 9.9



We also see by this definition that

$\log x < 0$ if $0 < x < 1$

$\log 1 = 0$

and

$\log x > 0$ if $x > 1$.

It 'also follows by definition that if.

$x_1 > x_2 > 0$, then $\log x_1 > \log x_2$. This shows that log x is strictly increasing. The reason for this is quite clear if we realise by $\log x_1$ as the area under the graph as shown in the Figure 9.10.

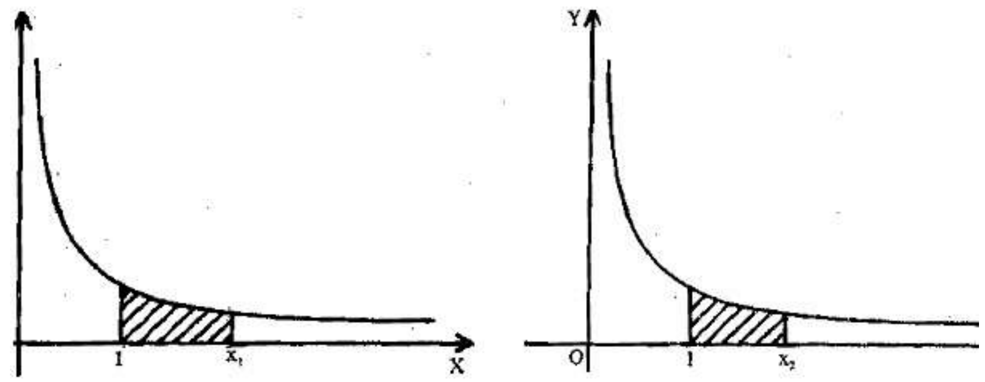
The logarithmic function defined here is called the Natural logarithmic function. For any x > 0, and for any positive real number $a \neq 1$, we can define

$$\log x = \frac{\log x}{\log a}$$

This function is called the logarithmic function with respect to the base a. If a = 10, then this function is called the common logarithmic function.

Notes

Figure 9.10



Logarithmic function to the base a has the following properties:

(i) log, $(x_1\ x_2)$ = log, X_1 + log, X_2

(ii) log a $\left[\dfrac{x_1}{x_2}\right] = \log_a x_1 - \log_a x_2$.

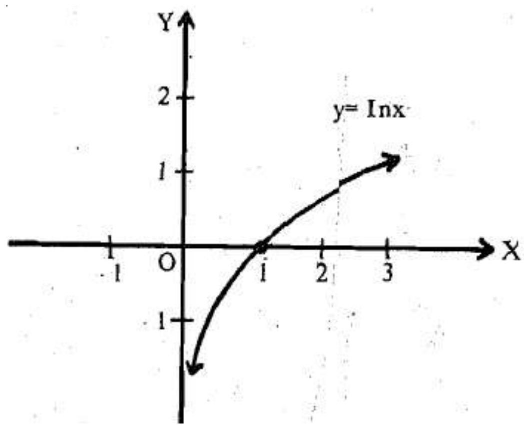
(iii) log, x^m = m log, x for every integer m.

(iv) $\log_a^{\ a} = I$.

(v) $\log_a^{\ 1} = 0$

By the definition of log x, we see that log 1 = 0 and as x becomes larger and larger, the area covered by the curve $f(t) = \dfrac{1}{t}$, X-axis and the ordinates at 1 and x, becomes larger and larger. Its graph is as shown in the Figure 9.11,

Figure 9.11



You already know what is meant, by inverse of s function. You had also seen in Unit 1 that if f is 1 – 1 and onto, then f is invertible. Let us apply that study to logarithmic function.

Notes

9.3.3 Exponential Function

We now come to define exponential function. We have seen that

$\log x :]0, \infty[\rightarrow R$

is strictly increasing function. The graph of the logarithmic function also shows that

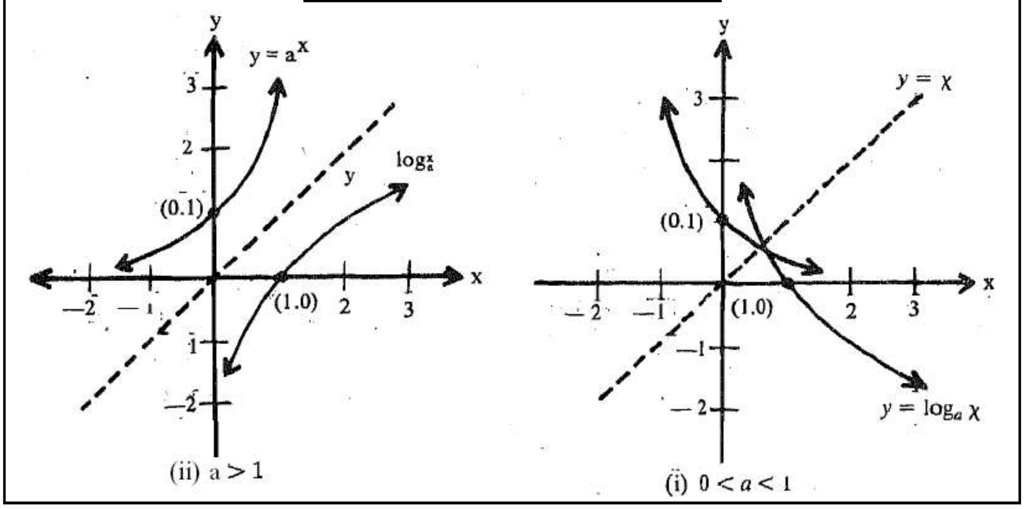
$\log x :]0, \infty[\rightarrow R$

is also onto. Therefore this function admits of inverse function. Its inverse function, called the Exponential function, Exp (x) has domain as the set R of all real numbers and range as $]0, \infty[$. If

$\log x = y$, then $\text{Exp}(y) = x$.

The graph of this function is the mirror image of logarithmic function as shown in the Figure 9.12.

Figure 9.12



The Exp (x) satisfies the following properties:

(i) E^p (x + y) = Exp x Exp y

(ii) Exp (x – y) = Exp x/Exp y

(iii) $(\text{Exp } x)^n = \text{Exp}(nx)$

(iv) Exp(0) = 1

We now come to define a^x for $a > 0$ and x any real number. We write

$a^x = \text{Exp}(x \log a)$

If x is any rational number, then we know that $\log a^x = x \log a$. Hence

$\text{Exp}(x \log a) – \text{Exp}(\log a^x) = a^x$. Thus our definition agrees with the already known definition of a in case x is a rational number. The function a^x satisfies the following properties

(i) $a^x a^y = a^{x+y}$

(ii) $\frac{a^x}{a^y} = a^{x-y}$

Notes

(iii) $(a^x)^y = a^{xY}$

(iv) $a^x b^x - (ab)^x$, a > 0, b > 0.

Denote E (I) = e, so that log e = 1. The number e is an irrational number and its approximation say up to five places of decimals is 2.71828. Thus

$$e^x = \text{Exp}(x \log e) = \text{Exp}(x).$$

Thus Exp (x) is also denoted as e^x and we write for each a > 0, $a^x = e^x \log a$

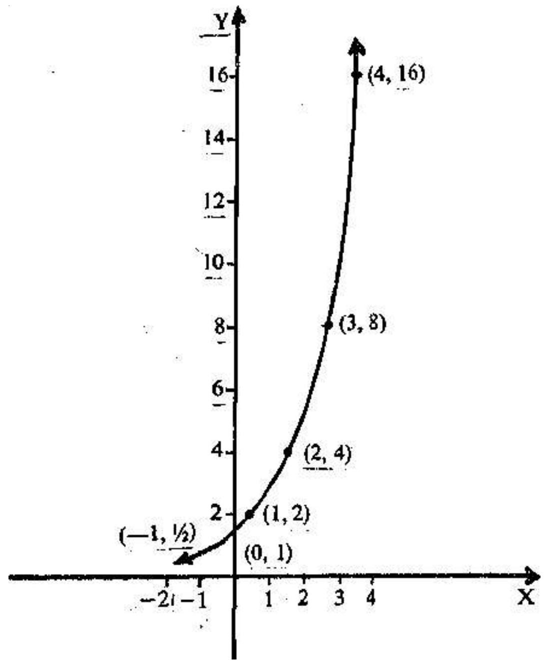
Example: Plot the graph of the function I : R → R defined by $f(x) = 2^x$.

Solution:

x	–2	–1	0	1	2
2^x	$\frac{1}{4}$	$\frac{1}{2}$	1	2	4

The required graph takes the shape as shown in the Figure 9.13.

Figure 9.13



9.4 Some Special Functions

So far, we have discussed two main classes of real functions – Algebraic and Transcendental. Some functions have been designated as special functions because of their special nature and behaviour. Some of these special functions are of great interest to us. We shall frequently use these functions in our discussion in the subsequent units and blocks.

9.4.1 Identity Function

We have already discussed some of the special functions. For example, the Identity function i : R → R, defined as i (x) = x, ∀ x ER has already been discussed as an algebraic function.

Notes

However, this function is sometimes, referred to as a special function because of its special characteristics, which are as follows:

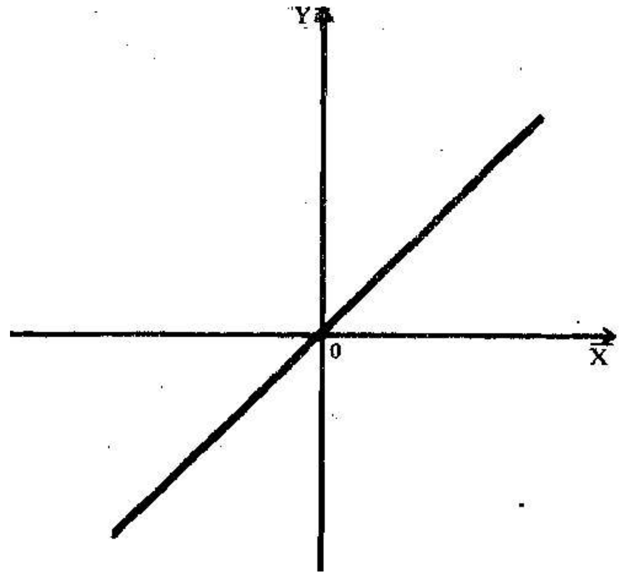
(i) Domain of i = Range of i = Codomain of i

(ii) The function i is one-one and onto. Hence it has an inverse i^{-1} which is also one-one and onto.

(iii) The function i is invertible

(iv) The graph of the identity function is a straight line through the origin which forms an angle of 45° along the positive direction of X-axis as shown in the Figure 9.14.

Figure 9.14



9.4.2 Periodic Function

You know that

$\sin(2\pi + x) = \sin(4\pi + x) = \sin x,$

$\tan(\pi + x) = \tan(2\pi + x) = \tan x.$

This leads us to define a special class of functions, known as Periodic functions. All trigonometric functions belong to this class.

A function $f : S \to R$ is said to be periodic if there exists a positive real number k such that

$f(x + k) = f(x),\ \forall x \in S$

where S C R.

The smallest such positive number k is called the period of the function.

You can verify that the functions sine, cosine, secant and cosecant are periodic each with a period $2n\pi$ while tangent and cotangent are periodic functions each with a period.

9.4.3 Modulus Function

The modulus or the absolute (numerical) value of a real number has already been defined in Unit 1. Here we define the modulus (absolute value) function as follows:

Let S be a subset of R. A function $f : S \to R$ defined by

$$f(x) = |x|, \ \forall x \in S$$

is called the modulus function.

In short, it is written as Mod function.

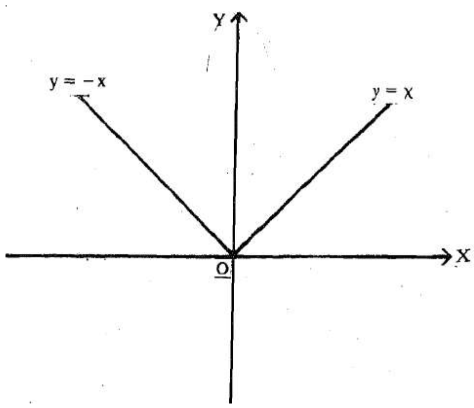
You can easily see the following properties of this function:

(i) The domain of the Modulus function may be a subset of R or the set R itself.

(ii) The range of this function is a subset of the set of non-negative real numbers.

(iii) The Modulus function $f : R \to R$ is not an onto function (Check why?).

Figure 9.15



(iv) The Modulus function $f : R \to R$ is not one-one. For instance, both 2 and –2 in the domain have the same image 2 in the range.

(v) The modulus function $f : R \to R$ does not have an inverse function (why)?

(vi) The graph of the Modulus function is $R \to R$ given in the Figure 9.15.

It consists of two straight lines:

(i) $y = x \ (y \geq 0)$

and (ii) $y = -x \ (y \geq 0)$

through 0, the origin, making an angle of $\pi/4$ and $3\pi/4$ with the positive direction of X-axis:

9.4.4 Signum Function

A function $f : R \to R$ defined by

$$f(x) = \begin{cases} \dfrac{|x|}{x} & \text{when } x \neq 0 \\ 0 \text{ if } x = 0 \end{cases}$$

Notes

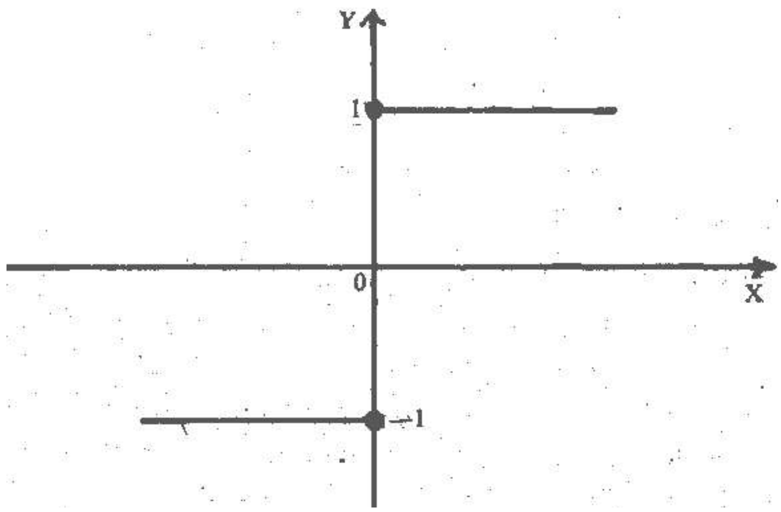
or equivalently by:

$$f(x) = \begin{cases} -1 \text{ if } x < 0 \\ 0 \text{ if } x = 0 \\ 1 \text{ if } x > 0 \end{cases}$$

is called the signum function. It is generally written as sgn (x).

Its range set is {–1, 0, 1}. Obviously sgn x is neither one-one nor onto. The graph of sgn x is shown in the Figure 9.16.

Figure 9.16



9.4.5 Greatest Integer Function

Consider the number 4.01. Can you find the greatest integer which is less than or equal to this number? Obviously, the required integer is 4 and we write it as [4.01] = 4.

Similarly, if the symbol [x] denotes the greatest integer contained in x then we have

[3/4] = 0, [5.01] –5,

[–.005] = –1 and [–3.96] = –4.

Based on these, the greatest integer function is defined as follows:

A function $f : R \to R$ defined by

$f(x) = [x], \ \forall x \in R,$

where [x] is the largest integer less than or equal to x is called the greatest integer function.

The following properties of this function are quite obvious:

(i) The domain is R and the range is the set Z of all integers.

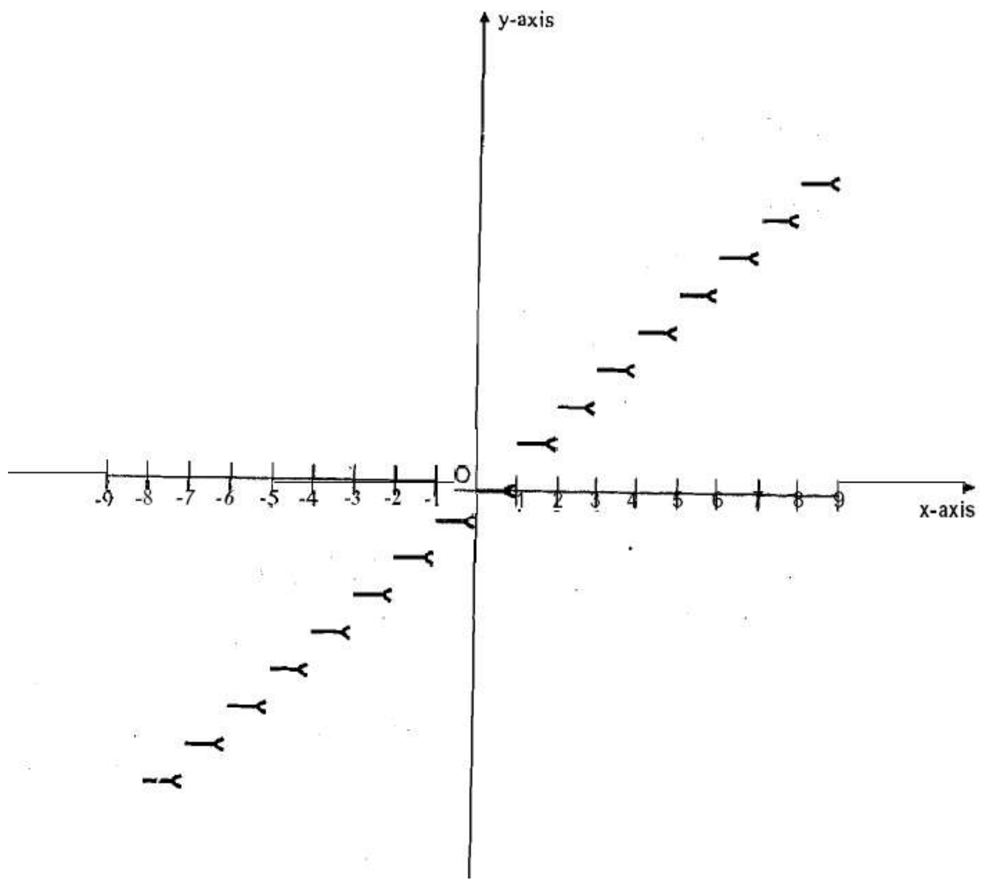
(ii) The function is neither one-one nor onto.

(iii) If n is any integer and x is any real number such that x is greater than or equal to n but less than n + 1 i.e., if $n \le x < n + 1$ (for some integer n), then [x] = n i.e.,

Notes

The graph of the greatest integer function is shown in the Figure 9.17.

Figure 9.17



Example: Prove that

$[x + m] = [x] + m, \ \forall \ x \in R$ and $m \in Z$,

Solution: You know that for every $x \in R$, there exists an integer n such that

$n < x < n + 1.$

Therefore,

$n + m < \ + m < n + 1 + m,$

and hence

$[x + m] = n + m = [x] + m,$

which proves the result.

Task Test whether or not the function $f : R \rightarrow R$ defined by $f(x) = x - [x] \ \forall \ x \in R$, is periodic. If it is so, find its period.

Notes

9.4.6 Even and Odd Functions

Consider a function f : R → R defined as

$f(x) = 2x, \ \forall x \in R.$

If you change x to –x, then you have

$f(-x) = 2(-x) = -2x - f(x).$

Such a function is called an odd function.

Now, consider a function f : R → R defined as

$f(x) = x^2 \ \forall x \in R$

Then changing x to –x we get

$q - x) = (-x)^2 = x^2 = f(x)$

Such a function is called an even function.

The definitions of even and odd functions are as follows:

A function f : R → R is called even if $f(-x) = f(x), \ \forall x \in R$,

It is called odd if $f(-x) = -f(x), \ \forall x \in R$

Example: Verify whether the function f : R ∈ R defined by

(i) $f(x) = \text{Sin}^2 x + \text{Cos}^3 2x$

(ii) $f(x) = \sqrt{a^2 + ax + x^2} - \sqrt{a^2 - ax + x^2}$ are even or odd.

Solution:

(i) $f(x) + \sin^2 x + \cos^3 2x, \forall x \in R$

$\Rightarrow \sin^2(-x) + \cos^3 2(-x)$

$\sin^2 x + \cos^3 2x = f(x), \ \forall x \in R$

$\Rightarrow$ f is an even function.

(ii) $f(x) = \sqrt{a^2 - ax + x^2} - \sqrt{a^2 - ax + x^2}, \forall x \in R$

$\Rightarrow f(-x) = \sqrt{a^2 - ax + x^2} - \sqrt{a^2 + ax + x^2}$

$= -(x), \ \forall x \in R$

$\Rightarrow$ f is an odd function.

Task Determine which of the following functions are even or odd or neither:

(i) $f(x) = x$

(ii) A constant function

(iii) sin x, cos x, tan x,

(iv) $f(x) = \dfrac{x-4}{x^2-9} = -(x), \ \forall x \in R$

9.4.7 Bounded Functions

Notes

In Unit 2, you were introduced to the notion of a bounded set, upper and lower bounds of a set. Let us now extend these notions to a function.

You know that if $f : S \rightarrow R$ is a function, (S C R), then

$\{f(x) : x \in S)\}$

is called the range set or simply the range of the function f.

A function is said to be bounded if its range is bounded.

Let $f : S \rightarrow R$ be a function. It is said to be bounded above if there exists a real number K such that

$f(x) \leq k \quad \forall x \in S$

The number K is called an upper bound of it. The function f is said to be bounded below if there exists a number k such that

$f(x) \geq k \quad \forall x \in S$

The number k (is called a lower bound of f).

A function $f : S \rightarrow R$, which is bounded above as well as bounded below, is said to be bounded. This implies that there exist two real numbers k and K such that

$k \leq f(x) \leq K \quad \forall x \in S.$

This is equivalent to say that a function $f : S \rightarrow R$ is bounded if there exists a real number M such that

$|f(x)| \leq M, \ \forall x \in S.$

A function may be bounded above only or may be bounded below only or neither bounded above nor bounded below.

Recall that sin x and cos x are both bounded functions. Can you say why? It is because of the reason that the range of each of these functions is [–1, 1].

Example: A function $f : R \rightarrow R$ defined by

(i) $f(x) = -x^2, \ \forall x \in R$ is bounded above with 0 as an upper bound

(ii) $f(x) = x, \ \forall x \geq 0$ is bounded below with 0 as a lower bound

(iii) f(x) = for |x| –l is bounded because |f(x)| ≤ 1 for |x|≤|.

Self Assessment

1. Test whether the following are rational numbers:

   (i) √I7 (ii) √8 (iii) √3 + √2

2. The inequality $x^2 - 5x + 6 < 0$ holds for

   (i) $x < 2, x < 3$ (ii) $x > 2, x < 3$

   (iii) $x < 2, x > 3$ (iv) $x > 2, x > 3$

Notes

3. Test whether the following statements are true or false:
   (i) The set Z of integers is not a NBD of any of its points.
   (ii) The interval]0, 1] is a NBD of each of its points
   (iii) The set]1, 3[$\cup$] 4, 5[is open.
   (iv) The set [a, ∞[$\cup$] $-\infty$, a] is not open.
   (v) N is a closed set.
   (vi) The derived set of Z is non-empty.
   (vii) Every real number is a limit point of the set Q of rational numbers.
   (viii) A finite bounded set has a limit point.
   (ix) [4, 5] $\cup$ [7, 8] is a closed set.
   (x) Every infinite set is closed.

9.5 Summary

- In this unit, we have discussed various types of real functions. We shall frequently use these functions in the concepts and examples to be discussed in the subsequent units throughout the course.
- We have introduced the notion of an algebraic function and its various types. A function f: $S \to R$ ($S \subset R$) defined as y = f(x), $\forall x \in S$ is said to be algebraic if it satisfies identically an equation of the form

  $p_0(x)\, y^n + p_1(x)\, y^{n-1} + p_2(x)\, y^{n-2} + ... + p(x)\, y + p_n(x) = 0,$
- where $p_0(x)$, $p_1(x)$, $p_n(x)$ are polynomials in x for all $x \in S$ and n is a positive integer. In fact, any function constructed by a finite number of algebraic operations—addition, subtraction, multiplication, division and root extraction – is an algebraic function. Some of the examples of algebraic functions are the polynomial functions, rational functions and irrational functions.
- But not all functions are algebraic. The functions which are not algebraic, are called transcendental functions. Some important examples of the transcendental functions are trigonometric functions, logarithmic functions and exponential functions which have been defined in this section. We have defined the monotonic functions also in this section.
- We have discussed some special functions. These are the identity function, the periodic functions, the modulus function, the signum function, the greatest integer function, even and odd functions. Lastly, we have introduced the bounded functions and discussed a few examples.

9.6 Keywords

Upper Bound: Let f : $S \to R$ be a function. It is said to be bounded above if there exists a real number K such that $f(x) \leq k \ \ \forall x \in S$. The number K is called an upper bound of it.

Lower Bound: The function f is said to be bounded below if there exists a number k such that $f(x) \geq k\ x \in S$. The number k is called a lower bound of f.

Cotangent Function: A function f : $S \to R$ defined.

9.7 Review Questions

Notes

1. Show the graph of $f : R \to R$ defined by $F(x) = (\frac{1}{2})^x$

2. Find the period of the function f where $f(x) = |\sin^3 x|$

3. Test which of the following functions with domain and co-domain as R are bounded and unbounded:

   (i) $f(x) = \tan x$

   (ii) $f(x) = [x]$

   (iii) $f(x) = e^x$

   (iv) $f(x) = \log x$

4. Suppose $t : S \to R$ and $g : S \to R$ are any bounded functions on S. Prove that f + g and f. g are also bounded functions on S.

5. If a, b, c, d are real numbers such that

   $$a^2 + b^2 = I, c^2 + d^2 = 1,$$

   then show that $ac + bd \leq 1$.

6. Prove that $|a + b + c| \leq |a|+|b|+|c|$ for all a, b, c, $\in$ R.

7. Show that

   $|a_1 + a_2 + ... + a_n| \leq |a_1|+|a_2| + ... +|a_n|$ for $a_1, a_2,.... a, \in R$.

8. Find which of the sets in question 8 are bounded below. Write the infemum if it exists.

9. Which of the sets in question 8 are bounded and unbounded.

10. Justify the following statements:

    (i) The identity function is an odd function.

    (ii) The absolute value function is an even function.

    (iii) The greatest integer function is not onto.

    (iv) The tangent function is periodic with period π.

    (v) The function $f(x) = |x|$ for $-2 \leq x \leq 3$ is bounded.

    (vi) The function $f(x) = e^x$ is not bounded

    (vii) The function $f(x) = \sin x$, for $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is monotonically increasing.

    (viii) The function $f(x) = \cos x$ for $0 \leq x \leq \pi$ is monotonically decreasing.

    (ix) The function $f(x) = \tan x$ is strictly increasing for $x \in [0, \frac{\pi}{2}]$.

    (x) $f(x) = \sqrt{\frac{2x^2 - 3x + 2}{3x - 2}}$ is an algebraic function.

Notes

Answers: Self Assessment

1. None is a rational number
2. For (ii) only since $2 < x < 3$.
3. (i) True
   (ii) False
   (iii) True
   (iv) False
   (v) True
   (vi) False
   (vii) True
   (viii) False
   (ix) True
   (x) False

9.8 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 10: Limit of a Function

Notes

CONTENTS

Objectives

Introduction

10.1 Notion of Limit

10.2 Sequential Limits

10.3 Algebra of Limits

10.4 Summary

10.5 Keywords

10.6 Review Questions

10.7 Further Readings

Objectives

After studying this unit, you will be able to:

- Define limit of a function at a point and find its value
- Know sequential approach to limit of a function
- Find the limit of sum, difference, product and quotient of functions

Introduction

In earlier unit, we dealt with sequences and their limits. As you know, sequences are functions whose domain is the set of natural numbers. In this unit, we discuss the limiting process for the real functions with domains as subsets of the set R of real numbers and range also a subset of R. What is the precise meaning for the intuitive idea of the values f(x) of a function f tending to or approaching a number A as x approaches the number a? The search for an answer to this question shall enable you to understand the concept of the limit which you have used in calculus. The effect of algebraic operations of addition, subtraction, multiplication and division on the limits of functions.

10.1 Notion of Limit

The intuitive idea of limit was used both by Newton and Leibnitz in their independent invention of Differential Calculus around 1675. Later this notion of limit was also developed by D'Alembert. "When the successive values attributed to a variable approach indefinitely a fixed value so as to end by differing from it by as little as one wishes, this last is called the limit of all the others."

Consider a simple example in which a function f is defined as

$$f(x) = 2x + 3, \quad \forall x \in R, x \neq 1.$$

Give x the values which are near to 1 in the following way:

When $\qquad x = 1.5, 1.4, 1.3, 1.2, 1.1, 1.01, 1.001$

Notes

$$f(x) = 6, 5.8, 5.6, 5.4, 5.2, 5.02, 5.002$$

When $x = .5, .6, .7, .8, .9, .99, .999$

$$f(x) = 4, 4.2, 4.4, 4.6, 4.8, 4.98, 4.998$$

You can form a table for these values as follows:

X	.5	.6	.7	.8	.9	.99	.999	1.001	1.01	1.1	1.2	1.3	1.4	1.5
f(x)	4	4.2	4.4	4.6	4.8	4.98	4.998	5.002	5.02	5.2	5.4	5.6	5.8	6

The limit of a function f at a point a is meaningful only if a is a limit point of its domain. That is, the condition $f(x) \to \alpha$ as $n \to a$ would make sense only when there does not exists a nbd. U of a for which the set U n Dom (f)\{a} is empty i.e., $a \in$ (Dom 0))′.

You see that as the values of x approach 1, the values of f(x) approach 5. This is expressed by saying that limit of f(x) is 5 as x approaches 1. You may note that when we consider the limit of f(x) as x approaches 1, we do not consider the value of f(x) at x = 1.

Thus, in general, we can say as follows:

Let f be a real function defined in a neighbourhood of a point x = a except possibly at a. Suppose that as x approaches a, the values taken by f approach more and more closely a value A. In other words, suppose that the numerical difference between A and the values taken by f can be made as small as we please by taking values of x sufficiently close to a. Then we say that f tends to the limit A as x tends to a. We write

$$f(x) \to A \text{ as } x \to a \text{ or } \lim_{x\to a} f(x) = A.$$

This intuitive idea of the limit of a function can be expressed mathematically as formulated by the German mathematician Karl Veierstrass in the 18th Century. Thus, we have the following definition:

Definition 1: Limit of a Function

Let a function f be defined in a neighbourhood of a point ‘a’ except possibly at ‘a’. The function f is said to tend to or approach a number A as x tends to or approaches a number ‘a’ if for any $\varepsilon > 0 >$ there exists a number $\delta > 0$ such that

$$|f(x) - A| < \varepsilon \text{ for } 0 < |x - a| < 6.$$

We write it as $\lim_{x\to a} f(x) = A$. You may note that

$$|f(x) - A| < \varepsilon \text{ for } 0 < |x - a| < 6.$$

can be equivalently written as

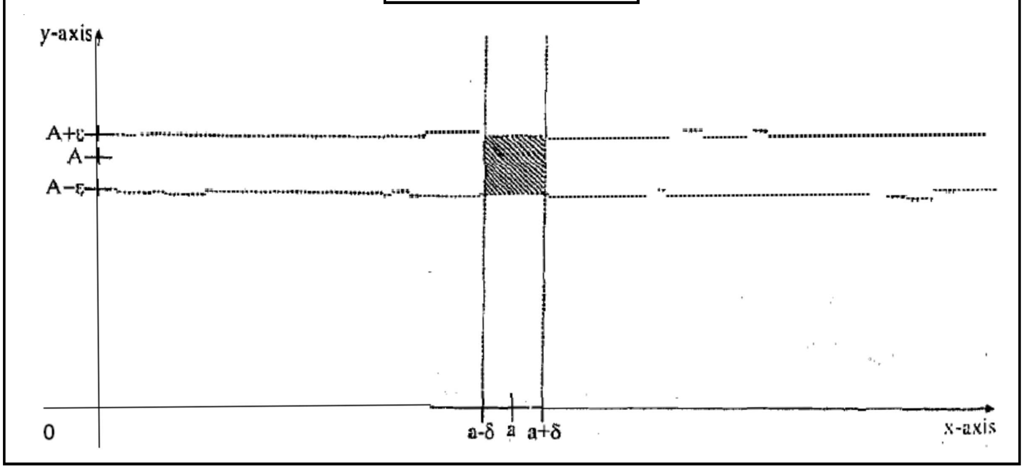
$$f(x) \in]A - \varepsilon, A + \varepsilon[\text{ for } x \in]a - 6, a + \delta[\text{ and } \neq a.$$

Geometrically, the above definition says that, for strip S_A of any given width around the point A, if it is possible to find a strip S_a of some width around the point a such that the values that f(x) takes, for x in the strip S_a $(x \neq a)$, lies in the shaded box formed by the intersection of strips S_A and S_a, then $\lim_{x\to a} f(x) = A$.

This is shown geometrically in Figure 10.1 below. The inequality $0 < |x - a| < 6$ determines the interval $]a - 6, a + \delta[$ minus the point ‘a’ along the x-axis and the inequality $|f(x) - A| < \varepsilon$ determines the interval $]A - \varepsilon, A + \varepsilon[$ along the y-axis.

Notes

Figure 10.1



Example: Let a function f: R → R be defined as

$$f(x) = x^2, \ \forall \ x \in R.$$

Find its limits when x → 2.

Solution: By intuition, it follows that

$$\lim_{x \to 2} f(x) = \lim_{x \to 2} x^2 = 4.$$

In other words, we have to show that for a given E > 0, there exists a δ > 0 such that

$$0 < |x - 2| < \delta \Rightarrow |f(x) - 4| < \varepsilon.$$

Suppose that an ε > 0 is fixed. Then consider the quantity $|f(x) - 4|$, which we can write as

$$|f(x) - 4| = |x^2 - 4| = |(x - 2)(x + 2)|.$$

Note that the term $|x - 2|$ is exactly the same that appears in the 6–inequality in the definition. Therefore, this term should be less than δ. In other words,

$$|x - 2| < \delta$$

$$\Rightarrow 2 - \delta < x < 2 + \delta$$

$$\Rightarrow x \in]\, 2 - \delta, 2 + \delta[.$$

We restrict δ to a value 2 so that x lies in the interval $]\, 2 - \delta, 2 + \delta[\subset]\, 0, 4[$. Accordingly, then $|x + 2| < \delta$. Thus, if $\delta \leq 2$, then

$$|x - 2| < 2 \Rightarrow 0 < |x + 2| < \delta,$$

and further that

$$|x - 2| < \delta \leq 2 \Rightarrow |x + 2| \; |x - 2| < \delta \; |x - 2| < 6\delta.$$

If 6 is small then so is 6δ. In fact it can be made less than ε by choosing δ suitably. Let us, therefore, select 6 such that δ = min.(2, ε/6). Then

$$0 < |x - 2| < \delta \Rightarrow |f(x) - 4| < \delta \; |x - 2| < \delta.\ \delta \leq 6.\ \varepsilon/\delta = \varepsilon.$$

Notes

This completes the solution.

Note that the first step is to manipulate the term $|f(x) - A|$ by using algebra. The second step is to use a suitable strategy to manipulate $|f(x) - A|$ into the form

$$|x - a| \text{ (trash)}$$

where the 'trash' is some expression which has the property that: it is bounded provided that 6 is sufficiently small. Why we use the term 'trash, for the expression as a multiple of $|x - a|$? The reason is that once we know that it is bounded, we can replace it by a number and forget about it.

The number 6 arose by virtue of this 'trash'. If you take $6 \leq 3$ (instead of $6 \leq 2$), you can still show that 6 will be replaced by 7. In that case you can set δ as

$$\delta = \min.(3, \varepsilon/7)$$

and the proof will be complete. Thus, there is nothing special about 6. The only thing is that such a number (whether 6 or 7) has to be evaluated by the restriction placed on 6.

Finally, note that in general, 6 will depend upon ε.

Task For a function f: R → R defined by $f(x) = x^2$, find its limit when x tends to 1 by the ε – δ approach.

In Unit 5, we proved that a convergent sequence cannot have more than one limit. In the same way, a function cannot have more than one limit at a single point of its domain. We prove it in the following theorem:

Theorem 1: If $\lim_{x \to a} f(x) = A$, $\lim_{x \to a} f(x) = B$, then A = B.

Proof: In short, we have to show that if $\lim_{x \to a} f(x)$ has two values say A and B, then A = B. Since $\lim_{x \to a} f(x) = A$, $\lim_{x \to a} f(x) = B$, given a number E > 0, there exists numbers $\delta_1, \delta_2 > 0$ such that

$$|f(x) - A| < \varepsilon/2 \text{ whenever } 0 < |x - a| < \delta_1$$

and

$$|f(x) - B| < \delta/2 \text{ whenever } 0 < |x - a| < \delta_2.$$

If we take δ equal to minimum of δ_1 and δ_2, then we have

$$|f(x) - A| < \varepsilon/2 \text{ and } |f(x) - B| < \varepsilon/2 \text{ whenever } 0 < |x - a| < \delta.$$

Choose an x_0 such that $0 < |x_0 - a| < \delta$. Then

$$|A - B| = (A - f(x_0) + f(x_0) - B| \leq (A - f(x_0)| + |f(x_0) - B|$$

$$< \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

E is arbitrary while A and B are fixed. Hence $|A - B|$ is less than every positive number ε which implies that $|A - B| = 0$ and hence A = B. (For otherwise, if A ≠ B then A – B = C ≠ 0 (say). We can choose $\varepsilon < |C|$ which will be a contradiction to the fact that $|A - B| < \varepsilon$ for every ε > 0.)

In the example considered before defining limit of a function, we allowed x to assume values both greater and smaller than 1. If we consider values of x greater than 1 that is on the right of 1, we see that values of f(x) approaches 5. We say that f(x) tends to 5 as x tends to 1 from the right.

Similarly you see that values of f(x) approach 5 as x tends to 1 from the left i.e. through values smaller than 1. This leads us to define right hand and left hand limits as under:

Definition 2: Right Hand Limits and Left Hand Limits

Let a function f be defined in a neighbourhood of a point 'a' except possibly at 'a'. It is said to tend to a number A as x tends to a number 'a' from the right or through values greater than 'a' if given a number ε > 0, there exists a number δ > 0 such that

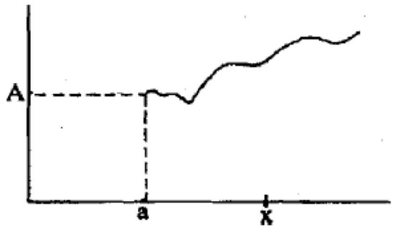
$$|f(x) - A| < \varepsilon \text{ for } a < x < a + \delta.$$

We write, it as

$$\lim_{x \to a+} f(x) = A \text{ or } \lim_{x \to a+0} f(x) = A \text{ or } f(a+) = A.$$

See Figure 10.2(a).

Figure 10.2(a)



The function f is said to tend to a number A as x tends to 'a' from the left or through values smaller than 'a' if given a number E > 0, there exists a number δ > 0 such that

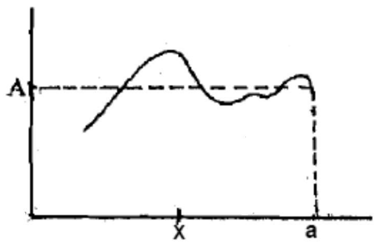
$$|f(x) - A| < \varepsilon \text{ for } a - 6 < x < a.$$

We write it as

$$\lim_{x \to a-} f(x) = A \text{ or } \lim_{x \to a-0} f(x) = A \text{ or } f(a-) = A.$$

See Figure 10.2(b).

Figure 10.2(b)



Since the definition of limit of a function employs only values of x different from 'a' it is totally immaterial what the value of the function is at x = a or whether f is defined at x = a at all. Also it is obvious that $\lim_{x \to a} f(x) = A$ if and only if f(a+) = A, f(a–) = A.

This we prove in the next theorem. First we consider the following example to illustrate it.

Example: Find the limit of the function f defined by

$$f(x) = \frac{x^2 + 3x}{2x} \text{ for } x \neq 0$$

Notes

when $x \to 0$.

Solution: The given function is not defined at x = 0 since $f(0) = \frac{0}{0}$ which is meaningless.

If $x \neq 0$, then $f(x) = \frac{x+3}{2}$. Therefore

$$\text{Right Hand Limit} = \lim_{x \to 0+0} f(x)$$

$$= \lim_{h \to 0} \frac{(0+h)+3}{2} \quad (h > 0)$$

$$= 3/2.$$

$$\text{Left Hand Limit} = \lim_{x \to 0-0} f(x)$$

$$= \lim_{h \to 0} f(x) = \frac{(0-h)+3}{2} \quad (h > 0)$$

$$= 3/2.$$

Since both the right hand and left hand limits exist and are equal,

$$\lim_{x \to 0} f(x) = 3/2.$$

We, now, discuss the theorem concerning the existence of limit and that of right and the left hand limits.

Theorem 2: Let f be a real function. Then

$$\lim_{x \to a} f(x) = A \text{ if and only if } \lim_{x \to a+} f(x) \text{ and } \lim_{x \to a-} f(x)$$

both exist and are equal to A.

Proof: If $\lim_{x \to a+} f(x) = A$, then corresponding to any $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - A| < \varepsilon \text{ whenever } 0 < |x - a| < \delta$$

i.e., $|f(x) - A| < \varepsilon$ whenever $a - \delta < x < a + \delta$, $x \neq a$

This implies that $|f(x) - A| < \varepsilon$ whenever $a - \delta < x < a$

and $|f(x) - A| < \varepsilon$ whenever $a < x < a + \delta$.

Hence both the left hand and right hand limits exist and are equal to A. Conversely, if f(a+) and f(a–) exist and are equal to A say, then corresponding to E > 0, there exist positive numbers δ_1 and δ_2 such that

$$|f(x) - A| < E \text{ whenever } a < x < a + \delta_1$$

and

$$|f(x) - A| < \varepsilon \text{ whenever } a - \delta_2 < x < a.$$

Let 6 be the minimum of δ_1 and δ_2. Then

$$|f(x) - A| < E \text{ whenever } a - \delta < x < a + \delta, x \neq a$$

i.e. $|f(x) - A| < \varepsilon$ if $0 < |x - a| < \delta$

which proves that

$\lim_{x \to a} f(x)$ exists and $\lim_{x \to a} f(x) = A$.

Notes

Example: Consider the function I defined by

$$f(x) = \frac{x^2 - 1}{x - 1}, x \in R, x \neq 1$$

Find its limit as $x \to 1$.

Solution: Note that f(x) is not defined at x = 1. (Why?).

For any $x \neq 1$, $$f(x) = \frac{x^2 - 1}{x - 1} = x + 1.$$

$$\lim_{x-1+} f(x) = \lim_{x-1+} (x + 1) = 2$$

$$\lim_{x-1-} f(x) = \lim_{x-1-} (x + 1) = 2$$

Since $$\lim_{x-1+} f(x) = \lim_{x-1-} f(x), \text{ by Theorem 2, } \lim_{x-1} f(x) = 2,$$

$\lim_{x-1} f(x) = 2$ can be seen by $\varepsilon - \delta$ definition as follows:

Corresponding to any number E > 0, we can choose 8 = ε itself. Then, it is clear that

$$0 < |x - 1| < \delta = \varepsilon \Rightarrow$$

$$|f(x) - 2| = \left|\frac{x^2 - 1}{x - 1} - 2\right| = |x + 1 - 2| = |x - 1| < \varepsilon.$$

From Theorem 2, it follows that f(l+) and f(l–) also exist and are both equal to 2.

Example: Let f: R → R be defined as

$$f(x) = \begin{cases} |x|, x \neq 0 \\ 3, x \approx 0. \end{cases}$$

Find its limit when $x \to 0$.

Solution: You are familiar with the graph of f as given in Unit 4. It is easy to see that $\lim_{x-0} f(x) = 0 = f(0+) = f(0-)$. The fact that f(0) = 3 has neither any bearing on the existence of the limit of f(x) as x tends to 0 nor on the value of the $\lim_{x-0} f(x)$.

Example: Define f on the whole of the real L in ε as follows:

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

Find its limit when x tends to 0.

Solution: Since f(x) = 1 for all x > 0,

$$f(0+) = \lim_{x \to 0+} f(x) = +1.$$

Similarly f(0–) = –1. Since $\lim_{x \to 1+} f(x) \neq \lim_{x \to 1-} f(x)$, $\lim_{x \to 0} f(x)$ does not exist.

Now, we give another proof using ε – S definition.

Notes

For, if $\lim_{x \to 0} f(x) = A$, then for a given $\varepsilon > 0$, there must exist some $\delta > 0$, such that $|f(x) - A| < \varepsilon$. Let us choose $x_1 > 0$, $x_2 < 0$ such that $|x_1| < S$ and $|x_2| < \delta$. Then

$$\begin{aligned} 2 &= |f(x_1) - f(x_2)| \\ &\le |f(x_1) - A| + |A - f(x_2)| \\ &< 2\varepsilon, \qquad \text{(because } |x_1 - 0| < \delta \text{ and } |x_2 - 0| < \delta) \end{aligned}$$

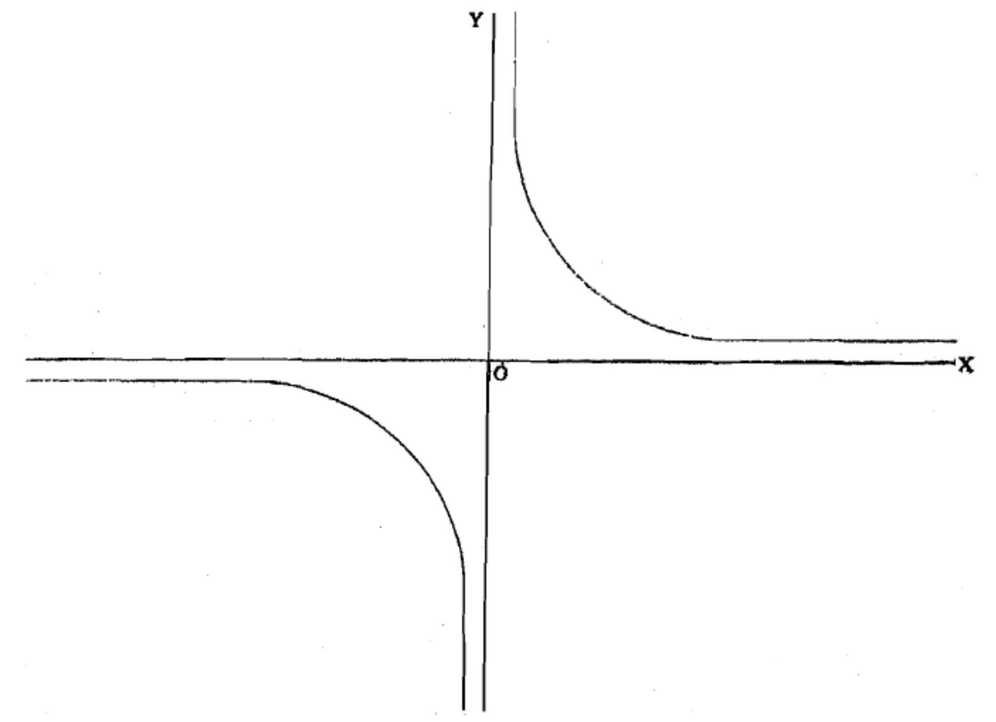
for every ε which is clearly impossible if $\varepsilon < 1$. Non-existence of $\lim_{x \to 0} f(x)$ also follows from Theorem 2, since $f(0+) \neq f(0-)$.

The above example shows clearly that the existence of both $f(a+)$ and $f(a-)$ alone is not sufficient for the existence of $\lim_{x \to 0} f(x)$. In fact, for $\lim_{x \to a} f(x)$ to exist, they both should be equal.

Now consider, the function f defined by $f(x) = \frac{1}{x}$ for $x \neq 0$.

The graph of f looks as shown in the Figure 10.3. You know that it is a rectangular hyperbola. Here none of the $\lim_{x \to 0+} f(x)$ and $\lim_{x \to 0+} f(x)$ exists. Hence $\lim_{x \to 0} f(x)$ does not exist.

Figure 10.3



This can be easily seen from the fact that $1/x$ becomes very large numerically as x approaches 0 either from the left or from the right. If x is positive and takes up larger and larger values, then values of $1/x$ i.e. $f(x)$ is positive and becomes smaller and smaller. This is expressed by saying that $f(x)$ approaches 0 as x tends to ∞. Similarly if x, is negative and numerically takes up larger and larger values, the values of $f(x)$ is negative and numerically becomes smaller and smaller and we say that $f(x)$ approaches 0 as x tends to $-\infty$. These two observations are related to the notion of the limit of a function at infinity.

Let us now discuss the behaviour of a function f when x tends to ∞.

Notes

Let a function f be defined for all values of x greater than a fixed number c. That is to say that f is defined for all sufficiently large values of x. Suppose that as x increases indefinitely, f(x) takes a succession of values which approach more and more closely a value A. Further suppose that the numerical difference between A and the values f(x) taken by the function can be made as small as we please by taking values of x sufficiently large. Then we say f tends to the limit A as x tends to infinity. More precisely, we have the following definition:

Definition 3: A function f tends to a limit A, as x tends to infinity if having chosen a positive number ε, there exists a positive number k such that

$$|f(x - A)| > \varepsilon \ \forall \ x \geq k.$$

The number E can be made as small as we like. Indeed, however small ε we may take, we can always find a number k for which the above inequality holds. We rewrite this definition in the following way:

A function f(x) → A as x → w if for every ε > 0, there exists k > 0 such that

$$|f(x) - A| < \varepsilon \text{ for all } x \geq k.$$

We write it as,

$$\lim_{x \to \infty} f(x) = A.$$

This notion of the limit of a function needs a slight modification when x tends to –∞. This is as follows:

We say that $\lim\limits_{x \to -\infty} f(x) = A$, if for a given ε > 0, there exists a number k < 0 such that

$$|f(x) - A| < E \text{ whenever } x \leq k.$$

We write it as $\lim\limits_{x \to -\infty} f(x) = A$.

Instead of f(x) approaching a real number A as x tends to +∞ or –m, we may also have f(x) approaching +∞ or –∞ as x tends to a real number 'a'. For example, if $f(x) = l/x^Z$, x ≠ 0 and x takes values near 0, the values of f(x) becomes larger and larger. Then we say that f(x) is tending to +∞ as x tends to 0. We can also have f(x) tending to +m or –∞ as x tends to +∞ or –∞. For example f(x) = x tends to +∞ or –∞ as x tends to +∞ or –∞. Again, the function f(x) = –x tends to +∞ or –∞ as x tends to –∞ or +∞. We formulate the following definition to cover all such cases of infinite limits.

Definition 4: Infinite Limits of a Function

Suppose a is a real number. We say that a function f tends to +m when x tends to a, if for a given positive real number M there exists a positive number δ such that

$$f(x) > M \text{ whenever } 0 < |x - a| < \delta.$$

We write it as

$$\lim_{x - a} f(x) = +\infty.$$

In this case we say that the function becomes unbounded and tends to +∞ as x tends to a.

In the same way, f is said to –∞ as x tends to a if for every real number –M, there is a positive number δ such that

$$f(x) < -M \text{ whenever } 0 < |x - a| < \delta.$$

We write it as

$$\lim_{x \to a} f(x) = -\infty.$$

Notes

In this case also f(x) is unbounded and tends to $-\infty$ as x tends to a. You can give similar definitions for $f(a+) = +\infty$, $f(a-) = +\infty$, $f(a+) = -\infty$, $f(a-) = -\infty$.

Now we define $\lim_{x \to \infty} f(x) = \infty$.

f is said to tend to ∞ as x tends to ∞ if given a number $M > 0$, there exists a number $k > 0$ such that

$$f(x) > M \text{ for } x \geq k.$$

We may similarly define

$$\lim_{x \to -\infty} f(x) = +\infty, \quad \lim_{x \to +\infty} f(x) = -\infty, \quad \lim_{x \to -\infty} f(x) = -\infty.$$

In all such cases we say that the function f becomes unbounded as x tends to $+\infty$ or $-\infty$ as the case may be.

It is easy to see from the definition of limit of a function that the limit of a constant function at any point in its domain is the constant itself. Similarly if $\lim_{x - a} f(x) = A$, then $\lim_{x - a} cf(x) = cA$ for any constant c where c is a real number.

Example: Justify that

$$\lim_{x - 2} \frac{1}{(x-2)^2} = \infty.$$

Solution: You have to verify that corresponding to a given positive number M, there exists a positive number 6, such that

$$\frac{1}{(x-2)^2} > M \text{ whenever } 0 < |x - 2| < 6.$$

Indeed for $x \neq 2$,

$$\frac{1}{(x-2)^2} > M \Rightarrow (x-2)^2 < \frac{1}{M}$$

$$\Rightarrow |x - 2| < \frac{1}{\sqrt{M}}.$$

Take $\delta = \frac{1}{\sqrt{M}}$. Then you can see that

$$\frac{1}{(x-2)^2} > M \text{ whenever } 0 < |x - 2| < 6.$$

Hence

$$\lim_{x - 2} \frac{1}{(x-2)^2} = \infty.$$

Task

1. Consider $f(x) = |x|$, $x \in R$. Show that $\lim_{x \to +\infty} f(x) = +\infty$, and $\lim_{x \to -\infty} f(x) = +w$ and $f(0+) = f(0-) = 0 = f(0)$
2. Let $f(x) = -|x|$, $x \in R$. Prove that $\lim_{x \to +\infty} f(x) = -\infty$ and $\lim_{x \to -\infty} f(x) = +\infty$ and $f(0) = f(0+) = f(0-) = 0$.

Notes

We have already stated that if a function t is define a by $f(x) = 1/x, x \neq 0$, then the limits f(0+) and f(0–) and $\lim_{x \to 0} f(x)$ do not exist. It simply means that these limits do not exist as real numbers. In other words, there is no (finite) real number A such that f (0+) = A f(0–) = A, or $\lim_{x \to 0} f(x) = A$.

10.2 Sequential Limits

In Unit 5, you studied the notion of the limit of a sequence. You also know that a sequence is also a function but a special type of function. What is special about a sequence? Do you remember it? Recall it from Unit 5. Naturally, you would like to know the relationship of a sequence and an arbitrary real function in terms of their limit concepts. Both require us to find a fixed number A as a first step. Both assume a small positive number ε as a test for closeness. For functions we need a positive number δ corresponding to the given positive number E and for sequences we need a positive integer m which depends on ε. So, then what is the difference between the two notions? The only difference is in their domains in the sense that the domain of a sequence is the set of natural numbers whereas the domain of an arbitrary function is any subset of the set of real numbers. In the case of a sequence, there are natural numbers only which exceed any choice of m. But for a function with a domain as an arbitrary set of real numbers, this is not necessary the case. Thus in a way, the notion of the limit of a function at infinity is a generalization of that of limit of a sequence.

Let us now, therefore, examine the connection between the limit of a function and the limit of a sequence called the sequential limit. We state and prove the following theorem for this purpose:

Theorem 3: Let a function f be defined in a neighbourhood of a point 'a' except possibly at 'a'. Then f(x) tends to a limit A as x tends to 'a' if and only if for every sequence (x_n), $x_n \neq a$ for any natural number n, converging to 'a', $f(x_n)$ converges to A.

Proof: Let, $\lim_{x \to a} f(x) = A$. Then for a number E > 0, there exists a 6 > 0 such that for 0 < | x – a | < 6 we have

$$|f(x) - A| < \varepsilon$$

Let (x,) be a sequence ($x_n \neq a$ for any $n \in N$) such that (x_n) converges to a i.e. $x_n \to a$.

Then corresponding to $\delta > 0$, there exists a natural number m such that for all $n \geq m$

$$|x_n - a| < \delta.$$

Consequently, we have

$$|f(X_n) - A| < \varepsilon, \ \forall \ n \geq m.$$

This implies that $f(x_n)$ converges to A.

Conversely, let $f(x_n)$ converge to A for every sequence x_n which converges to a, $x_n \neq a$ for any n.

Suppose $\lim_{x \to a} f(x) \neq A$.

Then there exists at least one ε, say $\varepsilon = \varepsilon_0$ such that for any $\delta > 0$ we have an x_δ such that

$$0 < |x_\delta - a| < \delta$$

and

$$|f(x_\delta) - A| \geq \varepsilon_0.$$

Let $\delta = \frac{1}{n}$, n = 1, 2, 3......

Notes

We get a sequence (x,) such that x, = x_δ where 6 = 1/n and

$$0 < |x_n - a| < \frac{1}{n} \text{ for } n = 1, 2, \ldots..$$

and

$$|f(x_n) - A| \geq \varepsilon_0.$$

$$0 < |x_n - a| \Rightarrow x, \neq a \text{ for any n.}$$

Since $\frac{1}{n} \to 0$ and $|x_n - a| < \frac{1}{n}$, it follows that x, $\to$ a.

But $|f(x_n) - A| \geq \varepsilon_0 \Rightarrow f(x_n) \to A$ i.e. $f(x_n)$ does not tend to A.

Therefore $x_n \neq a \;\forall$ a and x_n tends to a as n tends to ∞ whereas $f(X_n)$ does not converge to A, contradicting our hypothesis. This completes the proof of the theorem.

You may note that the above theorem is true even when either a or A is infinite or both a and A are infinite (i.e. $+\infty$ or $-\infty$).

By applying this theorem, we can decide about the existence or non-existence of limit of a function at a point. Consider the following examples:

Example: Let $f(x) = \begin{cases} 0 & \text{if x is rational} \\ 1 & \text{if x is irrational} \end{cases}$

Show that an point a in the real line R $\lim\limits_{x \to R} f(x)$ exists.

Solution: Consider any point 'a' of the real line. Let (p_n) be a sequence of rational numbers converging to the point 'a'. Since p_n is a rational number, $f(P_n) = 0$ for all n and consequently lim $f(P_n) = 0$, Now, consider a sequence (q) of irrational numbers converging to 'a'. Since q, is an irrational number, $f(q_n) = 1$ for all n and consequently lim $f(q_n) = 1$. So for two sequences (p_n) and (q,) converging to 'a'; sequences $(f(p_n))$ and $(f(q_n))$ do not converge to the same limit. Therefore $\lim\limits_{x-a} f(x)$ cannot exist for if it exists and is equal to A, then both $(f(p_n))$ and $(f(q_n))$ would have converged to the same limit A.

Example: Show that for the function f: R $\to$ R defined by f(x) = x Q x $\in$ R, $\lim\limits_{x-p} f(x)$ exists for every a $\in$ R.

Solution: Consider any point a $\in$ R. Let (x,) be a sequence of points of R converging to 'a'. Then $f(x_n)$ = x, and consequently lim $f(x_n)$ = lim (x_n) = a. So for every sequence $<x_n>$ converging to 'a' $(f(X_n))$ converges to 'a'. So by Theorem 3, $\lim\limits_{x-a} f(x) = a$. Consequently $\lim\limits_{x-a} f(x)$ exists for every a $\in$ R.

> *Task* Show that $\lim\limits_{x \to 1} 2^x = 2$ by proving that for any sequence (x_n), $x_n \neq 1$, converging to 1, 2^{xn} converges to 2.

10.3 Algebra of Limits

We discussed the algebra of limits of sequences. In this section, we apply the same algebraic operations to limits of functions. This will enable us to solve the problem of finding limits of functions. In other words we discuss limits of sum, difference, product and quotient of functions.

Notes

Definition 5: Algebraic Operations on Functions

Let f and g be two functions with domain D ⊂ R. Then the sum, difference, product, quotient of f and g denoted by f + g, f – g, fg, f/g are functions with domain D defined by

$$(f + g)(x) = f(x) + g(x)$$

$$(f - g)(x) = f(x) - g(x)$$

$$(fg)(x) = f(x).\ g(x)$$

$$(f/g)(x) = f(x)/g(x)$$

provided in the last case $g(x) \neq 0$ for all x in D.

Now we prove the theorem.

Theorem 4: If $\lim_{x\to a} f(x) = A$ and $\lim_{x\to a} g(x) = B$, where A and B are real numbers,

(1) $\lim_{x\to a} (f + g)(x) = A + B = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$,

(ii) $\lim_{x\to a} (f - g)(x) = A - B = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$,

(iii) $\lim_{x\to g} (f \cdot g)(x) = A \cdot B = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$,

(iv) If further $\lim_{x\to a} g(x) \neq 0$, then $\lim_{x\to g} f/g(x)$ exists and $\lim_{x\to g} \frac{f}{g}(x) = A/B = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$.

Proof: Since $\lim_{x\to a} f(x) = A$ and $\lim_{x\to a} g(x) = B$, corresponding to a number ε > 0. There exist numbers

$\delta_1 > 0$ and $\delta_2 > 0$ such that

$$0 < |x - a| < \delta_1 \Rightarrow |f(x) - A| < \varepsilon/2 \qquad (1)$$

$$0 < |x - a| < \delta_2 \Rightarrow |g(x) - B| < \varepsilon/2 \qquad (2)$$

Let δ = minimum (δ_1, δ_2). Then from (1) and (2) we have that

$$0 < |x - a| < \delta \Rightarrow |f(x) + g(x) - (A + B)| \leq |f(x) - A| + |g(x) - B|$$
$$< \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Which shows that $\lim_{x\to a} (f + g)(x) = \lim_{x\to a} f(x) + g(x) = A + B$

This proves part (i).

The proof of (ii) is exactly similar. Try it yourself.

(iii) $$|f(x) g(x) - AB| = |(f(x) - A) g(x) + A (g(x) - B)|$$
$$\leq |f(x) - A|\ |g(x)| + |A|.\ |(g(x) - B)|. \qquad (3)$$

Since $\lim_{x\to a} g(x) = B$ corresponding to 1, there exists a number $\alpha_0 > 0$
such that

$$0 < |x - a| < \alpha_0 \Rightarrow |g(x) - B| < 1.$$

which implies that $|g(x) \leq |g(x) - B| + |B| \leq 1 + |B| = K$ (say) (4)

Since f(x) = A, corresponding to ε < 0, there exists a number $\delta_1 > 0$ such that number $\delta_1 < 0$ such that

$$0 < |x - a| < \delta_1 \Rightarrow |f(x) - A| < \varepsilon/2K \qquad (5)$$

Notes

Since $\lim_{x \to a} g(x) = B$, corresponding to a number $\varepsilon > 0$, there exists a number $\delta_2 > 0$ such that

$$0 < |x - a| < \delta_2 \Rightarrow |f(x) - B| < \frac{E}{2(|A|+1)} \tag{6}$$

Let $\delta = \min(a_1, \delta_1, \delta_2)$. Then using (4), (5) and (6) in (3), we have for $0 < |x - a| < \delta$,

$$|f(x)\, g(x) - AB| \le |f(x) - A|\ |g(x)| + |A|\ |g(x) - B|$$
$$\le |f(x) - A|\,.\,K + |A|\ |(g(x) - B)|$$
$$< \frac{\varepsilon}{2K}\,.\,K + \frac{\varepsilon}{2(|A|+1)}\ |A| < \varepsilon/2 + \varepsilon/2 = \varepsilon. \tag{7}$$

Therefore, $\lim_{x \to a} g(x) = AB$ i.e. $\lim_{x \to a} (fg)(x) = AB = \lim_{x \to a} f(x).\ \lim_{x \to a} g(x)$, which proves part (iii) of the theorem.

(iv) First we show that *g* does not vanish in a neighbourhood of a.

$\lim_{x \to a} g(x) = B$ and $B \neq 0$. Therefore $|B| > 0$. Then corresponding to $\frac{|B|}{2}$ we have a number $\mu > 0$ such that for $0 < |x - a| < \mu$, $|g(x) - B| < \frac{|B|}{2}$.

Now by triangle inequality, we have

$$|\,|g(x)| - |B|\,| \le |g(x) - B| < \frac{|B|}{2}.$$

i.e.,
$$|B| - \frac{|B|}{2} < |g(x)| < |B| + \frac{|B|}{2}. \tag{8}$$

In other words, $0 < |x - a| < \mu \Rightarrow |g(x)| > \frac{|B|}{2}$.

Again since $\lim_{x \to a} g(x) = B$, for a given number $\varepsilon > 0$, we have a number $\mu' > 0$ such that $0 < |x - a| < \mu'$ implies that

$$|g(x) - B| < \frac{|B|^2 \varepsilon}{2}.$$

Let 6 = min (μ, p′). Then if $0 < |x - a| < \delta$, from (7) and (8) we have

$$\left|\frac{1}{g(x)} - \frac{1}{B}\right| = \frac{|B - g(x)|}{|g(x)||B|} < \frac{2|B - g(x)|}{|B|^2} < \frac{2|B|^2 \varepsilon}{2|B|^2} = \varepsilon.$$

This proves that $\lim_{x \to a} \frac{1}{g(x)} = \frac{1}{B}$.

Now by part (iii) of this theorem, we get that

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} f(x).\frac{1}{g(x)} = \lim_{x \to a} f(x).\lim_{x \to a} \frac{1}{g(x)}$$
$$= A.\frac{1}{B} = A/B.$$

i.e.,
$$\lim_{x \to a} \left(\frac{f}{g}\right)(x) = A/B = \frac{\lim_{x \to a} f(x)}{\lim_{x \to a} g(x)}.$$

Notes

This completes the proof of the theorem. You may note the theorem is true even when a = ±∞. You may also see that while proving (iv), we have proved that if

$$\lim_{x\to a} g(x) = B \neq 0, \text{ then } \lim_{x\to a} \frac{1}{g(x)} = \frac{1}{B}.$$

Before we solve some examples, we prove two more theorems.

Theorem 5: Let f and g be defined in the domain D and let $f(x) \le g(x)$ for all x in D. Then if $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist,

$$\lim_{x\to n} f(x) \le \lim_{x\to a} g(x).$$

Proof: Let $\lim_{x\to a} f(x) = A$, $\lim_{x\to a} g(x) = B$. If possible, let A > B.

for $\varepsilon = \frac{A-B}{2}$, there exist $\delta_1, \delta_2 > 0$ such that

$$0 < |x-a| < \delta_1 \Rightarrow |f(x) - A| < E$$

and $$0 < |x-a| < \delta_2 \Rightarrow |g(x) - B| < E.$$

If 6 = min. (δ_1, δ_2), then for $0 < |x-a| < 6$, $g(x) \in]B-\varepsilon, B+\varepsilon[$ and $f(x) \in]A-\varepsilon, A+\varepsilon[$. But $B + \varepsilon = A - E = \frac{A+B}{2}$. Therefore $g(x) < f(x)$ for $0 < |x-a| < 6$ which contradicts the given hypothesis. Thus $A \le B$.

Theorem 6: Let S and T be non-empty subsets of the real set R, and let f: S → T be a function of S onto T. Let g: U → R be a function whose domain U C R contains T. Let us assume that $\lim_{x\to a} f(x)$ exists and is equal to b and $\lim_{x\to b} g(y)$ exists and is equal to c. Then $\lim_{x\to a} g(f(x))$ exists and is equal to c.

Proof: Since $\lim_{x\to b} g(y) = c$, given a number E > 0, there exists a number $\alpha_0 > 0$ such that

$$0 < |y-b| < \alpha_0 \Rightarrow |g(y) - c| < \varepsilon.$$

Since $\lim_{x\to a} f(x) = b$, corresponding to $\alpha_0 > 0$, there exists $\delta > 0$ such that

$$0 < |x-a| < \delta \Rightarrow |f(x) - b| < \alpha_0.$$

Hence, taking y = f(x) and combining the two we get that for

$$0 < |x-a| < \delta, \ |g(f(x)) - c| = |g(y) - c| < \varepsilon$$

$$(\text{since } |f(x) - b| < \alpha_0).$$

This completes the proof of the theorem. Finally we give one more result without proof.

Result: If $\lim_{x\to a} f(x) = A$, $A > 0$ and $\lim_{x\to a} g(x) = B$ where A and B are finite real numbers then

$$\lim_{x\to a} f(x)^{g(x)} = A^B.$$

Now we discuss some examples. You will see how the above results help us in reducing the problem of finding limit of complicated functions to that of finding limits of simple functions.

Example: Find $\lim_{x\to\infty} \frac{(2x+7)(3x-11)(4x+5)}{4x^3+x-1}$

Solution:

$$\lim_{x\to\infty} \frac{(2x+7)(3x-11)(4x+5)}{4x^3+x-1}$$

Notes

$$= \lim_{x\to\infty} \frac{x^3\left[\left(2+\frac{7}{x}\right)\left(3-\frac{11}{x}\right)\left(4+\frac{5}{x}\right)\right]}{x^3\left(4+\frac{1}{x^2}-\frac{1}{x^3}\right)}$$

We divide the numerator and denominator by x^3 since x^3 is neither zero nor ∞.

$$= \lim_{x\to\infty} \frac{(2x+7)(3x-11)(4x+5)}{4x^3+x-1}$$

$$= \lim_{x\to\infty} \frac{\left(2+\frac{7}{x}\right)\left(3-\frac{11}{x}\right)\left(4+\frac{5}{x}\right)}{4+\frac{1}{x^2}-\frac{1}{x^3}} - \frac{2\times 3\times 4}{4} = .6.$$

Example: Find $\lim_{x\to 3} \frac{x^2-9}{x^2-4x+3}$

Solution: $$\lim_{x\to 3} \frac{x^2-9}{x^2-4x+3} = \lim_{x\to 3} \frac{(x-3)(x+3)}{(x-3)(x-1)}$$

Hence $$\lim_{x\to 3} \frac{x^2-9}{x^2-4x+3} = \lim_{x\to 3} \frac{x+3}{x-1}$$

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

The function $f(x) = \frac{x^2-9}{x^2-4x+3}$ is not defined at x = 3. But we are considering only the values of the function at those points x in a neighbourhood of 3 for which $x \neq 3$ and hence we can cancel x – 3 factor.

Example: Evaluate $\lim_{x\to 0} \frac{(1+x)^{1/2}-1}{(1+x)^{1/3}-1}$.

Solution: To make the problem easier, we make a substitution which enables us to get rid of fractional powers 1/2 and 1/3. L.C.M. of 2 and 3 is 6. So, we put $1 + x = y^6$.

Then we have

$$\lim_{x\to 0} \frac{(1+x)^{1/2}-1}{(1+x)^{1/3}-1} = \lim_{y\to 1} \frac{y^3-1}{y^2-1} = \lim_{y\to 1} \frac{(y-1)(y^2+y+1)}{(y-1)(y+1)}$$

$$= \lim_{y\to 1} \frac{y^2+y+1}{y+1} - \frac{3}{2}.$$

Self Assessment

Fill in the blanks:

1. The intuitive idea of limit was used both by Newton and Leibnitz in their independent invention of Differential Calculus around ....
2. The limit of a function that a point a is meaningful only if a is a limit point of its ....
3. For a function f: $R \to R$ defined by $f(x) = x^2$, find its limit when x tends to 1 by the ....

Notes

4. The function f is said to tend to a number A as x tends to 'a' from the left or through values smaller than 'a' if given a number E > 0, there exists a number $\delta > 0$ such that ....

10.4 Summary

- We started with the intuitive idea of a limit of a function. Then we derived the rigorous definition of the limit of a function, popularly called $\varepsilon - \delta$ definition of a limit. Further, we gave the notion of right and left hand limits of a function. It has been proved that $\lim_{x\to a} f(x) = A$ if and only if both right hand and left hand limits are equal to A i.e. $\lim_{x\to a+} f(x) = \lim_{x\to a-} f(x) = A$. In the same section we discussed the limit of a function as x tends to $+\infty$ or $-\infty$. Also we discussed the infinite limit of a function.
- We studied the idea of sequential limit of a function by connecting the idea of limit of an arbitrary function with the limit of a sequence. It has been shown how this relationship helps in finding the limits of functions.
- We defined the algebraic operations of sum, difference, product, quotient of two functions. We proved that the limit of the sum, difference, product and quotient of two functions at a point is equal to the sum, difference, product and quotient of the limits of the functions at the point provided in the case of quotient, the limit of the function in the denominator is non-zero. Finally in the same section, the usefulness of the algebra of limits in finding the limits of complicated functions has been illustrated.

10.5 Keywords

Function: A function f tends to a limit A, as x tends to infinity if having chosen a positive number ε, there exists a positive number k such that

$$f(x - A)| > \varepsilon \;\forall\; x \geq k.$$

Infinite Limits of a Function: Suppose a is a real number. We say that a function f tends to +m when x tends to a, if for a given positive real number M there exists a positive number δ such that

$$f(x) > M \text{ whenever } 0 < |x - a| < \delta.$$

10.6 Review Questions

1. Show that $\operatorname{Lim}_{x\to 2} \frac{x^2 - x + 18}{3x - 1} = 4$, using the $\varepsilon - \delta$ definition.
2. Find the limit of the function f defined as

   $f(x) = \frac{2x^2 + x}{3x}$, $x \neq 0$ when x tends to 0.
3. Find, if possible, the limit of the following functions.

   (i) $f(x) = \frac{|x - 2|}{x - 2}$, $x \neq 2$

   when x tends to 2.

   (ii) $f(x) = \frac{-1}{e^{1/x} + 1}$, $x \neq 0$

   when x tends to 0.

Notes

4. (i) Let $f(x) = \frac{1}{|x|}$, $x \neq 0$. Show that $\lim_{x \to 0+} f(x) = +\infty$, $\lim_{x \to 0-} f(x) = \infty$ and $\lim_{x - 0} f(x) = +\infty$.

   (ii) Let $f(x) = -\frac{1}{|x|}$, $x \neq 0$. Show that $\lim_{x \to 0+} f(x) = -\infty$, $\lim_{x \to 0-} f(x) = -\infty$ and $\lim_{x - 0} f(x) = -\infty$.

   (iii) Let $f(x) = \frac{1}{x}$, $x \neq 0$. Prove that $\lim_{x \to 0+} f(x) = +\infty$, $\lim_{x \to 0-} f(x) = -\infty$.

   (iv) Let $f(x) = -\frac{1}{x}$, $x \neq 0$. Prove that $\lim_{x \to 0+} f(x) = -\infty$, $\lim_{x \to 0-} f(x) = \infty$.

5. Show that for the function f: R → R defined by

$$f(x) = x^2,$$

f(x) exists for every a ∈ R.

6. Find

   (i) $\lim_{x - w} \frac{(2x+3)^3 (3x-2)^2}{x^5 + 5}$

   (ii) $\lim_{x - w} \frac{(x^3+1)^{1/3}}{x+1}$.

7. If $g(x) = \begin{cases} 2x & \text{for } 0 \text{ s } x < 1 \\ 4 & \text{for } x = 1 \\ 5 - 3x & \text{for } 1 < x \leq 2. \end{cases}$

   find $\lim_{x \to 1} g(x)$

8. Find

   (i) $\lim_{x \to 4} \frac{\sqrt{x} - 2}{x - 4}$

   (ii) Find $\lim_{x \to 2} \frac{3x^2 - x - 10}{x^2 + 5x - 14}$

   (iii) $\lim_{x \to 0} \frac{1 - \cos x}{x^2}$

   (iv) $\lim_{x \to a} \frac{\sin x - \sin a}{x - a}$

   (v) $\lim_{x \to \infty} \left(\frac{x-1}{x+1}\right)^x$

   (vi) $\lim_{x \to 0} \left(\frac{\sin 2x}{x}\right)^{1+x}$

   (vii) $\lim_{x \to \infty} \left(\frac{x+1}{2x+1}\right)^{x^2}$

Answers: Self Assessment

1. 1675
2. domain
3. ε – δ approach
4. | f(x) – A | < ε for a – 6 < x < a

10.7 Further Readings

Notes

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Notes

Unit 11: Continuity

CONTENTS

Objectives
Introduction
11.1 Continuous Functions
11.2 Algebra of Continuous Functions
11.3 Non-continuous Functions
11.4 Summary
11.5 Keywords
11.6 Review Questions
11.7 Further Readings

Objectives

After studying this unit, you will be able to:

- Define the continuity of a function at a point of its domain
- Determine whether a given function is continuous or not
- Construct new continuous functions from a given class of continuous functions

Introduction

The values of a function f(x) approaching a number A as the variable x approaches a given point a. When there is break (or jump) in the graph, then this property fails at that point. This idea of continuity is, therefore, connected with the value of $\lim_{x \to a} f(x)$ and the value of the function f at the point a. We define in this unit the continuity of a function at a given point a in precise mathematical language. Therefore extend it to the continuity of a function on a non-empty subset of the domain of f which could be the whole of the domain of f also. We study the effect of the algebraic operations of addition, subtraction, multiplication and division on continuous functions.

Here we discuss the properties of continuous functions and the concept of uniform continuity.

11.1 Continuous Functions

We have seen that the limit of a function f as the variable x approaches a given point a in the domain of a function f does not depend at all on the value of the function at that point a but it depends only on the values of the function at the points near a. In fact, even if the function f is not defined at a then $\lim_{x \to a} f(x)$ may exist.

For example $\lim_{x \to 1} f(x)$ exists when

$$f(x) = \frac{x^2 - 1}{x - 1} \text{ though f is not defined at } x = 1.$$

We have also seen that $\lim_{x \to a} f(x)$ may exist, still it need not be the same as f(a) when it exists. Naturally, we would like to examine the special case when both $\lim_{x \to a} f(x)$ and f(a) exist and are

Notes

equal. If a function has these properties, then it is called a continuous function at the point a. We give the precise definition as follows:

Definition 1: Continuity of a Function at a Point

A function f defined on a subset S of the set R is said to be continuous at a point a ∈ S, if

(i) $\lim_{x \to a} f(x)$ exists and is finite

(ii) $\lim_{x \to a} f(x) = f(a)$.

Note that in this definition, we assume that S contains some open interval containing the point a. If we assume that there exists a half open (semi-open) interval [a, c[contained in S for some c∈ R, then in the above definition, we can replace $\lim_{x \to a} f(x)$ by $\lim_{x \to a+} f(x)$ and say that the function is continuous from the right of a or f is right continuous at a.

Similarly, you can define left continuity at a, replacing the role of $\lim_{x \to a} f(x)$ by $\lim_{x \to a-} f(x)$. Thus, f is continuous from the right at a if and only if

$$f(a+) = f(a)$$

It is continuous from the left at a if and only if

$$f(a-) = f(a).$$

From the definition of continuity of a function f at a point a and properties of limits it follows that f(a+) = f(a–) = f(a) if and only if, f is continuous at a. If a function is both continuous from the right and continuous from the left at a point a, then it is continuous at a and conversely.

The definition X is popularly known as the Limit-Definition of Continuity.

Since $\lim_{x \to a} f(x)$ is also defined, in terms of ε and δ, we also have an equivalent formulation of the definition X. Note that whenever we talk of continuity of a function f at a in S, we always assume that S contains a neighbourhood containing a. Also remember that if there is one such neighbourhood there are infinitely many such neighbourhoods. An equivalent definition of continuity in terms of ε and δ is given as follows:

Definition 2: (●●-Definition of Continuity

A function f is continuous at x = a if f is defined in a neighbourhood of a and corresponding to a given number E > 0, there exists some number δ > 0 such that |x – a | < δ implies |f(x) – f(a)| < E.

Note that unlike in the definition of limit, we should have

$$|f(x) - f(a)| < E \text{ for } |x - a| < 6.$$

The two definitions are equivalent. Though this fact is almost obvious, it will be appropriate to prove it.

Theorem 1: The limit definition of continuity and the (ε, δ)-definition of continuity are equivalent.

Proof: Suppose f is continuous at a point a in the sense of the limit definition. Then given ε > 0, we have a δ > 0 such that 0 < |x – a| < δ implies |f(x) – f(a)| < ε. When x = a, we trivially have

$$|f(x) - f(a)| = 0 < \varepsilon.$$

Hence, $|x - a| < \delta \Rightarrow |f(x) - f(a)| < E$

which is the (ε, δ)-definition.

Conversely we now assume that f is continuous in the sense of (ε, δ)-definition. Then for every E > 0 there exists a δ > 0 such that

$$|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Notes

Leaving the point 'a', we can write it as

$$0 < (x - a) < \delta \Rightarrow |f(x) - f(a)| < E.$$

This implies the existence, of $\lim_{x \to a} f(x)$ and that $\lim_{x \to a} f(x) = f(a)$.

Note that δ in the definition 2, in general, depends on the given function f, ε and the point a. Also $|x - a| < \delta$ if and only if $a - \delta < x < a + \delta$ and $]a - \delta, a + \delta[$ is an open interval containing a. Similarly $|f(x) - f(a)| < \varepsilon$ if and only if

$$f(a) - \varepsilon < f(x) < f(a) + E.$$

We see that f is continuous at a point a, if corresponding to a given (open) ε-neighbourhood U of f(a) there exists a (open) δ-neighbourhood V of a such that $f(V) \subset U$. Observe that this is the same as $x \in V \Rightarrow f(x) \in U$. This formulation of the continuity at a is more useful to generalise this definition to more general situations in Higher Mathematics.

A function f is said to be continuous on a set S if it is continuous at every point of the set S. It is clear that a constant function defined on S is continuous on S.

Example: Examine the continuity of the following functions:

(i) The absolute value (Modulus) function,

(ii) The signum function.

Solution:

(i) You know that the absolute value function

$f: R \to R$ is defined as $f(x) = |x|, \forall x \in R$.

The function is continuous at every point $x \in R$. For given $\varepsilon > 0$, we can choose $\delta = \varepsilon$ itself. If $a \in R$ be any point them $|x - a| < \delta = \varepsilon$ implies that

$$|f(x) - f(a)| = ||x| - |a|| \leq |x - a| < \varepsilon.$$

(ii) The signum function, as you know a function $f: R \to R$ defined as

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

This function is not continuous at the point x = 0. We have already seen that $f(0+) = 1$, $f(0-) = -1$. Since $f(0+) \neq f(0-)$, $\lim_{x \to 0} f(x)$ does not exist and consequently the function is not continuous at x = 0. For every point $x \neq 0$ the function f is continuous. This is easily seen from the graph of the function f. There is a jump at the point x = 0 in the values of f(x) defined in a neighbourhood of 0.

Note that if $f: R \to R$ is defined as,

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

then, it is easy to see that this function is continuous from the right at x = 0 but not from the left. It is continuous at every point $x \neq 0$.

Similarly, if f is defined by $f(x) = 1$ if $x > 0$

$= -1$ if $x \leq 0$

then f is continuous from the left at x = 0 but not from the right.

Notes

Example: Discuss the continuity of the function sin x on the real line R.

Solution: Let f(x) = Sin x $\forall$ x $\in$ R.

We show by the (ε, δ)-definition that f is continuous at every point of R.

Consider an arbitrary point a $\in$ R. We have

$$|f(x) - f(a)| = |\sin x - \sin a| = \left|2 \sin \frac{x-a}{2} \cos \frac{x+a}{2}\right|$$

$$= 2\left|\sin \frac{x-a}{2}\right| \left|\cos \frac{x+a}{2}\right|$$

$$\leq 2\left|\sin \frac{x-a}{2}\right| \left(\text{since} \left|\cos \frac{x+a}{2}\right| \leq I\right)$$

From Trigonometry, you know that $|\text{Sin } \theta| \leq |\theta|$.

Therefore $\left|\sin \frac{x-a}{2}\right| \leq \left|\frac{x-a}{2}\right| = \frac{|x-a|}{2}$

Consequently $|f(x) - f(a)| \leq |x - a|$

$< \varepsilon$ if $|x - a] < \delta$ where $\delta = \varepsilon$.

So f is continuous at the point a. But a is any point of R. Hence Sin x is continuous on the real line R.

Task Discuss the continuity cos x on the real R.

As we have connected the limit of a function with the limit of a sequence of real numbers. In the same way, we can discuss the continuity of a function in the language of the sequence of real numbers in the domain of the function. This is explained in the following theorem.

Theorem 2: A function f: S $\rightarrow$ R is continuous at point a in S if and only for every sequence (x_n), ($x_n \in S$) converging do a, f(x_n) converges to f(a).

Proof: Let us suppose that f is continuous at a. Then $\lim_{x \to a} f(x) = f(a)$.

Given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|x - a| < 6 \Rightarrow |f(x) - f(a)| < \varepsilon.$$

If x_n is a sequence converging to 'a', then corresponding to $\delta > 0$, there exists a positive integer M such that

$$|x_n - a| < \delta \text{ for } n \geq M.$$

Thus, for n $\geq$ M, we have $|x_n - a| < \delta$ which, in turn, implies that

$$|f(x_n) - f(a)| < \varepsilon,$$

proving thereby f(x_n) converges to f(a).

Notes

Conversely, let us suppose that whenever x_n converges to a, $f(x_n)$ converges to f(a). Then we have to prove that f is continuous at a. For this, we have to show that corresponding to an E > 0, there exists some $\delta > 0$ such that

$$|f(x) - f(a)| \geq \varepsilon, \text{ whenever } |x - a| < \delta.$$

If not, i.e., if f is not continuous at a, then there exists an E > 0 such that whatever $\delta > 0$ we take there exists an x_δ such that

$$|x_\delta - a| < \delta \text{ but } |f(x_\delta) - f(a)| \geq \varepsilon.$$

By taking $\delta = 1, 1/2, 1/3,....$ in succession we get a sequence (x_n), where $x_n = x_\delta$ for $\delta = 1/n$, such that $|f(x_n) - f(a)| \geq \varepsilon$. The sequence (x_n) converges to a. For, if m > 0, these exists M such that $1/n < m$ for $n \geq M$ and therefore $|x_n - a| < m$ for $n \geq M$. But $f(x_n)$ does not converge to f(a), a contradiction to our hypothesis. This completes the proof of the theorem.

Theorem 2 is sometimes used as a definition of the continuity of a function in terms of the convergent sequences. This is popularly known as the Sequential Definition of Continuity which we state as follows:

Definition 3: Sequential Continuity of a Function

Let f be a real-valued function whose domain is a subset of the set R. The function f is said to be continuous at a point a if, for every sequence (x_n) in the domain of f converging to a, we have,

$$\lim_{n \to \infty} f(X_n) = f(a)$$

The next example illustrates this definition.

Example: Let f: $R \to R$ be defined as

$$f(x) = 2x^2 + 1, \ \forall \ x \in R$$

Prove that f is continuous on R by using the sequential definition of the continuity of a function.

Solution: Suppose (x_n) is a sequence which converges to a point 'a' of R. Then, we have

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} (2x_n^2 + 1) = 2(\lim_{n - m} x_n)^2 + 1 = 2a^2 + 1 = f(a)$$

This shows that f is continuous at a point $a \in R$. Since a is an arbitrary element of R, therefore, f is continuous everywhere on R.

Task Prove by sequential definition of continuity that the function $f : R \to R$ defined by $f(x) = \sqrt{x}$ is continuous at x = 0.

11.2 Algebra of Continuous Functions

As we have proved limit theorems for sum, difference, product, etc. of two functions, we have similar results for continuous functions also. These algebra operations on the class of continuous functions can be deduced from the corresponding theorems on limits of functions, using the limit definition of continuity. We leave this deduction as an exercise for you. However, we give a formal proof of these algebraic operations by another method which illustrates the use of Theorem 2. We prove the following theorem:

Notes

Theorem 3: Let f and g be any real functions both continuous at a point $a \in R$. Then,

(i) αf defined by $(\alpha f)(x) = \alpha f(x)$, is continuous for any real number α,

(ii) f + g defined by $(f + g)(x) = f(x) + g(x)$ is continuous at a,

(iii) f – g defined by $(f - g)(x) = f(x) - g(x)$ is continuous at a,

(iv) fg defined by $(fg)(x) = f(x)\, g(x)$ is continuous at a,

(v) f/g defined by $(f/g)(x) = \dfrac{f(x)}{g(x)}$, is continuous at a provided $g(a) \neq 0$.

Proof: Let x_n be an arbitrary sequence converging to a. Then the continuity of f and g imply that the sequences $f(x_n)$ and $g(x_n)$ converge to f(a) and g(a) respectively. In other words, $\lim f(x_n) = f(a)$, $\lim g(x_n) = g(a)$.

Using the algebra of sequences, we can conclude that

$$\lim \alpha f(x_n) = \alpha f(a),$$

$$\lim (f + g)(x_n) = \lim f(x_n) + \lim g(x_n) = f(a) + g(a),$$

$$\lim (f - g)(x_n) = \lim f(x_n) - \lim g(x_n) = f(a) - g(a),$$

$$\lim (f \cdot g)(x_n) = \lim f(x_n) \lim g(x_n) = f(a)\, g(a).$$

If infinite number of x_n's are such that $g(x_n) = 0$, then $g(X_n) - g(a)$ implies that $g(a) = 0$, a contradiction.

This proves the parts (i), (ii), (iii) and (iv). To prove the part (v) we proceed as follows:

Since $g(a) \neq 0$, we can find $a > 0$ such that the interval $]g(a) - \alpha, g(a) + \alpha[$ is either entirely to the right or to the left of zero depending on whether $g(a) > 0$ or $g(a) < 0$. Corresponding to $a > 0$, there exists a $\delta_1 > 0$ such that $|x - a| < \delta_1$ implies $|g(x) - g(a)| < \alpha$, i.e., $g(a) - \alpha < g(x) < g(a) + at$. Thus, for x such that $|x - a| < \delta_i$, $g(x) \neq 0$. If (x_n) converges to a, omitting a finite number of terms of the sequence if necessary, then we can assume that $g(x_n) \neq 0$, for all n. Hence, $\dfrac{f(x_n)}{g(x_n)}$ converges to $\dfrac{f(a)}{g(a)}$ and so $\dfrac{f}{g}$ is continuous at a. This completes the proof of the theorem.

In part (v) if we define f by $f(x) = 1$, then it follows that if g is continuous at 'a' and $g(a) \neq 0$, then its reciprocal function 1/g is continuous at 'a'.

Now, we prove another theorem, which shows that a continuous function of a continuous function is continuous.

Theorem 4: Let f and g be two real functions such that the range of g is contained in, the domain of f. If g is continuous at x = a, f is continuous at b = g(a) and h(x) = f(g(x)), for x in the domain of g, then h is continuous at a.

Proof: Given $\varepsilon > 0$, the continuity of f at b = g(a) implies the existence of an $\eta > 0$ such that for

$$|y - b| < \eta, \ |f(y) - f(b)| < \varepsilon \qquad \text{...(l)}$$

Corresponding to $\eta > 0$, from the continuity of g at x = a, we get a $\delta > 0$ such that

$$|x - a| < \delta \text{ implies } |g(x) - g(a)| < \eta \qquad \text{...(2)}$$

Combining (1) and (2) we get that

$|x - a| < \delta$ implies that

$$|h(x) - h(a)| = |f(g(x)) - f(g(a))|$$

$$= |f(y) - f(b)| < \varepsilon,$$

Notes

where we have taken y = g(x). Hence h is continuous at a which proves the theorem.

Let us now study the following example:

Example: Examine for continuity the following functions:

(i) The polynomial function f R → R defined by

$$f(x) = a, + a_1x + a_2x^2 +... + a_xx^n.$$

(ii) The rational function f: R → R defined as

$$f(x) = \frac{f(x)}{q(x)}, \forall \text{ x for which } q(x) \neq 0.$$

Solution:

(i) It is obvious that the function f(x) = x, x ∈ R, is continuous on the whole of the real line. It follows from theorem 3 that the functions x^Z, x^3,...., are all continuous. The fact that constant functions are continuous, we get that any polynomial f(x) in x, i.e., the function f defined by

$$f(x) = a, + a_1x+ a_2x^2 +... + a_n x^n,$$

is continuous on R.

(ii) It follows from theorem 3(v) that a rational function f, defined by,

$$f(x) = \frac{p(x)}{q(x)} = \frac{a_0 + a_1x + \ldots + a_nx^n}{b_0 + b_1x + \ldots + b_mx^m}$$

is continuous at every point a ∈ R for which q(a) ≠ 0.

11.3 Non-continuous Functions

You have seen that a function may or may not be continuous at a point of the domain of the function. Let us now examine why a function fails to be continuous.

A function f: S → R fails to be continuous on its domain S if it is not continuous at a particular point of S. This means that there exists a point a ∈ S such that, either

(i) $\lim_{x-a} f(x)$ does not exist, or

(ii) $\lim_{x-a} f(x)$ exists but is not equal to f(a).

But you know that a function f is continuous at a point a if and only if

$$f(a+) = f(a–) = f(a).$$

Thus, if f is not continuous at a, then one of the following will happen:

(i) either f(a+) or f(a–) does not exist (this includes the case when both f(a+) and f(a–) do not exist).

(ii) both f(a+) and f(a–) exist but f(a+) ≠ f(a–).

(iii) both f(a+) and f(a–) exist and f(a+) = f(a–) but they are not equal to f(a).

If a function f: S → R is discontinuous for each b ∈ S, then we say that totally discontinuous an S. Functions which are totally discontinuous are often encountered but by no means rare. We give an example.

Notes

Example: Examine whether or not the function f: R → R defined as,

$$f(x) = \begin{cases} 1, & \text{if x is irrational} \\ 0, & \text{if x is rational} \end{cases}$$

is totally discontinuous.

Solution: Let b be an arbitrary but fixed real number. Choose ε = 1/2. Let δ > 0 be fixed. Then the interval defined by

$$|x - b| < \delta$$

is $(x: b - \delta < x < b + \delta)$

or $]b - \delta, b + \delta[$

This interval contains both rational as well as irrational numbers. Why?

If b is rational, then choose x in the interval to be irrational, If b is irrational then choose x in the interval to be rational. In either case,

$$0 < |x - b| < \delta$$

and

$$|f(x) - f(b)| = I > \varepsilon.$$

Thus, f is not continuous at b. Since b is an arbitrary element of S, f is not continuous at any point of S and hence is totally discontinuous.

There are certain discontinuities which can be removed. These are known as removable discontinuities. A discontinuity of a given function f: S → R is said to be removable if the limit of f(x) as x tends to a exits and that

$$\lim_{x \to a} f(x) \neq f(a)$$

In other words, f has removable discontinuity at x = a if f(a+) = f(a–) but none is equal to f(a).

The removable discontinuities of a function can be removed simply by changing the value of the function at the point a of discontinuity. For this a function with removable discontinuities can be thought of as being almost continuous. We discuss the following example to illustrate a few cases of removable discontinuities.

Example: Discuss the nature of the discontinuities of the following functions:

(i) $f(x) = \dfrac{x^2 - 4}{x - 4}$, $x \neq 2$

$= 1$ $x = 2$

at x = 2.

(ii) $f(x) = 3$, $x \neq 3$

$= 1$ $x = 3$

at x = 3.

(iii) $f(x) = x^2$, $x \in]-2, 0 (U) 0, 2[$

$= 1$ $x = 0$

at x = 0.

Notes

Solution:

(i) This function is discontinuous at $x = 2$. This is a removable discontinuity, for if we redefine $f(x) = 4$, then we can restore the continuity of f at $x = 2$.

(ii) This is again a case of removable discontinuity at 3. Therefore, if f is defined by $f(x) = 3 \ \forall x \in R$, then it is continuous at $x = 3$.

(iii) This function is discontinuous at $x = 0$. Why? This is a case of discontinuity which is removable. To remove the discontinuity, set $f(0) = 0$. In other words, define f as

$$f(x) = x^2, x \in]-2, 0 \ [\cup] \ 0, 2 [$$

$$= 0, \qquad x = 0$$

This is continuous at $x = 0$. Verify it.

Example: Let a function f: $R \to R$ be defined as,

(i) $f(x) = \frac{1}{x}, \qquad x \neq 0$

$= 0, \qquad x = 0$

(ii) $f(x) = \frac{1}{x}, \qquad \text{if } x > 0$

$= 1, \qquad \text{if } x < 0$

(iii) $f(x) = \frac{1}{x}, \qquad \text{if } x < 0$

$= 1, \qquad \text{if } x > 0$

Test the continuity of the function. Determine the type of discontinuity if it exists.

Solution:

(i) Here $f(0+)$ and $f(0-)$ both do not exist (as finite real numbers) and so function is discontinuous. This is not a case of removable discontinuity.

(ii) In this case, $f(0)$ does not exist whereas $f(0+)$ exists and $f(0-) = f(0) = 1$. This is not a case of removable discontinuity.

Task Prove that the function f defined by $f(x) = x \sin 1/x$ if $x \neq 0$ and $f(0) = 1$ has a removable discontinuity at $x = 0$.

Self Assessment

Fill in the blanks:

1. A function f is said to be .... on a set S if, it is continuous at every point of the set S. It is clear that a constant function defined on S is continuous on S.

2. A function f: $S \to R$ is continuous at point a, in S if and only 31 far every sequence (x_n), $(x_n \in S)$ converging do a, $f(x_n)$ .... to f(a).

Notes

3. Let f be a .... whose domain is a subset of the set R. The function f is said to be continuous at a point a if, for every sequence (x_n) in the domain of f converging to a, we have, $\lim_{n\to\infty} f(X_n) = f(a)$.

4. Let f and g be two real functions such that the range of g is contained in, the domain of f. If g is continuous at x = a, f is continuous at b = g(a) and ...., for x in the domain of g, then h is continuous at a.

5. A function f : S → R fails to be .... on its domain S if it is not continuous at a particular point of S.

11.4 Summary

The concept of the continuity of a function at a point of its domain and on a subset of its domain. The limit definition and (ε, –δ)-definition of continuity. It has been proved that both the definitions are equivalent. Sequential definition of continuity has been discussed and illustrations regarding its use for solving problems have been given. The algebra of continuous functions is considered and it has been proved that the sum, difference, product and quotient of two continuous functions at a point is also continuous at the point provided in the case of quotient, the function occurring in the denominator is not zero at the point. In the same section, we have proved that a continuous function of a continuous function is continuous. Finally in Section 9.4, discontinuous and totally discontinuous functions are discussed. Also in this section, one kind of discontinuity that is removable discontinuity has been studied.

11.5 Keywords

Continuity: A function f is continuous at x = a if f is defined in a neighbourhood of a and corresponding to a given number E > 0, there exists some number δ > 0 such that |x – a | < δ implies |f(x) – f(a)| < E.

Sequential Continuity of a Function: Let f be a real-valued function whose domain is a subset of the set R. The function f is said to be continuous at a point a if, for every sequence (x_n) in the domain of f converging to a, we have, $\lim_{n\to\infty} f(X_n) = f(a)$

11.6 Review Questions

1. Examine the continuity of the following functions:

   (i) The function f: R – (0) → R defined as

   $$f(x) = \frac{|x|}{x},$$

   at the point x = 0

   (ii) The function f: R → (1) – R defined as

   $$f(x) = \frac{x^2-1}{x-1},$$

   (iii) The function f: R – (0) – R defined as

   $$f(x) = \frac{1}{x}.$$

Notes

2. Examine the continuity of the function f: R → R defined as,

   (i) $f(x) = x^3$ at a point $a \in R$;

   (ii) $f(x) = \begin{cases} \dfrac{x^2 - 4}{x - 2}, & \text{if } x \neq 2 \\ 1, & \text{if } x = 2 \end{cases}$

3. Show that the function f: R → R defined by

$$f(x) = \begin{cases} 1, & \text{if x is rational} \\ 0, & \text{if x is irrational} \end{cases}$$

   is totally discontinuous. Does f(a+) and f(a–) exist at any point $a \in R$?

4. Prove that the function |f| defined by |f|(x) = |f(x)| for every real x is continuous on R whenever f is continuous on R.

5. (i) Find the type of discontinuity at x = 0 of the function f defined by

$$f(x) = x + 1 \text{ if } x > 0,\ f(x) = -(x + 1) \text{ if } x < 0 \text{ and } f(0) = 0.$$

   (ii) The function f is defined by

$$f(x) = \sin\frac{1}{x}, \qquad x \neq 0$$
$$= 0, \qquad x = 0$$

   Is f continuous at 0?

Answers: Self Assessment

1. continuous
2. converges
3. real-valued function
4. h(x) = f(g(x))
5. continuous

11.7 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol : Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik : Mathematical Analysis.

H.L. Royden : Real Analysis, Ch. 3, 4.

Unit 12: Properties of Continuous Functions

CONTENTS

Objectives

Introduction

12.1 Continuity on Bounded Closed Intervals

12.2 Pointwise Continuity and Uniform Continuity

12.3 Summary

12.4 Keywords

12.5 Review Questions

12.6 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss the properties of continuous functions on bounded closed intervals
- Explain the important role played by bounded closed intervals in Real Analysis
- Describe the concept of uniform continuity and its relationship with continuity

Introduction

Having studied in the last two units you have studied about limit and continuity of a function at a point, algebra of limits and continuous functions, the connection between limits and continuity, etc., we now take up the study of the behaviour of continuous functions and bounded closed intervals on the real line. You will learn that continuous functions on such intervals are bounded and attain their bounds; they take all values in between any two values taken at points of such intervals. You will also be introduced to the concept of uniform continuity and further you will see that a continuous function on a bounded closed interval is uniformly continuous. This means that continuous functions are well-behaved on bounded closed intervals. Thus, we will see that bounded closed intervals form an important subclass of the class of subsets of the real line which are known as compact subsets of the real line. You will study more about this in higher mathematics at a later stage. We will henceforth call bounded closed intervals of R as compact intervals.

The results of this unit play an important and crucial role in Real Analysis and so for further study in analysis, you must understand clearly the various theorems given in this unit.

It may be noted that an interval of R will not be a compact interval if it is not a bounded or closed interval.

12.1 Continuity on Bounded Closed Intervals

We now consider functions continuous on bounded closed intervals. They have properties which fail to be true when the intervals are not bounded or closed. Firstly, we prove the properties and then with the help of examples we will show the failures of these properties. To prove these properties, we need an important property of the real line that was discussed in Unit 1.

Notes

This property called the completeness property of R states as follows:

Any non-empty subset of the Seal Hue R which is bounded above has the least upper bound or equivalently, any non-empty subset of R which is bounded below has the greatest lower bound.

In the following theorems we prove the properties of functions continuous on bounded closed intervals. In the first two theorems we show that a continuous function on a bounded closed interval is bounded and attains its bounds in the interval. Recall that f is bounded on a set S, if there exists a constant $M > 0$ such that $|f(x)| \leq M$ for all $x \in S$. Note also that a real function f defined on a domain D (whether bounded or not) is bounded if and only if its range f(D) is a bounded subset of R.

Theorem 1: A function f continuous on a bounded and closed interval [a, b] is necessarily a bounded function.

Proof: Let S be the collection of all real numbers c in the interval [a, b] such that f is bounded on the interval [a, c]. That is, a real number c in [a, b] belongs to S if and only if there exists a constant M_c such that $|f(x)| \leq M_c$ for all x in [a, c]. Clearly, $S \neq \phi$ since $a \in S$ and b is an upper bound for S. Hence, by completeness property of R, there exists a least upper bound for S. Let it be k (say). Clearly, $k \leq b$. We prove that $k \in S$ and $k = b$ which will complete the proof of the theorem.

Corresponding to $\varepsilon = 1$, by the continuity of f at $k(\leq b)$ there exists a $d > 0$ such that

$$|f(x) - f(k)| < \varepsilon = 1 \text{ whenever } |x - k| < d, x \in [a, b].$$

By the triangle inequality we have

$$|f(x)| - |f(k)| \leq |f(x) - f(k)| < 1$$

Hence, for all x in [a. b] for which $|x - k| < d$, we have that

$$|f(x)| < |f(k)| + 1 \quad ...(1)$$

Since k is the least upper bound of S, $k - d$ is not an upper bound of S. Therefore, there is a number $c \in S$ such that

$$k - d < c \leq k$$

Consider any t such that $k \leq t < k + d$. If x belongs to the interval [c, t] then $|x - k| < d$. For,

$$x \in [c, t] = \Rightarrow c \leq x \leq t \Rightarrow k - d < c \leq x \leq t < k + d \quad ...(2)$$

Now $c \in S$ implies that there exists $M_c > 0$ such that for all

$$x \in [a, c], |f(x)| \leq M_c$$

$$x \in [a, t] = [a, c] \cup [c, t] \Rightarrow \text{either } x \in [a, c] \text{ or } x \in [c, t].$$

If $x \in [a, c]$, by (3) we have

$$|f(x)| \leq M_c < M_c + |f(k)| + 1.$$

If, however, $x \in [c, t]$ then by (1) and (2) we have

$$|f(x)| < |f(k)| + 1 < M_c + |f(k)| + 1$$

In any case we get that $x \in [a, t]$ implies that

$$|f(x)| < M_c + |f(k)| + 1$$

This shows that f is bounded in the interval [a, t] thus proving that $t \in S$ whenever $k \leq t < k + d$. In particular $k \in S$. In such a case $k = b$. For otherwise we can choose a 't' such that $k < t < k + d$ and $t \in S$ which will contradict the fact that k is an upper bound. This completes the proof of the theorem.

Having proved the boundedness of the : function continuous on a bounded closed interval, we now prove that the function attains its bounds, that is, it has the greatest and the smallest values.

Theorem 2: If f is a continuous function on the bounded closed interval [a, b] then there exists points x_1 and x_2 in [a, b] such that $f(x_1)$ S $f(x) \leq f(x_2)$ for all $x \in [a, b]$ (i.e. f attains its bounds).

Proof: From Theorem 1, we know that f is bounded on [a, b].

Therefore there exists M such that $|f(x)| \leq M \ \forall x \in [a, b]$.

Hence, the collection {f(x) : a a x ≤ b] has an upper bound, since $f(x) \leq |f(x)| \leq M \ \forall x \in [a, b]$.

So by the completeness property of R, the set (f(x) : a ≤ x ≤ b) has a least upper bound.

Let us denote by K the least upper bound of {f(x) : a ≤ x ≤ b].

Then f(x) ≤ K for all x such that a ≤ x ≤ b. We claim that there exists x_2 in [a, b] such that $f(x_2) = K$. If there is no such x_2, then K – f(x) > 0 for all a ≤ x ≤ b. Hence, the function g given by,

$$g(x) = \frac{1}{K - f(x)}$$

is defined for all x in [a, b] and g is continuous since f is continuous. Therefore by Theorem 1, there exists a constant M′ > 0 such that

$$|g(x)| \leq M' \ \forall x \in [a, b]$$

Thus, we get

$$|g(x)| = \frac{1}{|K - f(x)|} = \frac{1}{K - f(x)} \leq M'$$

i.e., $$f(x) \leq K - \frac{1}{M} \ \forall x \in [a, b].$$

But this contradicts the choice of K as the least upper bound of the set (f(x) : a | x ≤ b). This contradiction, therefore, proves the existence of an x_2 in [a, b] such that $f(x_2) = K \geq f(x)$ for a ≤ x ≤ b. The existence of x_1 in [a, b] such that $f(x_1) \leq f(x)$ for a ≤ x | b can be proved on exactly similar lines by taking the g.l.b. of {f(x) : a ≤ x ≤ b} instead of the l.u.b. or else by considering –f instead of f.

Theorems 1 and 2 are usually proved using what is called the Heine-Borel property on the real line or other equivalent properties. The proofs given in this unit straightaway appeal to the completeness property of the red line (Unit 2) namely that any subset of the real line bounded above has least upper bound. These proofs may be slightly longer than the conventional ones but it does not make use of any other theorem except the property of the real line stated above.

As remarked earlier, the properties of continuous functions fail if the intervals are not bounded or closed, that is, the intervals of the type

]a, b[, 1a, b], [a, b[, [a,′∞[, 1a, ∞[,] –∞, a],] –∞, a [or] –∞, ∞[.

Example: Show that the function f defined by f(x) = 3 V x ∈ [0, ∞[is continuous but not bounded.

Solution: The function f being a polynomial function is continuous in [0, ∞[. The domain of the function is an unbounded closed interval. The function is not bounded since the set of values of the function that is the range of the function is $\{x^2 : x \in [0, \infty[\ \} = [0, \infty[$ which is not bounded.

Notes

Example: Show that the function f defined by f(x) = $\frac{1}{x}$ V x ∈] 0, 1[is continuous but not bounded.

Solution: The function f is continuous being the quotient of continuous functions F(x) = 1 and G(x) = x with

$G(x) \neq 0, x \in]0, 1[$

Domain of f is bounded but not a closed interval. The function is not bounded since its range is (1/x : x ∈] 0, l [ ] =]1, ∞[which is not a bounded set.

Example: Show that the function f such that f(x) = x ∀ x ∈]0, 1[is continuous but does not attain its bounds.

Solution: As mentioned the identity function f is continuous in]0, 1[. Here the domain of f is bounded but is not a closed interval. The function f is bounded with least upper bound (1.u.b) = 1 and greatest lower bound (g.l.b) = 0 and both the bounds are not attained by the function, since range of f =]0, 1[.

Example: Show that the function f such that

$$f(x) = \frac{1}{x^2} \quad \forall x \in [0, 1[.$$

is continuous but does not attain its g.l.b.

Solution: The function G given by $G(x) = x^2$ ∀ x ∈]0, 1[is continuous and G(x) ≠ 0 ∀ x ∈]0, 1[therefore its reciprocal function $f(x) = 1/x^2$ is continuous in]0, 1[. Here the domain f is bounded but is not a closed interval.

Further l.u.b. of f does not exist whereas its g.l.b. is 1 which is not attained by f.

Task Show that the function f given by f(x) = sin x, x ∈]0, π/2[is continuous but does not attain any of its bounds.

Task Prove that the function f given by $f(x) = x^2$ ∀ x ∈] –∞, 0[is continuous but does not attain its g.l.b.

We next prove another important property known as the intermediate value property of a continuous function on an interval I. We do not need the assumption that I is bounded and losed. This property justifies our intuitive idea of a continuous function namely as a function f which cannot jump from one value to another since it takes on between any two values f(a) and f(b) all values lying between f(a) and f(b).

Theorem 3: (Intermediate Value Theorem). Let f be a continuous function on an interval containing a and b. If K is any number between f(a) and f(b) then there is a number c, a ≤ c S b such that f(c) = K.

Proof: Either f(a) = f(b) or f(a) < f(b) or f(b) < f(a). If f(a) = f(b) then K = f(a) = f(b) and so c can be taken to be either a or b. We will assume that f(a) < f(b). (The other case can be dealt with similarly.) We can, therefore, assume that f(a) < K < f(b).

Notes

Let S denote the collection of all real numbers x in [a, b] such that $f(x) < K$. Clearly S contains a, so $S \neq \phi$ and b is an upper bound for S. Hence, by completeness property of R, S has least upper bound and let us denote this least upper bound by c. Then $a \leq c \leq b$. We want to show that $f(c) = K$.

Since f is continuous on [a, b], f is continuous at c. Therefore, given $\varepsilon > 0$, there exists a $6 > 0$ such that whenever x is in [a, b] and $|x - c| < 6$, $|f(x) - f(c) (< G$,

i.e., $\quad f(c) - \varepsilon < f(x) < f(c) + \varepsilon$. .. . (4)

If $c \neq b$, we can clearly assume that $c + 6 < b$. Now c is the least upper bound of S. So $c - \delta$ is not 'an upper bound' of S. Hence, there exists a y in S such that $c - 6 < y \leq c$. Clearly $|y - c| < \delta$ and so by (4) above, we have

$$f(c) - \varepsilon < f(y) < f(c) + \varepsilon.$$

Since y is in S, therefore $f(y) < K$. Thus, we get

$$f(c) - S < K$$

If now $c = b$ then $K - \varepsilon < K < f(b) = f(c)$, i.e., $K < f(c) + E$. If $c \neq b$, then $c < b$; then there exists an x such that $c < x < c + 6, 6, x \in [a, b]$ and for this x, $f(x) < f(c) + \varepsilon$ by (4) above. Since $x > c$, $K \leq f(x)$, for otherwise x would be in S which will imply that c is not an upper bound of S. Thus, again we have $K \leq f(x) < f(c) + E$.

In any case,

$$K < f(c) + \varepsilon \qquad ...(6)$$

Combining (5) and (6), we get for every $\varepsilon > 0$

$$f(c) - \varepsilon < K < f(c) + \varepsilon$$

which proves that $K = f(c)$, since ε is arbitrary while K, f(c) are fixed. In fact, when $f(a) < K < f(b)$ and $f(c) = K$, then $a < c < b$.

Corollary 1: If f is a continuous function on the closed interval [a, b] and If f(a) and f(b) have opposite signs (i.e., $f(a) f(b) < 0$), then there is a point x_0 in]a, b[at which f vanishes. (i.e., $f(x_0) = 0$).

Corollary follows by taking $K = 0$ in the theorem.

Corollary 2: Let f be a continuous function defined on a bounded closed interval [a, b] with values in [a, b]. Then there exists a point c in [a, b] such that $f(c) = c$. (i.e., there exists a fixed point c for the function f on [a, b]).

Proof: If $f(a) = a$ or $f(b) = b$ then there is nothing to prove. Hence, we assume that $f(a) \neq a$ and $f(b) \neq b$.

Consider the function g defined by $g(x) = f(x) - x$, $x \in [a, b]$. The function being the difference of two continuous functions, is continuous on [a, b]. Further, since f(a), f(b) are in [a, b], $f(a) > a$ (since $f(a) \neq a$, $f(a) \in [a, b]$) and $f(b) < b$. (Since $f(b) \neq b$, $f(b) \in [a, b]$). So, $g(a) > 0$ and $g(b) < 0$. Hence, by Corollary 1, there exists a c in]a, b[such that $g(c) = 0$, i.e., $f(c) = c$. Hence, there exists ac in [a, b] such that $f(c) = c$.

The above Corollary 1 helps us sometimes to locate some of the roots of polynomials. We illustrate this with the following example.

Example: The equation $x^4 + 2x - 11 = 0$ has a real root lying between 1 and 2.

Solution: The function $f(x) = x^4 + 2x - 11$ is a continuous function on the closed interval [1, 2], $f(1) = -8$ and $f(2) = 9$. Hence, by Corollary 1, there exists an $x_0 \in]1, 2[$ such that $f(x_0) = 0$, i.e., x_0 is a real root of the equation $x^4 + 2x - 11 = 0$ lying in the interval]1, 2[.

Notes

> *Task* Show that the equation $16x^4 + 64x^3 - 32x^2 - 117 = 0$ has a real root > 1.

12.2 Pointwise Continuity and Uniform Continuity

Here you will be introduced with the concept of uniform continuity of a function. The concept of uniform continuity is given in the whole domain of the function whereas the concept of continuity is pointwise that is it is given at a point of the domain of the function. If a function f is continuous at a point a in a set A, then corresponding to a number E > 0, there exists a positive number $\delta(a)$ (we are denoting 6 as $\delta(a)$ to stress that 6 in general depends an the point a chosen) such that $|x - a| < \delta(a)$ implies that $|f(x) - f(a)| < \delta$. The number $\delta(a)$ also depends on E. When the point a varies $\delta(a)$ also varies. We may or may not have a 6 which serves for all points a in A. If we have such a 6 common to all points a in A, then we say that f is uniformly continuous on A. Thus, we have the following definition of uniform continuity.

Definition 1: Uniform Continuity of a Function

Let f be a function defined on a subset A contained in the set R of all reals. If corresponding to any number $\varepsilon > 0$, there exists a number $\delta > 0$ (depending only on G) such that

$$|x - y| < \delta, x, y \in A \Rightarrow |f(x) - f(y)| < \varepsilon^*$$

then we say that f is uniformly continuous on the subset A.

An immediate consequence of the definition of uniform continuity is that uniform continuity in a set A implies pointwise continuity in A. This is proved in the following theorem.

Theorem 4: If a function f is uniformly continuous in a set A, then it is continuous in A.

Proof: Since f is uniformly continuous in A, given a positive number E, there corresponds a positive number 6 such that

$$|x - y| < \delta; x, y \in A \Rightarrow |f(x) - f(y)| < \varepsilon \qquad \text{... (7)}$$

Let a be any point of A. In the above result (1), take y = a. Then we get,

$$|x - a| < \delta; x \in A \Rightarrow |f(x) - f(a)| < \varepsilon$$

which shows that f is continuous at 'a'. Since 'a' is any point of A, it follows that f is continuous in A.

Now we consider some examples.

Example: Show that the function f : R

$f(x) = x$ V $x \in R$,

is uniformly continuous on R

Solution: For a given $\varepsilon > 0$, 6 can be chosen to be ε itself so that

$$|x - y| < 6 = G \Rightarrow |f(x) - f(y)| = |x - y| < \varepsilon.$$

Example: Show that the function f : R – R given by

$f(x) = x^2 \ \forall x \in R$

is not uniformly continuous on R.

Notes

Solution: Let ε be any positive number. Let δ > 0 be any arbitrary positive number. Choose x > ε/δ and y = x + δ/2. Then

$$|x - y| = \frac{\delta}{2} < \delta.$$

$$(\text{fix}) - f(y)| = |x^2 - y^2| = |x + y|\,|x - y|$$

$$= \left(\frac{\delta}{2}\right)|x + y| = \left(\frac{\delta}{2}\right)\left|2x + \frac{\delta}{2}\right|$$

$$> \frac{\delta}{2}\left(\frac{2\varepsilon}{\delta} + \frac{\delta}{2}\right) = \varepsilon + \frac{\delta^2}{4} > \varepsilon$$

That is whatever δ > 0 we choose, there exist real numbers x, y such that |x – y| < δ but |f(x) – f(y)| > G which proves that f is not uniformly continuous.

But we know that f is a continuous function on R.

Example: In the above example if we restrict the domain of f to be the closed interval [–1, 1], then show that f is uniformly continuous on [–1, 1].

Solution: Given E > 0, choose $\delta < \frac{8}{2}$. If |x – y| < δ and x, y ∈[–1, 1],

then using the triangle inequality for | | we get,

$$|f(x) - f(y)| = |x^2 - y^2| = |x + y|\,|x - y|$$

$$< \in (|x| + |y|)$$

$$\leq 2\delta \text{ (since } |x| \leq 1, |y| \leq 1)$$

You should be able to solve the following exercises:

Task

1. Show that $f(x) = x^n$, n > 1 is not uniformly continuous on R even though for each a > 1, it is a continuous function on R.
2. Show that the function $f(x) = \frac{1}{x}$ for 0 < x < 1 is continuous for every x but not uniformly on]0, 1[.
3. Show that the function $f(x) = \sin\frac{1}{x}$ is not uniformly continuous on the interval]0, 1[even though it is continuous in that interval.
4. Show that f(x) = cx where c is a fixed non-zero real number is a uniformly continuous function on R.

We have seen that the function defined by f(x) = 1/x on the open interval]0, 1 [is not uniformly continuous on]0, 1[even though it is a continuous function on]0, 1[. Similarly the function f defined as $f(x) = x^2$ is continuous on the entire real line R but is not uniformly continuous on R.

However, if we restrict the domain of this function to the bounded closed interval [–1, 1], then it is uniformly continuous. This property is not a special property of the function f, where $f(x) = x^2$

but is common to all continuous functions defined on bounded closed intervals of the real line. We prove it in the following theorem.

Theorem 5: If f is a continuous function on a bounded and closed interval [a, b] then f is uniformly continuous on [a, b].

Proof: Let f be a continuous function defined on the bounded closed interval [a, b]. Let S be the set of all real numbers c in the interval [a, b] such that for a given $\varepsilon > 0$, there exists positive number d_c such that for points x_1, x_2 belonging to closed interval [a, c],

$|f(x_1) - f(x_2)| < \varepsilon$ whenever $|x_2 - x_2| < d_r$.

(In other words f is uniformly continuous on the interval [a, c]. Clearly a ε S so that S is non-empty. Also b is an upper bound of S. Prom completeness property of the real line S has least upper bound which we denote by k. $k \le b$.

f is continuous at k. Hence given E > 0, there exists positive real number d_k such that

$$|f(x) - f(k)| < \varepsilon/2 \text{ whenever } |x - k| < d_k \quad ...(8)$$

Since k is the least upper bound of S, $k - \frac{1}{2}$ is not an upper bound of S.

Therefore there exists a point $c \in S$ such that

$$k - 1/2\, d_k < c \le k. \quad ...(9)$$

Since $c \in S$; from the definition of S we see that there exists d_c such that

$$|f(x_1) - f(x_2)| < \varepsilon \text{ whenever } |x_1 - x_2| < d_c, x_1, x_2 \in [a, c], \quad ...(10)$$

Let $\quad d = \min((1/2)\, d_k, d_c)$ and $b' = \min. (k + (1/2)\, d_k, b)$.

Now let $x_1, x_2 \in [a, b']$ and $|x_1 - x_2|$. Then if $x_1, x_2 \in [a, c]$, $|x_1 - x_2| < d \le d_c$ by the choice of d and d_c, then $|f(x_1) - f(x_2)| < \varepsilon$ by (10). If one of x_1 x_2 is not in fa, cl, then both x_1, x_2 belong to the interval $]k - d_k, k + d_k[$. For $x_1 \notin [a, c]$, implies $b' \ge x_1 > c > k - (1/2)d_k > k - d_k$ by (9) above. This means $x_1 \le b'$ implies $x_1 \le k + (1/2)d_k < k - d_k$ by the choice of b′. i.e.

$$k - d_k < k - (1/2)\, d_k < x_1 < k + (1/2)d_k < k + d_k \quad ...(11)$$

$|x_1 - x_2| < d$ implies that $x_1 - (1/2)\, d_k < x_2 < x_1 + (1/2)d_k$ since $d \le (1/2)d_k$ by this choice of d. Thus we get from (11) above that

$$|x_1 - x_2| < x_1 - (1/2)d_k < x_2 < x_1 + (1/2)d_k < k + \left(\frac{1}{2}\right) d_k + \frac{1}{2} d_k = k + d_k \quad ...(12)$$

Then (11) and (12) show that $x_1, x_2 \in]k - d_k, k + d_k[$.

Thus we get that $|x_1 - k| < d_k$ and $|x_2 - k| < d_y$, which in turn implies, by (8) above, that $|f(x_1) - (k)| < \varepsilon/2$ and $|f(x_2) - f(k)| < \varepsilon/2$.

Thus $|f(x_1) - f(x_2)| < |f(x_1) - f(k)| + |f(k) - f(x_2)| < \varepsilon/2 + \varepsilon/2 = E$. In other words, if $|x_1 - x_2| <$ d and x_1, x_2 are in [a, b′] then $|f(x_1) - f(x_2)| < E$ which proves that $b' \in S$ i.e. $b' \le k$. But $k \le b'$ by the choice of b′ since $k \le k + (1/2)\, d_k$ and k a b. Thus we get that k = b′. This can happen only when k = b. For if k < b. i.e. $k = b' = \min(k + (l/2)\, d_k, b) < b$, then it implies that $\min(k + (l/2)\, d_k, b) = (k + (1/2)\, d_k = b'$, where $b' \varepsilon S$ i.e. $k + (1/2)\, d_k$ is in S and is greater than k which is a contradiction to the fact that k is the l.u.b of S. Thus we have shown that $k = b\ \varepsilon\ S$. In other words there exists a positive number d_k (corresponding to b) such that $|x_1 - x_2| < d_k$, $x_1, x_2 \in [a, b]$ implies $|f(x_1) - f(x_2)| < \varepsilon$. Therefore f is uniformly continuous in [a, b].

You may note that uniform continuity always implies continuity but not conversely. Converse is true when continuity is in the bounded closed interval.

Notes

Before we end this unit, we state a theorem without proof regarding the continuity of the inverse function of a continuous function.

Theorem 6: Inverse Function Theorem

Let f : I – J be a function which is both one-one and onto. If f is continuous on I, then $f^{-1} : J \to I$ is continuous on J. For example the function.

$$f : \left[\frac{-\pi}{2}, \frac{\pi}{2}\right] [-1, 1] \text{ defined by}$$

$$f(x) = \sin x,$$

is both one-one and onto. Besides f is continuous on $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$. Therefore, by Theorem 6, the function

$$f^{-1} : -11 \to \left[\frac{-\pi}{2}, \frac{\pi}{2}\right] \text{ defined by}$$

$$f^{-1}(x) = \sin -x$$

is continuous on [–1, 1].

Self Assessment

1. Give an example of the following:

   (i) A function which is nowhere continuous but its absolute value is everywhere continuous.

   (ii) A function which is continuous at one point only.

   (iii) A linear function which is continuous and satisfies the equation f(x + y) = f(x) + f(y).

   (iv) Two uniform continuous functions whose product is not uniformly continuous.

2. State whether the following are true or false:

   (i) A polynomial function is continuous at every point of its domain.

   (ii) A rational function is continuous at every point at which it is defined.

   (iii) If a function is continuous, then it is always uniformly continuous.

   (iv) The functions e^x and log x are inverse functions for x > 0 and both are continuous for each x > 0.

   (v) The functions cos x and $\cos^{-1} x$ are continuous for all real x.

   (vi) Every continuous function is bounded.

   (vii) A continuous function is always monotonic.

   (viii) The function sin x is monotonic as well as continuous for $s \in [0, 3]$

   (ix) The function cos x is continuous as well as monotonic for every $x \in R$.

   (x) The function $|x|, x \in R$ is continuous.

Notes

12.3 Summary

- In this unit you have been introduced to the properties of continuous functions on bounded closed intervals and you have seen the failure of these properties if the intervals are not bounded and closed. These properties have been studied. It has been proved that if a function f is continuous on a bounded and closed interval, then it is bounded and it also attains its bounds. In the same section we proved the Intermediate Value Theorem that is if f is continuous on an interval containing two points a and b, then f takes every value between f(a) and f(b). The notion of uniform continuity is discussed. We have proved that if a function f is uniformly continuous in a set A, then it is continuous in A. But converse is not true. It has been proved that if a function is continuous on a bounded and closed interval, then it is uniformly continuous in that interval. These properties fail if the intervals are not bounded and closed. This has been illustrated with a few examples.

12.4 Keywords

Bounded Function: A function f continuous on a bounded and closed interval [a, b] is necessarily a bounded function.

Boundedness: If f is a continuous function on the bounded closed interval [a, b] then there exists points x_1 and x_2 in [a, b] such that $f(x_1)$ S $f(x) \le f(x_2)$ for all $x \in [a, b]$.

Intermediate Value Theorem: Let f be a continuous function on an interval containing a and b. If K is any number between f(a) and f(b) then there is a number c, $a \le c$ S b such that f(c) = K.

12.5 Review Questions

1. Find the limits of the following functions:

   (i) $f(x) = x \cos \frac{1}{x}$, $x \ne 0$, as $x \to 0$.

   (ii) $f(x) = \frac{|x|}{x}$, $x \ne 0$, as $x \to \infty$.

   (iii) $f(x) = \frac{\sin x}{x}$, $x \ne 0$, as $x \to \infty$.

2. For the following functions, find the limit, if it exists:

   (i) $f(x) = \frac{\sqrt{x} - \sqrt{b}}{x - b}$ for $x \ne b$ where $b > 0$, as $x \to b$

   (ii) $f(x) = \frac{1}{1 + e^{-1/x}}$ for $x \ne 0$, as $x \to 0$

   (iii) $f(x) = \begin{cases} 3 - x \text{ when } x \le 1 \\ 2x \text{ when } x > 1 \end{cases}$ as $x \to 1$.

3. Test whether or not the limit exists for the following:

   (i) $f(x) = \begin{cases} 3 - x \text{ when } x > 1 \\ 1 \text{ when } x = 1, \\ 2x \text{ when } x < 1 \end{cases}$ as $x \to 1$.

Notes

(ii) f(x) = $\dfrac{x^2-4}{x^2+4}$, x ∈ R, as x → 1.

(iii) f(x) = $\dfrac{\sqrt{4+x}-2}{x}$, x ≠ O as x → 0.

(iv) f(x) = $\dfrac{1}{x-1}\left(\dfrac{1}{x+3}-\dfrac{2}{3x+5}\right)$ as x → 1.

4. Discuss the continuity of the following functions at the points noted against each.

(i) f(x) = $\begin{cases} x^2 \text{ for } x \neq 1 \\ 0 \text{ for } x = 1 \end{cases}$ as x → 1.

(ii) f(x) = $\begin{cases} 1 \text{ for } 0 \leq x < 1 \\ 0 \text{ otherwise} \end{cases}$ as x → 1.

(iii) f(x) = $\dfrac{x^2-4}{x-1}$ when x ≠ 1.

f(1) = 2

as x → 1.

(iv) f(x) = $\begin{cases} (1+x)^{1/x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$ as x → 1.

(v) f(x) = $\begin{cases} x^2 \sin\dfrac{1}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$ as x → 1.

5. Show that the function f : R → R defined as

$\dfrac{1}{1+|x|}$

does not attain its infimum.

6. Show that the function f : R → R such that

f(x) = x is not bounded but is continuous in [1, ∞[.

7. Which of the following functions are uniformly continuous in the interval noted against each? Give reasons.

(i) f(x) = tan x, x ∈[0, π/4J

(ii) f(x) = $\dfrac{1}{x^2-3}$ on [1, 4].

Answers: Self Assessment

1. (i) $\begin{cases} f(x) = 1 & \text{if } x \text{ is rational} \\ = -1 & \text{if } x \text{ is irrational} \end{cases}$

(ii) $\begin{cases} f(x) = x & \text{if } x \text{ is rational} \\ = -x & \text{if } x \text{ is irrational} \end{cases}$

the only point of continuity is 0.

Notes

(iii) $f(x) = Cx, \ \forall x \in R$ where C is a fixed constant.

(iv) $f(x) = x, g(x) = \sin x, \ \forall x \in R$

Both f(x) and g(x) are uniformly continuous but their product

$f(x)\, g(x) = x \sin x$

is not uniformaly continuous on R.

2. (i) True (ii) True

(iii) False (iv) True

(v) True (vi) False

(vii) False (viii) True

(ix) False (x) True

12.6 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 13: Discontinuities and Monotonic Functions

Notes

CONTENTS

Objectives

Introduction

13.1 Discontinuous Functions

13.2 Classification of Discontinuities

13.3 Monotone Function

13.4 Discontinuities of Monotone Functions

13.5 Discontinuities of Second Kind

13.6 Summary

13.7 Keywords

13.8 Review Questions

13.9 Further Readings

Objectives

After studying this unit, you will be able to:

- Define Discontinuous Functions
- Describe Classification of Discontinuities
- Explain Monotone Function
- Describe the Discontinuities of Monotone Functions
- Discuss the Discontinuities of Second Kind

Introduction

In mathematics, a monotonic function (or monotone function) is a function that preserves the given order. This concept first arose in calculus, and was later generalized to the more abstract setting of order theory. In calculus, a function *f* defined on a subset of the real numbers with real values is called monotonic (also monotonically increasing, increasing or non-decreasing), if for all *x* and *y* such that *x* d" *y* one has *f*(*x*) d" *f*(*y*), so *f* preserves the order. Likewise, a function is called monotonically decreasing (also decreasing or non-increasing) if, whenever *x* d" *y*, then *f*(*x*) e" *f*(*y*), so it *reverses* the order.

13.1 Discontinuous Functions

If a function fails to be continuous at a point c, then the function is called **discontinuous** at c, and c is called a **point of discontinuity**, or simply a discontinuity.

Notes

Task **Consider the following functions:**

1. $k(x) = \begin{cases} \dfrac{x^2 - 9}{x - 3} & \text{if } x \neq 3 \\ 1, & \text{if } x = 3 \end{cases}$

2. $h(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{If } x = 0 \\ -1, & \text{if } x < 0 \end{cases}$

3. $f(x) = \begin{cases} \sin(1/x), & \text{if } x \neq 0 \\ 0, & \text{If } x \neq 0 \end{cases}$

4. $g(x) = \begin{cases} 1, & \text{if } x = \text{rational} \\ 0, & \text{If } x = \text{irrational} \end{cases}$

Which of these functions, without proof, has a 'fake' discontinuity, a 'regular' discontinuity, or a 'difficult' discontinuity?

13.2 Classification of Discontinuities

Suppose f is a function with domain D and c ∈ D is a point of discontinuity of f.

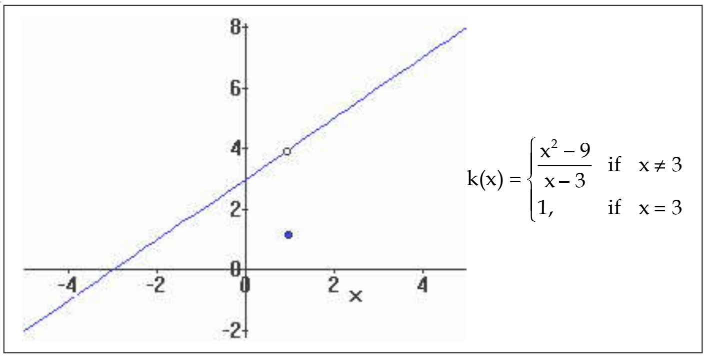
1. If $\lim_{x \to c} f(x)$ exists, then c is called removable discontinuity.

2. If $\lim_{x \to c} f(x)$ does not exist, but both $\lim_{x \to c^-} f(x)$ and $\lim_{x \to c^+} f(x)$ exist, then c is called a discontinuity of the first kind, or jump discontinuity.

3. If either $\lim_{x \to c^-} f(x)$ or $\lim_{x \to c^+} f(x)$ does not exist, then c is called a discontinuity of the second kind, or essential discontinuity.

Example: Prove that k(x) has a removable discontinuity at x = 3, and draw the graph of k(x).

Solution:



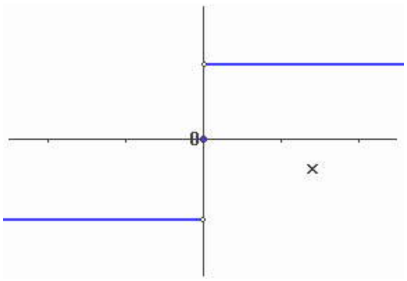
We can easily check that the limit as *x* approaches *3* from the right and from the left is equal to *4*. Hence, the limit as *x* approaches *3* exists, and therefore the function has a removable

Notes

discontinuity at $x = 3$. If we define $k(3) = 4$ instead of $k(3) = 1$ then the function in fact will be continuous on the real line

Example: Prove that $h(x)$ has a jump discontinuity at $x = 0$, and draw the graph of $h(x)$

Solution:

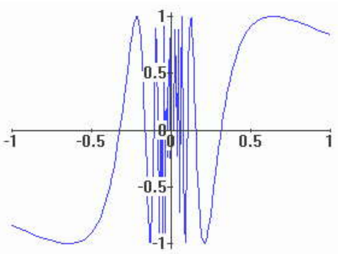


$$h(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{If } x = 0 \\ -1, & \text{if } x < 0 \end{cases}$$

It is easy to see that the limit of h(x) as x approaches 0 from the left is –1, while the limit of h(x) as x approaches 0 from the right is +1. Hence, the left and right handed limits exist and are not equal, which makes x = 0 a jump discontinuity for this function.

Example: Prove that f(x) has a discontinuity of second kind at x = 0

Solution:



$$f(x) = \begin{cases} \sin(1/x), & \text{if } x \neq 0 \\ 0, & \text{If } x \neq 0 \end{cases}$$

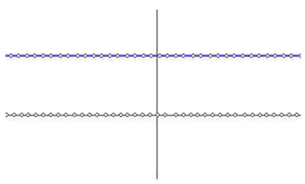
This function is more complicated. Consider the sequence $x_n = 1/(2n\pi)$. As n goes to infinity, the sequence converges to zero from the right. But $f(x_n) = \sin(2n\pi) = 0$ for all k. On the other hand, consider the sequence $x_n = 2/(2n + 1)$. Again, the sequence converges to zero from the right as n goes to infinity. But this time $f(x_n) = \sin((2n + 1)/2)$ which alternates between +1 and –1. Hence, this limit does not exist. Therefore, the limit of f(x) as x approaches zero from the right does not exist.

Since f(x) is an odd function, the same argument shows that the limit of f(x) as x approaches zero from the left does not exist.

Therefore, the function has an essential discontinuity at x = 0.

Example: What kind of discontinuity does the function g(x) have at every point (with proof).

Solution:



$$g(x) = \begin{cases} 1, & \text{if } x = \text{rational} \\ 0, & \text{If } x = \text{irrational} \end{cases}$$

Notes

This function is impossible to graph. The picture above is only a poor representation of the true graph. Nonetheless, take an arbitrary point x_0 on the real axis. We can find a sequence $\{x_n\}$ of rational points that converge to x_0 from the right. Then $g(x_n)$ converges to 1. But we can also find a sequence $\{x_n\}$ of irrational points converging to x_0 from the right. In that case $g(x_n)$ converges to 0. But that means that the limit of g(x) as x approaches x_0 from the right does not exist. The same argument, of course, works to show that the limit of g(x) as x approaches x_0 from the left does not exist. Hence, x_0 is an essential discontinuity for g(x).

It is clear that any function is either continuous at any given point in its domain, or it has a discontinuity of one of the above three kinds. It is also clear that removable discontinuities are 'fake' ones, since one only has to define $f(c) = \lim_{x \to c} f(x)$ and the function will be continuous at c. Of the other two types of discontinuities, the one of second kind is hard. Fortunately, however, discontinuities of second kind are rare, as the following results will indicate.

13.3 Monotone Function

A function f is monotone increasing on (a, b) if $f(x) \leq f(y)$ whenever $x < y$. A function f is monotone decreasing on (a, b) if $f(x) \geq f(y)$ whenever $x < y$.

A function f is called monotone on (a, b) if it is either always monotone increasing or monotone decreasing.

Some basic applications and results

The following properties are true for a monotonic function $f : R \to R$:

- f has limits from the right and from the left at every point of its domain;
- f has a limit at infinity (either ∞ or $-\infty$) of either a real number, ∞, or $-\infty$.
- f can only have jump discontinuities;
- f can only have countably many discontinuities in its domain.

These properties are the reason why monotonic functions are useful in technical work in analysis. Two facts about these functions are:

- if f is a monotonic function defined on an interval I, then f is differentiable almost everywhere on I, i.e. the set of numbers x in I such that f is not differentiable in x has Lebesgue measure zero.
- if f is a monotonic function defined on an interval [a, b], then f is Riemann integrable.

An important application of monotonic functions is in probability theory. If *X* is a random variable, its cumulative distribution function

$$F_X(x) = \text{Prob}\,(X \leq x)$$

is a monotonically increasing function.

A function is *unimodal* if it is monotonically increasing up to some point (the *mode*) and then monotonically decreasing.

Notes

Note If f is increasing if -f is decreasing, and visa versa. Equivalently, f is increasing if

- $f(x)/f(y) \leq 1$ whenever $x < y$
- $f(x) - f(y) \leq 0$ whenever $x < y$

These inequalities are often easier to use in applications, since their left sides take a very nice and simple form. Next, we will determine what type of discontinuities monotone functions can possibly have. The proof of the next theorem, despite its surprising result, is not too bad.

13.4 Discontinuities of Monotone Functions

If f is a monotone function on an open interval (a, b), then any discontinuity that f may have in this interval is of the first kind.

If f is a monotone function on an interval [a, b], then f has at most countably many discontinuities.

Proof: Suppose, without loss of generality, that f is monotone increasing, and has a discontinuity at x_0. Take any sequence x_n that converges to x_0 from the left, i.e. x_n x_0. Then $f(x_n)$ is a monotone increasing sequence of numbers that is bounded above by $f(x_0)$. Therefore, it must have a limit. Since this is true for every sequence, the limit of f(x) as x approaches x_0 from the left exists. The same prove works for limits from the right.

Notes This proof is actually not quite correct. Can you see the mistake? Is it really true that if x_n converges to x_0 from the left then $f(x_n)$ is necessarily increasing? Can you fix the proof so that it is correct?

As for the second statement, we again assume without loss of generality that f is monotone increasing. Define, at any point c, the jump of f at x = c as:

$$j(c) = \lim_{x \to c^+} f(x) - \lim_{x \to c^-} f(x)$$

Note that j(c) is well-defined, since both one-sided limits exist by the first part of the theorem. Since f is increasing, the jumps j(c) are all non-negative. Note that the sum of all jumps can not exceed the number f(b) – f(a). Now let J(n) be the set of all jumps c where j(c) is greater than 1/n, and let J be the set of all jumps of the function in the interval [a, b]. Since the sum of jumps must be smaller than f(b) – f(a), the set J(n) is finite for all n. But then, since the union of all sets J(n) gives the set J, the number of jumps is a countable union of finite sets, and is thus countable.

This theorem also states that if a function wants to have a discontinuity of the second kind at a point x = c, then it can not be monotone in any neighbourhood of c.

13.5 Discontinuities of Second Kind

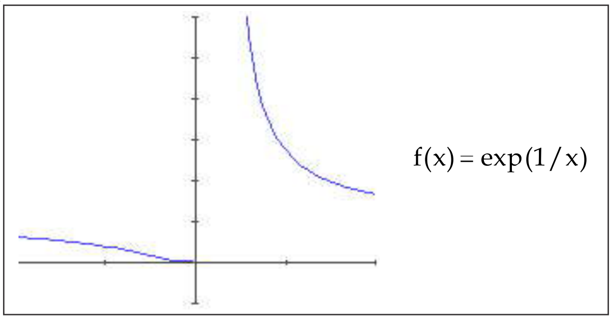
If f has a discontinuity of the second kind at x = c, then f must change from increasing to decreasing in every neighbourhood of c.

Proof: Suppose not, i.e. f has a discontinuity of the second kind at a point x = c, and there does exist some (small) neighbourhood of c where f, say, is always decreasing. But then f is a monotone function, and hence, by the previous theorem, can only have discontinuities of the first kind. Since that contradicts our assumption, we have proved the corollary.

Notes

In other words, f must look pretty bad if it has a discontinuity of the second kind.

Example: What kind of discontinuity does the function f(x) = exp(1/x) have at x = 0?

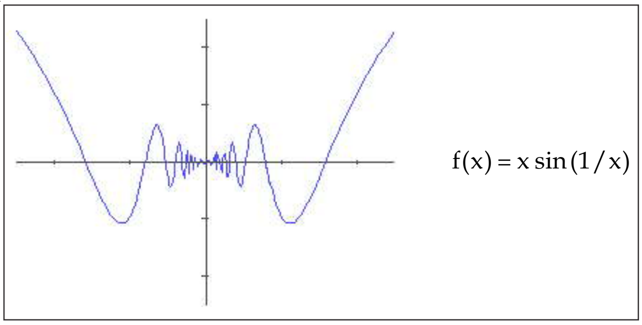


As x approaches zero from the right, 1/x approaches positive infinity. Therefore, the limit of f(x) as x approaches zero from the right is positive infinity.

As x approaches zero from the left, 1/x approaches negative infinity. Therefore, the limit of f(x) as x approaches zero from the left is zero.

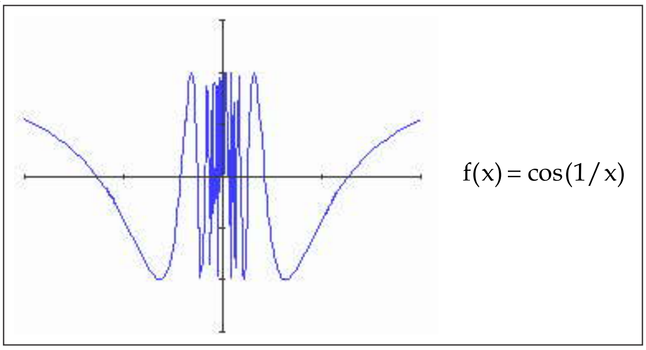
Since the right-handed limit fails to exist, the function has an essential discontinuity at zero.

Example: What kind of discontinuity does the function f(x) = x sin(1/x) have at x = 0?



Since $|x \sin(1/x)| < |x|$, we can see that the limit of f(x) as x approaches zero from either side is zero. Hence, the function has a removable discontinuity at zero. If we set f(0) = 0 then f(x) is continuous.

Example: What kind of discontinuity does the function f(x) = cos (1/x) have at x = 0?



Notes

By looking at sequences involving integer multiples of π or π/2 we can see that the limit of f(x) as x approaches zero from the right and from the left both do not exist. Hence, f(x) has an essential discontinuity at x = 0.

Self Assessment

Fill in the blanks:

1. The class of monotonic functions consists of both the ………………..
2. ……………….. have no discontinuities of second kind.
3. Let f be monotonic on (a; b), then the set of points of (a; b) at which f is ………………… at most countable.
4. A function *f*(*x*) is said to be ……………….. over an interval (*a*, *b*) if the derivatives of all orders of *f* are nonnegative at all points on the interval.
5. The term ……………….. can also possibly cause some confusion because it refers to a transformation by a strictly increasing function

13.6 Summary

- If a function fails to be continuous at a point c, then the function is called discontinuous at c, and c is called a point of discontinuity, or simply a discontinuity.
- In calculus, a function f defined on a subset of the real numbers with real values is called monotonic (also monotonically increasing, increasing or non-decreasing), if for all x and y such that x ≤ y one has f(x) ≤ f(y), so f preserves the order (see Figure 1).
- Suppose f is a function with domain D and ∈ D is a point of discontinuity of f.

  if $\lim_{x \to c} f(x)$ exists, then c is called removable discontinuity.

  if $\lim_{x \to c} f(x)$ does not exist, but both $\lim_{x \to c^-} f(x)$ and $\lim_{x \to c^+} f(x)$ exit, then c is called a discontinuity of the first kind, or jump discontinuity

  if either $\lim_{x \to c^-} f(x)$ or $\lim_{x \to c^+} f(x)$ does not exist, then c is called a discontinuity of the second kind, or essential discontinuity

  If f is a monotone function on an open interval (a, b), then any discontinuity that f may have in this interval is of the first kind.

  If f is a monotone function on an interval [a, b], then f has at most countably many discontinuities.

13.7 Keywords

Monotonic Transformation: The term monotonic transformation can also possibly cause some confusion because it refers to a transformation by a strictly increasing function.

Monotonically Decreasing: A function is called monotonically decreasing (also decreasing or non-increasing) if, whenever x ≤ y, then f(x) ≥ f(y), so it reverses the order.

Monotonic Function: In mathematics, a monotonic function (or monotone function) is a function that preserves the given order.

Notes

13.8 Review Questions

1. Define Discontinuous Functions.
2. Describe Classification of Discontinuities.
3. Explain Monotone Function.
4. Describe the Discontinuities of Monotone Functions.
5. Discuss the Discontinuities of Second Kind.

Answers: Self Assessment

1. Increasing and decreasing functions
2. Monotonic functions
3. Discontinuous
4. Absolutely monotonic
5. Monotonic transformation

13.9 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol : Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik : Mathematical Analysis.

H.L. Royden : Real Analysis, Ch. 3, 4.

Unit 14: Sequences and Series of Functions

Notes

CONTENTS

Objectives

Introduction

14.1 Sequences of Functions

14.2 Uniform Convergence

14.3 Series of Functions

14.4 Summary

14.5 Keywords

14.6 Review Questions

14.7 Further Readings

Objectives

After studying this unit, you will be able to:

- Define sequence and series of functions
- Distinguish between the pointwise and uniform convergence of sequences and series of functions
- Know the relationship of uniform convergence with the notions of continuity, differentiability, and integrability

Introduction

In earlier unit, we have studied about, convergence of the infinite series of real numbers. In this unit, we want to discuss sequences and series whose members are functions defined on a subset of the set of real numbers. Such sequences or series are known as sequences or series of real functions. You will be introduced to the concepts of pointwise and uniform convergence of sequences and series of functions. Whenever they are convergent, their limit is a function called limit function. The question arises whether the properties of continuity, differentiability, integrability of the members of a sequence or series of functions are preserved by the limit function. We shall discuss this question also in this unit and show that these properties are preserved by the Uniform convergence and not by the pointwise convergence.

14.1 Sequences of Functions

As you have studied that a sequence is a function from the set N of natural numbers to a set B. In that unit, sequences of real numbers have been considered in detail. You may recall that for sequences of real numbers, the set B is a sub-set of real numbers. If the set B is the set of real functions defined on a sub-set A of R, we get a sequence called sequence of functions. We define it in the following way:

Definition 1: Sequence of Functions

Let A be a non-empty sub-set of R and let B be the set of all real functions each defined on A. A mapping from the set N of natural numbers to the set B of real functions is called a sequence of functions.

Notes

The sequences of functions are denoted by (f_n), (g) etc., it (f_n) is a sequence of functions defined on A, then its members f_1, f_2, f_3,.... are real functions with domain as the set A. These are also called the terms of the sequence (f_n).

Example: Let $f_n(x) = x^n$, n = 1, 2, 3,...., where $x \in A = \{x : 0 \le x \le I\}$. Then (f_n) is a sequence of functions defined on the closed interval [0, 1].

Similarly consider (f_n), where $f_n(x) = \sin {}_n x$, n = 1, 2, 3,.... $x \in R$. Then $\{f_n\}$, is a sequence of functions defined on the set R of real numbers.

Suppose (f_n) is a sequence of functions defined on a set A and we fix a point x of A, then the sequence $(f_n(x))$, formed by the values of the members of (f_n), is a sequence of real numbers. This sequence of real numbers may be convergent or divergent. For example suppose that $f_n(x) = x^n$, $x \in [-1, 1]$. If we consider the point $x = \frac{1}{2}$, then the sequence $(f_n(x))$ is $((\frac{1}{2})^n)$ which converges to 0. If we take the point x = –1, the sequence $(f_n(x))$ is the constant sequence (1, 1, 1.....) which converges to 1. If x = –1, the sequence (f, (x)) is (–1, I, –1, 1,....) which is divergent.

Thus, you have seen that the sequence $(f_n(x))$ may or may not be convergent. If for a sequence (f_n) of functions defined on a set A, the sequence of numbers $(f_n(x))$ converges for each x in A, we get a function f with domain A whose value f(x) at any point x of A is $\lim_{n\to\infty}$ fn(x). In this case (f_n) is said to be pointwise convergent to f. We define it in the following way:

Definition 2: Pointwise Convergence

A sequence of functions (f_n) defined on a set A is said to be convergent pointwise to f if for each x in A, we have $\lim f_n(x) = f(x)$. Generally, we write $f_n \to f$ (pointwise) on A.

or $\lim_{n\to\infty} f_n(x) = f(x)$ pointwise on A. Also f is called pointwise limit or limit function of (f_n) on A.

Equivalently, we say that a sequence $\{f_n\}$ converges to f pointwise on the set A if, for each $\varepsilon > 0$ and each $x \in A$, there exists a positive integer in depending both on $\in$ and x such that

$$|f_n(x) - f(x)| < \varepsilon, \text{ whenever } n \ge m,$$

Now we consider some examples.

Example: Show that the sequence (f_n) where $f_n(x) = x^n$, $x \in [0, 1]$ is pointwise convergent. Also find the limit.

Solution: If $0 \le x < 1$, then $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = 0$.

If x = 1, then $\lim_{n\to\infty} f(x) = \lim_{n\to\infty} 1 = 1$.

Thus (f,) is pointwise convergent to the limit function f where f(x) = 0 for $0 \le x \le 1$ and f(x) = 1 for x = 1.

Task Show that the sequence of functions (f_n) where $f_n(x) = x^n$, for $x \in [-1, 1]$ is not pointwise convergent.

Notes

Example: Define the function f_n, n = 1, 2,....., as follows:

$$f_n(x) = \begin{cases} 0, & \text{if } x = 0 \\ 2n^2 x, & \text{if } 0 < x < \frac{1}{2n} \\ 2n - 2n^2 x, & \text{if } \frac{1}{2n} \leq x \leq \frac{1}{2n} \\ 0, & \text{if } \frac{1}{n} < x \leq 1 \end{cases}$$

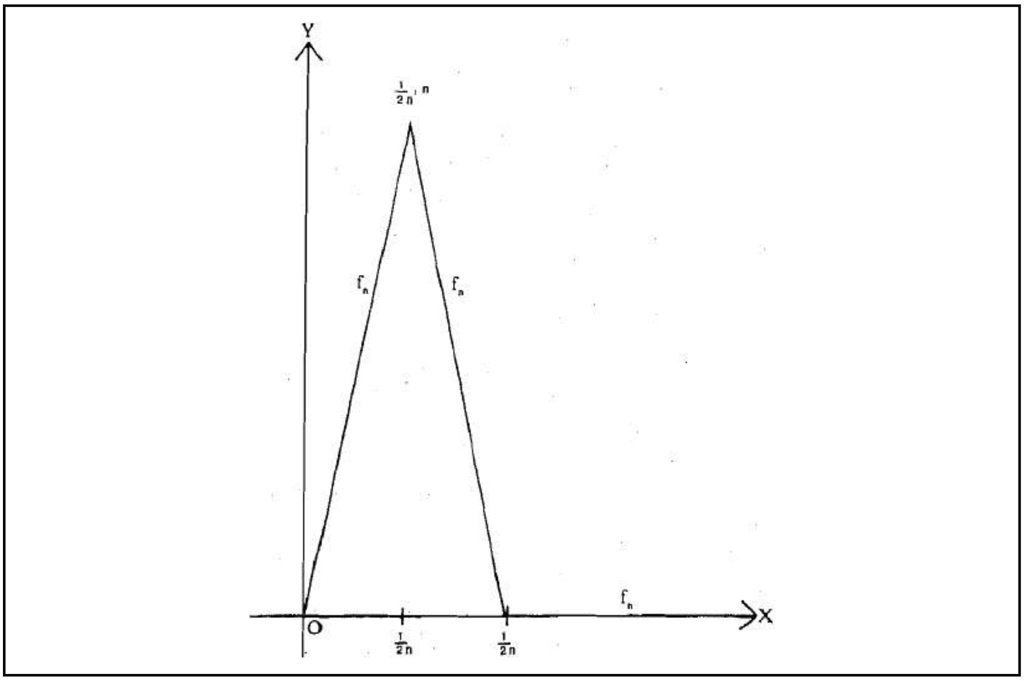
Show that the sequence (f_n) is pointwise convergent.

Solution: The graph of function f_n looks as shown in the Figure.

When x = 0, $f_n(x) = 0$ for n = 1, 2,.....

Therefore, the sequence $(f_n(0))$ tends to 0.

If x is fixed such that 0 < XIP; then choose m large enough so that $\frac{1}{m} < x$ or $m > \frac{1}{X}$. Then $f_m(x) = f_{m+1}(x) = ... = 0$. Consequently the sequence $(f_n(x)) \to 0$ as n – 3 ∞.



Thus, we see that $f_n(x)$ tends to 0 for every x in $0 < x \leq 1$ and consequently (f_n) tends pointwise to f where f(x) = 0 $\forall x \in [0, I]$.

Example: Consider the sequence of functions f_n defined by $f_n(x) = \cos nx$ for $-\infty < x < \infty$ i.e. $x \in R$. Show that the sequence is not convergent pointwise for every real x.

Solution: If x = π/4 then $(f_n(x))$ is the sequence

(1/√2, 0, – 1/√2, – 1, – 1/√2, 0,.....) which is not convergent.

Notes

Task Show that the sequence (f_n) where $f_n(x) = \frac{\sin nx}{\sqrt{n}}$, $x \in R$, is pointwise convergent. Also find the pointwise limit.

If the sequence of functions (f_n) converges pointwise to a function f on a subset A of R, then the following question arises: "If each member of (f_n) is continuous, differentiable or integrable, is the limit function f also continuous; differentiable or integrable?". The answer is no if the convergence is only pointwise. For instance each of the functions f_n is continuous (in fact uniformly continuous) but the sequence of these functions converges to a limit function f(x)

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

which is not continuous. Thus, the pointwise convergence does not preserve the property of continuity. To ensure the passage of the properties of continuity, differentiability or integrability to the limit function, we need the notion of uniform convergence which we introduce in the next section.

14.2 Uniform Convergence

From the definition of the convergence of the sequence or real numbers, it follows that the sequences (f_n) of functions converges pointwise to the function f on A if and only if for each $x \in A$ and for every number $\in > 0$, there exists a positive integer m such that

$$|f_n(x) - f(x)| < \in \text{ whenever } n \ge m.$$

Clearly for a given sequence $(f_{,,})$ of functions, this m will, in general, depend on the given $\in$ and the point x under consideration. Therefore it is, sometimes, written as m $(\in, x)$. The following example illustrates this point.

Example: Define $f_n(x) = \frac{x}{n}$ for $< x < \infty$.

For each fixed x the sequence $(f_n(x))$ clearly converges to zero. For a given $\in > 0$, we must show the existence of an m, such that for all $n \ge m$,

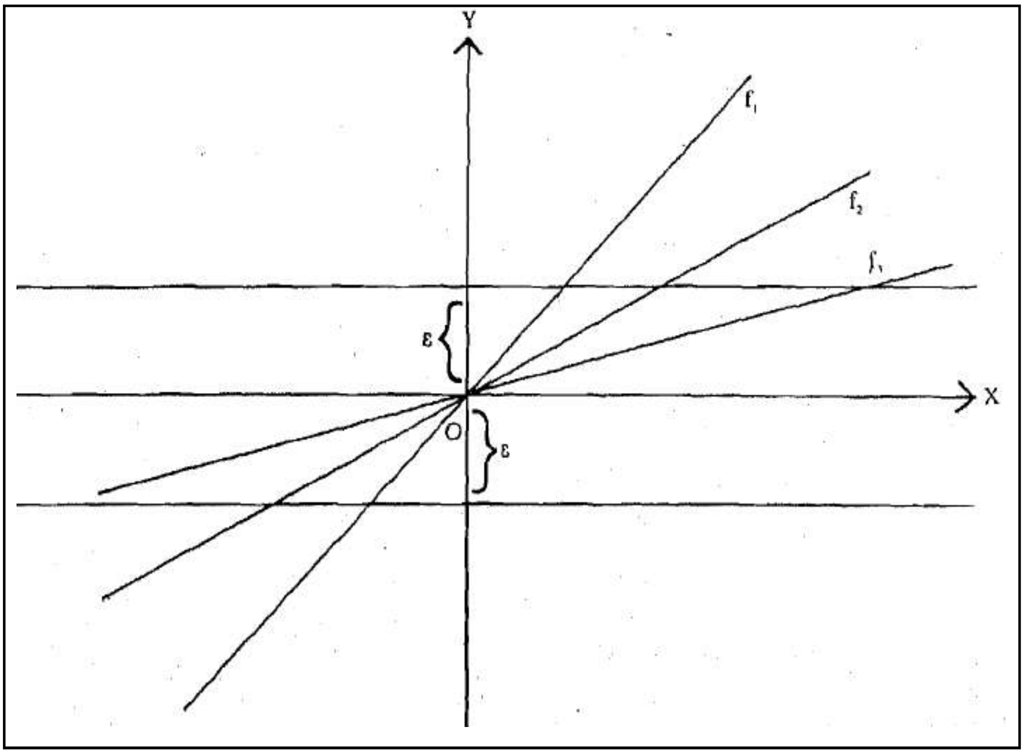
$$|f_n(x) - f(x)| = \left|\frac{x}{n} - 0\right| = \frac{|x|}{n} < \in$$

This can be achieved by choosing $m = \left[\frac{|x|}{\in}\right] + 1$ where $\left[\frac{|x|}{\in}\right]$ denotes the integral part of $\frac{|x|}{\in}$ (i.e. the integer m is next to $|x|$ in the real line). Clearly this choice of m depends both on $\in$ and x.

For example, let $\varepsilon = \frac{1}{10^3}$. If $x = \frac{1}{10^3}$ then $\frac{|x|}{\in} = 1$ and, so, m can be chosen to be 2. If x = 1, then $|x| = 10^3$ and, so, m should be larger than 10^6. Note that it is impossible to find an m that serve for all x. For, if it were, then $\frac{|x|}{m} > E$, for all x,

Consequently $|x|$ is smaller than ε m, which is not possible. Geometrically, the f_n's can be described as shown in the Figure.

Notes



By putting $y = f_n(x)$, we see that $y = \frac{1}{?} x$ is the line with slope $\frac{1}{n}$. f_n is the line y = x with slope 1, f_2 is the line with slope $\frac{1}{2}$ and so on. As n tends to ∞, the lines approach the X-axis. But if we take any strip of breadth 2ε around X-axis, parallel to the X-axis as shown in the figure, it is impossible to find a stage m such that all the lines after the stage m, i.e. f_m, f_{m+1},.... lie entirely in this strip.

If it is possible to find m which depends only on ε but is independent of the point x under consideration, we say that (f_n) is uniformly convergent to f. We define uniform convergence as follows:

Definition 3: Uniform Convergence

A sequence of functions (f_n) defined on a set A is said to be uniformly convergent to a function f on A if given a number $\in > 0$, there exists a positive integer m depending only on $\in$ such that

$$f_n(x) - f(x) | < \in \text{ for } n \geq m \text{ and } \forall x \in A.$$

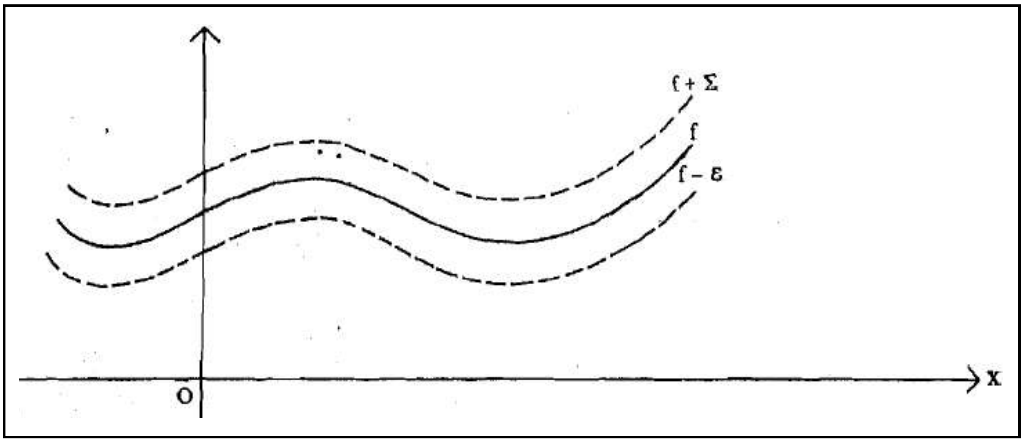
We write it as $f_n \to f$ uniformly on A or lim $f_n(x) = f(x)$ uniformly on A. Also f is called the uniform limit of f on A.

Note that if $f_n \to f$ uniformly on the set A, for a given $\in > 0$, there exists m such that

$$f(x) - \in < f_n(x) < f(x) + \in$$

for all $x \in A$ and $n \geq m$. In other words, for $n \geq m$, the graph of f_n lies in the strip between the graphs of $f - \in$ and $f + \in$. As shown in the figure below, the graphs of f_n Tor $n \geq m$ will all lie between the dotted lines.

Notes



From, the definition of uniform convergence, it follows that uniform convergence of a sequence of functions implies its pointwise convergence and uniform limit is equal to the pointwise limit. We will show below by suitable examples that the converse is not true.

Example: Show that the sequence (f_n) where $f_n(x) = \frac{x}{n}$, $x \in R$ is pointwise but not uniformly convergent in R.

Solution: You have seen that (f_n) is pointwise convergent to f where $f(x) = 0\ \forall x \in R$. In the same example, at the end, it is remarked that given $\in > 0$, it is not possible to find a positive integer m such that $\frac{|x|}{n} < \in$ for $n \geq m$ and $\forall x \in R$ i.e., $|f_n(x) - f(x)| < \in$ for $n \geq m$ and $\forall x \in R$. Consequently (f_n) is not uniformly convergent in R.

Example: Show that the sequence (f_n) where $f_n(x) = x^n$ is convergent pointwise but not uniformly on [0, 1].

Solution: You have been shown that (f_n) is pointwise convergent to, f on [0, l] where

$$f(x) = 0\ \forall x \in [0, 1[\text{ and } f(l) = l$$

Let $\in > 0$ be any number. For x = 0 or x = 1, $|f_n(x) - f(x)| < \in$ for $n \geq 1$.

For $0 < x < l$, $|f_n(x) - f(x)| < \in$ if $x^n < \in$ i.e. $n \log x < \log \in$ i.e. $n > \frac{\log \in}{\log x}$.

since log x is negative for 0 < x < 1. If we choose $m = \left[\frac{\log \in}{\log x}\right] + 1$, then $|f_n(x) - f(x)| < \in$ for $n \in \geq m$.

Clearly m depends upon $\in$ and x.

We will now prove that the convergence is not uniform by showing that it is not possible to find an m independent of x.

Let us suppose that $0 < \in < l$. If there exists m independent of x in [0, 1] so that

$$< |f_n(x) - f(x)| < \in \text{ for all } n \geq m,$$

then $x''< \in$ for all $n \geq m$, whatever may be x in $0 < x < 1$.

Notes

If the same m serves for all x for a given $\in > 0$ then $x^m < \in$ for all x, $0 < x < 1$. This implies that $m > \frac{\log \in}{\log x}$ (since log x is negative). This is not possible since log x decreases to zero as x tends to 1 and so $\log \in / \log x$ is unbounded.

Thus we have shown that the sequence (f_n) does not converge to the function f uniformly in [0, 1] even though it converges pointwise.

Example: Show that the sequence (g_n) where $g_n(x) = \frac{x}{1+nx}$, $x \in [0, \infty[$ is uniformly convergent in $[0, \infty[$.

Solution: $\lim_{n\to\infty} g_n(x) = 0$ for all x in the interval $[0, \infty[$. Thus (g_n) is pointwise converge of where $f(x) = 0 \ \forall x \in [0, \infty[$.

Now $|g_n(x) - f(x)| = \frac{x}{1+nx} < \frac{1}{n}$ for all x in $[0, \infty[$.

Since $\lim_{n\to\infty} \frac{1}{n} = 0$, therefore given $\in > 0$, there exists a positive integer m such that $\frac{1}{n} < \in$ for $n \geq m$

Thus m depends only on $\in$. Therefore,

$$|g_n(x) - f(x)| < \in \text{ for } n \geq m \text{ and } \forall x \in [0, \infty[.$$

Therefore $(g_n) \to f$ uniformly in $[0, \infty[$.

Just as you have studied Cauchy's Criterion for convergence of sequence of real numbers, we have Cauchy's Criterion for uniform convergence of sequence of functions which we now state and prove.

Theorem 1: Cauchy's Principle of Uniform Convergence

The necessary and sufficient condition for a sequence of functions (f_n) defined on A to converge uniformly on A is that for every $\in > 0$, there exists a positive integer m such that

$$|f_n(x) - f_k(x)| < \in \text{ for } n > k \geq m \text{ and } \forall x \in A$$

Proof: Condition is necessary. It is given that (f_n) is uniformly convergent on A.

Let $f_n - f$ uniformly on A. Then given $\in > 0$, there exists a positive integer m such that

$$|f_n(x) - f(x)| < \in/2 \text{ for } n \geq m \text{ and } \forall x \in A.$$

$$\therefore |f_n(x) - f_k(x)| = |f_n(x) - f(x)| + |f(x) - f_k(x)|$$

$$< |f_n(x) - f(x)| + |f(x) - f_k(x)| \text{ (By triangular inequality)}$$

$$< \frac{\in}{2} + \frac{\in}{2} + \text{ for } n > k \geq m \text{ and } \forall x \in A$$

This proves the necessary part. Now we prove the sufficient part.

Condition is sufficient: It is given that for every $\in > 0$, there exists a positive integer m such that $|f_n(x) - f_k(x)| < \in$ for $n > k \geq m$ and for all x in A. But by Cauchy's principle of convergence of sequence of real numbers, for each fixed point x of A, the sequence of numbers $(f_n(x))$ converges. In other words, (f_n) is pointwise convergent say to f on A. Now for each $\in > 0$, there exists a positive integer m such that

Notes

$$|f_n(x) - f(x)| < \frac{\in}{2} \text{ for } n > k \geq m.$$

Fix k and let $n \to \infty$. Then $f_n(x) \to f(x)$ and we get

$$|f(x) - f_k(x)| < \frac{\in}{2} \text{ i.e., } |f_k(x) - f(x)| < \in.$$

This is true for $k \geq m$ and for all x in A. This shows that (f) is uniformly convergent to f on A, which proves the sufficient part.

As remarked in the introduction, uniform convergence is the form of convergence of the sequence of function (f_n) which preserves the continuity, differentiability and integrability of each term f_n of the sequence when passing to the limit function f. In other words if each member of the sequence of functions (f_n) defined on a set A is continuous on A, then the limit function f is also continuous provided the convergence is uniform. The result may not be true if the convergence is only pointwise. Similar results hold for the differentiability and integrability of the limit function f. Before giving the theorems in which these results are proved, we discuss some examples to illustrate the results.

Example: Discuss for continuity the convergence of a sequence of functions (f), where $f(x) = 1 - |1 - x^2|^n$, $x \in \{x \mid |1 - x^2| \leq 1\} = [-\sqrt{2}, \sqrt{2}]$.

Solution: Here $\lim_{n\to\infty} f(x) = \begin{cases} 1, & \text{when } |1-x^2|<1 \\ 0, & \text{when } 1-x^2|=1 \text{ i.e. } x = 0 \pm \sqrt{2} \end{cases}$

Therefore the sequence (f_n) is pointwise convergent to f where

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

Now each member of the sequence (f,) is continuous at 0 but f is discontinuous at 0. Here (f_n) is not uniformly convergent in $[-\sqrt{2}, \sqrt{2}]$ as shown below.

Suppose (f_n) is uniformly convergent in $[-\sqrt{2}, \sqrt{2}]$, so that f is its uniform limit.

Taking $\in = \frac{1}{2}$, there exists an integer m such that

$$f(x) = < \frac{1}{2} \text{ for } n \geq m \text{ and } \forall\, x \in [\sqrt{2}, \sqrt{2}].$$

in particular $|f_m(x) - f(x)| < \frac{1}{2}$ for $x \in [\sqrt{2}, \sqrt{2}]$

$$\text{Now } |f_m(x) - f(x)| = \begin{cases} |1-x^2|^{m} & \text{when } |1-x^2|<1 \\ 0 & \text{when } |1-x^2|=1 \end{cases}$$

Since $\lim_{x\to 0} |1 - x^2|^m = 1$, $\exists$ a + v no. δ such that

$$|1 - x^2|^m - 1| < 1/4 \text{ for } 0 < |x| < \delta$$

$$\text{i.e. } 3/4 < |1 - x^2|^m < 5/4 \text{ for } |x| < \delta$$

Notes

So $(1 - x^2|^m > \frac{1}{2}$ for $|x| < \delta$ which is a contradiction.

Consequently (f_n) is not uniformly convergent in $[-\sqrt{2}, \sqrt{2}]$.

Example: Discuss, for continuity, the convergence of the sequence $((f_n)$ where

$$f_n(x) = \frac{x}{1+nx} \quad x \in [0, \infty[.$$

Solution: As you have seen that $(f_n) \to f$ uniformly in $[0, \infty[$ where $f(x) = 0$, $x \in [0, \infty[$.

Here each f_n is continuous in $[0, \infty[$ and also the uniform limit is continuous in $[0, \infty[$.

Example: Discuss for differentiability the sequence (f_n) where

$$f_n(x) = \frac{\sin nx}{\sqrt{n}}, \quad \forall x \in [0, \infty[.$$

Solution: Here $(f_n) \to f$ uniformly where $f(x) = 0 \;\; \forall x \in R$. You can see that each f_n and fare differentiable in R and

$$f_n'(x) = \sqrt{n} \cos nx \text{ and } f'(x) = 0 \quad \forall x \in R.$$

$$f_n'(0) = \sqrt{n} \to \infty \text{ whereas } f'(0) = 0$$

$$\lim_{n\to\infty} f'_n(0) \neq f'(0)$$

i.e. limit of the derivatives is not equal to the derivative of the limit.

As you will see in the theorem for the differentiability of f and the equality of the limit of the derivatives and the derivative of the limit, we require the uniform convergence of the sequence (f_n).

Example: Discuss for integrability the sequence (f_n) where

$$f(x) = n\,x\,e^{-nx^2}, \; x \in [0, 1],$$

Solution: If x = 0, then $f_n(0) = 0$

and $\lim f_n(0) = 0$. If $x \neq 0$, $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{nx}{e^{nx^2}}$ which is of the form $\frac{\infty}{\infty}$.

Applying L' Hopital's Rule, we have

$$\lim_{n\to\infty} \frac{nx}{e^{nx^2}} = \lim_{n\to\infty} \frac{x}{2nxe^{nx^2}} = 0$$

So $(f_n) \to f$, pointwise, where $f(x) = 0, \;\; \forall x \in [0, 1]$

You may find that $\int_0^1 f_n(x)\,dx = \frac{1}{2}(1 - e^{-n})$ and $\int_0^1 f(x)\,dx = 0$

Notes

Therefore, $\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = \frac{1}{2} \neq \int_0^1 f(x)\,dx = \int_0^1 f(x)\,dx = \int_0^1 \lim_{n\to\infty}(f_n(x))dx$. That is, the integral of the limit is not equal to be limit of the sequence of integrals. In fact, (f_n) is not uniformly convergent to f in [0, 1]. This we prove by the contradiction method. If possible, let the sequence be uniformly convergent in [0, 1]. Then, for $\in = \frac{1}{4}$, there exists a positive integer m such that $|f(x) - f(x)| < \frac{1}{4}$, for $n \geq m$ and $\forall\, x \in [0, 1]$.

i.e., $\frac{nx}{e^{nx^2}} < \frac{1}{4}$, for $n \geq m$ and $V\ x \in [0, 1]$

Choose a positive integer $M \geq m$ such that $\frac{1}{M} \in [0, 1]$,

Take n = M and x $\frac{1}{\sqrt{M}}$. We get

$$\frac{1}{\sqrt{M}} < \frac{1}{4} \text{ i.e., } M < \frac{e^2}{16} < 1.$$

which is a contradiction. Hence (f) is not uniformly convergent in [0, 1].

Now we give the theorems without proof which relate uniform convergence with continuity, differentiability and integrability of the limit function of a sequence of functions.

Theorem 2: Uniform Convergence and Continuity

If (f_n) be a sequence of continuous functions defined on [a, b] and $(f_n) \to f$ uniformly on [a, b], then f is continuous on [a, b].

Theorem 3: Uniform Convergence and Differentiation

Let (f_n) be a sequence of functions, each differentiable on [a, b] such that $(f_n(x_0))$ converges for some point x_0 of [a, b]. If (f_n) converges uniformly on [a, b] then (f_n) converges uniformly on [a, b] to a function f such that

$$f'(x) = \lim_{n\to\infty} f'_n(x);\ x \in [a, b].$$

Theorem 4: Uniform Convergence and Integration

If a sequence (f_n) converges uniformly to f on [a, b] and each function f_n is integrable on [a, b], then f is integrable on [a, b] and

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\int_a^b f_n(x)\,dx$$

14.3 Series of Functions

Just as we have studied series of real numbers, we can study series formed by a sequence of functions defined on a given set A. The ideas of pointwise convergence and uniform convergence of sequence of functions can be extended to series of functions.

Definition 4: Series of Functions

A series of the forth $f_1 + f_2 + f_3 + ... + f_n + ...$ where the f_n are real functions defined on a given set ACR is called a series of functions and is denoted by $\sum_{n=1}^{\infty} f_n$. The function f_n is called nth term of the series.

For each x in A, $f_1(x) + f_2(x) + f_3(x) + ... +$ is a series of real numbers. We put $S_n(x) = \sum_{k=1}^{n} f_k(x)$. Then we get a sequence (S,) of real functions defined on A. We say that the given series $f_1 + f_1 + ... + f_n + ...$ of functions converges to a function pointwise if the sequence (S„) associated to the given series of functions converges pointwise to the function f. i.e. (S„ (x)) converges to f(x) for every x in A.

We also say that f is the pointwise sum of the series Σf_n on A.

If the sequence (S_n) of functions converges uniformly to the function f, then we say that the given series $f_1 + f_2 + ... + f_n + ...$, of functions converges uniformly to the function f on A and f is called uniform sum of $\sum_{i=1}^{\infty} f_n$ on A. The function S_n is called the sum of n terms of the given series or the n partial sum of the series and the sequence (S_n) is called the sequence of partial sums of the series $\sum_{i=1}^{\infty} f_n$. To make the ideas clear, we consider some examples.

Example: Let $f_n(x) = x^{n-1}$ where $x_0 = 1$ and $-r \le x \le r$ where $0 < r < 1$. Then the associated series is $1 + x + x^2 + ...$

In this case, $S_n(x) = 1, + x + x^2 + ... + x^{n-1}$. It is clear that $S_n(x) = \frac{1-x^n}{1-x}$.

This sequence (S, (x)) of functions is easily seen to converge pointwise to the function $f(x) = \frac{1}{I-x}$, since $x^n \to 0$ as $n \to \infty$, since $|x| < r < I$ but the convergence is not uniform as shown below:

Let $\varepsilon > 0$ be given.

$$|S_n(x) - f(x)| = \frac{|x|^n}{|1-x|} \le \frac{r^n}{1-r} \text{ if } r^n < \varepsilon(1-r)$$

$$\text{i.e. } n > \frac{\log(\in(1-r))}{\log r}$$

$$\text{If } m = \left[\frac{\log(\in(1-r))}{\log r}\right] = 1, \text{ then}$$

$|s,(x) - f(x)| < \varepsilon$ if $n \ge m$ and for $-r \le x \le r$.

Therefore (S_n) converges uniformly in [–r, r]. Thus the geometric series $1 + x + x^2 + ...$ converges uniformly in [–r, r] to the sum function $f(x) = \frac{I}{1-x}$.

Notes

Example: Let $f_n(x) = n\,x\,e^{-nx^2} - (n-1)\,x\,e^{-(n-1)x^2}$, $x \in [0, 1]$.

Consider the series $\sum_{k=1}^{n} f_n(x)$.

In this case $S_n(x) = \sum_{k=1}^{n} (k x^{-kx^2} - (k-1) x e^{-(k-1)x^2}) = n\,x\,e^{-nx^2}$

As you have seen that this sequence (S_r) is pointwise but not uniformly convergent to the function f where f(x) = 0, $x \in (0, 1)$. Thus the series $\Sigma\, f_n(x)$ is pointwise convergent but not uniformly to the function f where f(x) = 0, $x \in [0, 1]$.

There is a very useful method to test the uniform convergence of a series of functions. In this method, we relate the terms of the series with those of a series with constant terms. This method is popularly called Weierstrass's M-test given by the German mathematician K.W.T. Weierstrass (1815-1 897). We state this test in the form of the following theorem (without proof) and illustrate the method by an example.

Theorem 5: Weierstrass M-Test

Let Σf_n be a series of functions defined on a subset A of R and let (M_n) be a sequence of real numbers such that ΣM_n is convergent and $|f_n(x)| \le M_n$, $\forall$ n and $\forall\, x \in A$. Then Σf_n is uniformly and absolutely convergent on A.

Example: Test the uniform convergence of the series $\sum_{n=1}^{\infty} \frac{x}{n^2(n+1)}$

Solution: Since $|f_n(x)| = \frac{x}{n^2(n+1)} \le \frac{k}{n^3}, \forall n$ and $\in [0, k]$.

Now the series $\Sigma M_n = k\,\Sigma\, \frac{k}{n^3}$ is known to be convergent, by p-test.

Therefore, by Weierstrass M-test, the given series is uniformly convergent in the set [0, k].

Task Show that the series $\sum_{n=1}^{\infty} \frac{1}{n^4 + x^2}$ converges uniformly, $\forall\, x \in R$.

Self Assessment

Fill in the blanks:

1. A sequence of functions (f_n) defined on a set A is said to be convergent pointwise to f if for each x in A, we have given fn (x) = f(x). Generally, we write .... on A.

2. A sequence of functions (f_n) defined on a set A is said to be uniformly convergent to a function f on A if given a number $\in > 0$, there exists a positive integer m depending only on $\in$ such that ....

3. If (f_n) be a sequence of continuous functions defined on [a, b] and $(f_n) \to f$ uniformly on [a, b], then f is continuous on [a, b] is known as ....

Notes

4. Let (f_n) be a sequence of functions, each differentiable on [a, b] such that $(f_n(x_0))$ converges for some point x_0 of [a, b]. If (f_n) converges uniformly on [a, b] then (f_n) converges uniformly on [a, b] to a function f such that ....

5. If a sequence (f_n) converges uniformly to f on [a, b] and each function f_n is integrable on [a, b], then f is integrable on [a, b] and ....

14.4 Summary

- In this unit you have learnt how to discuss the pointwise and uniform convergence of sequences and series of functions. Sequence of functions is defined and pointwise convergence of the sequence of functions has been discussed. We say that a sequence of functions (f_n) is pointwise convergent to f on a set A if given a number e > 0, there is a positive integer m such that

  $|f_n(x) - f(x)| < \varepsilon$ for $n \geq m, x \in A$.

  m in general depends on ε and the point x under consideration. If it is possible to find m which depends only on s and not the point x under consideration, then (f_n) is said to be uniformly convergent are f on A. Cauchy's criteria for uniform convergence are discussed. Also in this section you have seen that if the sequence of functions (f_n) is uniformly convergent to a function f on [a, b] and each f_n' is continuous or integrable, then f is also continuous or integrable on [a,b]. Further it has been discussed that if (f_n) is a sequence of functions, differentiable on [a,b] such that $(f_n(x,))$ converges for some point x_0 of [a, b] and if (f_n) converges uniformly on [a, b], then (f_n) converges uniformly to a differentiable function f such that $f'(x) = \lim_{n\to\infty} f'_n(x); x \in [a, b]$.

- Finally pointwise and uniform convergence of series of functions is given. The series of functions is said to be pointwise or uniformly convergent on a set A according as the sequence of partial sums (s_n) of the series is pointwise or uniformly convergent on A.

14.5 Keywords

Uniform Convergence and Continuity: If (f_n) be a sequence of continuous functions defined on [a, b] and $(f_n) \to f$ uniformly on [a, b], then f is continuous on [a, b].

Uniform Convergence and Differentiation: Let (f_n) be a sequence of functions, each differentiable on [a, b] such that $(f_n(x_0))$ converges for some point x_0 of [a, b]. If (f_n) converges uniformly on [a, b] then (f_n) converges uniformly on [a, b] to a function f such that

$$f'(x) = \lim_{n\to\infty} f'_n(x); x \in [a, b].$$

Uniform Convergence and Integration: If a sequence (f_n) converges uniformly to f on [a, b] and each function f_n is integrable on [a, b], then f is integrable on [a, b] and

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx$$

14.6 Review Questions

1. Examine which of the following sequences of functions converge pointwise

   (i) $f_n(x) = \sin nx, -\infty < x < +\infty$

   (ii) $f_n(x) = \dfrac{mx}{1+n^2x^2}, -\infty < x < +\infty$

Notes

2. Test the uniform convergence of the following sequence of functions in the specified domains

   (i) $f_n(x) = \frac{1}{nx}$ in $0 < x < \infty$

   (ii) $f_n(x) = \frac{nx}{1+n^2x^2}$, $-\infty < < x < \infty$

   (iii) $f_n(x) = \frac{x^n}{1+x^n}$, $0 \le x \le 1$

   (iv) $f_n(x) = \frac{1}{n}$, $0 \le x < \infty$

3. Show that the limit function of the sequence (f,) where $(f_n)(x) = \frac{x}{n}$, $x \in R$, is continuous in R while (f,,) is not uniformly convergent.

4. Show that for the sequence (f,) where $(f_n)(x) = nx(1 - x^2)^n$, $x \in [0, 1]$, the integral of the limit is not equal to the limit of the sequence of integrals.

5. Show that the series

   $\frac{x}{x+1} + \frac{x}{(x+1)(2x+1)} + \frac{x}{(2x+1)(3x+1)} + \ldots\ldots$ is uniformly convergent in]k, ∞[where k is a positive number.

6. Show that the series $\Sigma \frac{x}{n(n+1)}$ is uniformly convergent in [0, k] where k is any positive number but it does not converge uniformly in [0, ∞].

Answers: Self Assessment

1. $f_n \to f$ (pointwise)
2. $f_n(x) - f(x)| < \in$ for $n \ge m$ and $\forall x \in A$.
3. Convergence and Differentiation
4. $f'(x) = \lim_{n\to\infty} f'_n(x)$; $x \in [a, b]$.
5. $\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx$

14.7 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 15: Uniform Convergence of Functions

CONTENTS

Objectives

Introduction

15.1 Uniform Convergence

15.2 Testing Pointwise and Uniform Convergence

15.3 Covers and Subcovers

15.4 Dini's Theorem

15.5 Summary

15.6 Keywords

15.7 Review Questions

15.8 Further Readings

Objectives

After studying this unit, you will be able to:

- Define uniform convergence
- Explain the testing pointwise and uniform convergence
- Discuss the covers and subcovers
- Explain Dini's theorem

Introduction

In earlier unit as you all studied about sequences in metric spaces. This unit will explain that pointwise convergence of a sequence of functions was easy to define, but was too simplistic of a concept. We would prefer a type of convergence that preserves at least some of the shared properties of a function sequence. Such a concept is uniform convergence.

15.1 Uniform Convergence

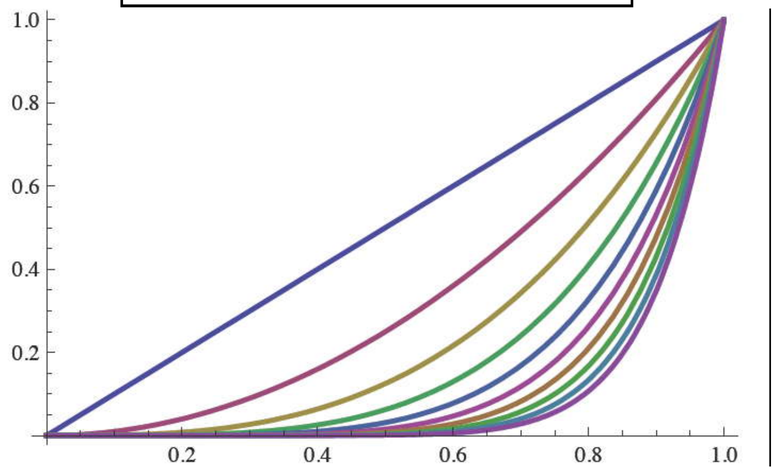
Definition 1: Let $I \subset \mathbb{R}$ and $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on *I*. We say that f_n converges to *f* **pointwise** on *I* as $n \to \infty$ if:

$$\forall x \in I : f_n(x) \to f(x) \text{ as } \to \infty$$

Example: Let *I* = (0, 1) and $f_n(x) = x^n, \left(f_1(x) = x^2, f_3(x) = x^3, ...\right)$. It can be observed that $\forall x \in I: \ f_n(x) = x^n \to 0$. So f_n converges pointwise to the zero function on *I*.

Notes

Figure 15.1: Pointwise convergence of $f_n = x^n$. First 10 terms of f_n.

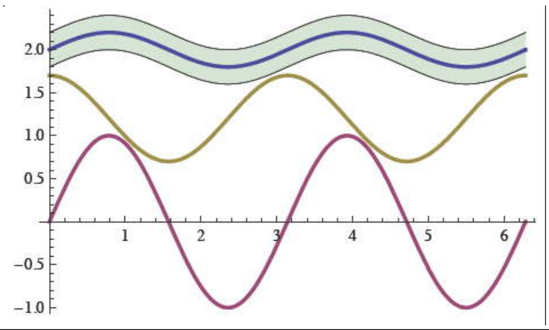


Definition 2: Let $I \subset \mathbb{R}$ and $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on *I*. We say that f_n converges of *f* **uniformly** on *I* if:

$$\forall_{\varepsilon} > 0 \exists N \in \mathbb{N} \forall_{x} \in I : |f_n(x) - f(x)| < \varepsilon$$

Meaning: For any ε-tube around *f* all functions f_n starting from some *N* will be lying inside the tube.

Figure 15.2: Uniform convergence. f_1 in red, f_2 in yellow, *f* in blue, ●-tube arund *f* in green.



Example: Let us take $I = (0,1)$ and $f_n = x^n$. Does it converge to $f(x) = 0$ uniformly on *I*? In this case answer is no. But the converge holds in general.

Theorem 1: If $f_n \to f$ uniformly then $f_n \to f$ pointwise.

Proof: Pointwise convergence (*N* is allowed to depend on *x*):

$$\forall x \in I \ \forall \varepsilon > 0 \ \exists \ N \in \mathbb{N} \forall n \geq N : |f_n(x) - f(x)| < \varepsilon$$

Uniform convergence ($\tilde{N}$ cannot depend on *x*):

$$\forall \tilde{\varepsilon} > 0 \ \exists \tilde{N} \in \mathbb{N} \ \forall n \geq \tilde{N} \ \forall x \in I : |f_n(x) - f(x)| < \varepsilon$$

Notes

Hence, for any given $\varepsilon > 0$ in pointwise definition it suffices to take $N = \tilde{N}_\varepsilon$ taken from uniform definition.

15.2 Testing Pointwise and Uniform Convergence

1. Test the pointwise convergence.
2. If there is no pointwise convergence, there is no uniform convergence.
3. If f_n converges pointwise to some f test the uniform convergence to f.

Example: Let $I = [0,1]$, $f_n(x) = x^n$. We see that f_n converges pointwise to $f(x) = \begin{cases} 0, & x \neq 1 \\ 1, & x = 1 \end{cases}$.

Does it converge uniformly to f on I? the answer is no.

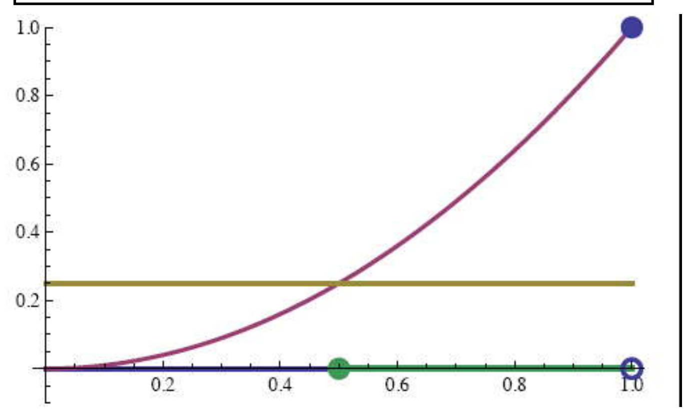
Proof: Negation of uniform convergence is

$$\exists \varepsilon > 0 \forall N \in \mathbb{N} \exists n \geq N \ \exists x \in I : |f_n(x) - f(x)| \geq \varepsilon$$

Take $\varepsilon = \frac{1}{4}$. Then $\forall N \in \mathbb{N}$ we have to find n and x such that $|f_n(x) - f(x)| \geq \varepsilon$. Take $n = N$. Now we want to find x. If we take any $x \in \left(\frac{1}{\sqrt[n]{4}}, 1\right)$ we get

$$|f_n(x) - f(x)| = |f_N(x) - f(x)| = |x^N - 0| = x^N \geq \frac{1}{4} = \varepsilon.$$

Figure 15.3: Suitable x's (green) for function $f_2(x) = x^2$ (red). Limit function f is blue, ●-tube and f in yellow.



Example: Let $I = [0,1]$ and $f_n(x) = \begin{cases} 1, & x \in [0, 1/n] \\ 0, & x \in [1, 2/n, 1] \end{cases}$. f_n converges pointwise to

$f(x) = \begin{cases} 1, & x = 0 \\ 0, & \text{otherwise} \end{cases}$ but does not converge uniformly to f.

Proof: If $x = 0$ then $f_n(0) \to 1$. If $x \in (0,1)$ then $f_n(x) \to 0$. Therefore $f_n \to f$ pointwise. To prove that f_n does not converge to f uniformly fix $\varepsilon = 1/2$. Then $\forall N \in \mathbb{N}$ choose $n = N$ and $x = 1/2N$. Then

Notes

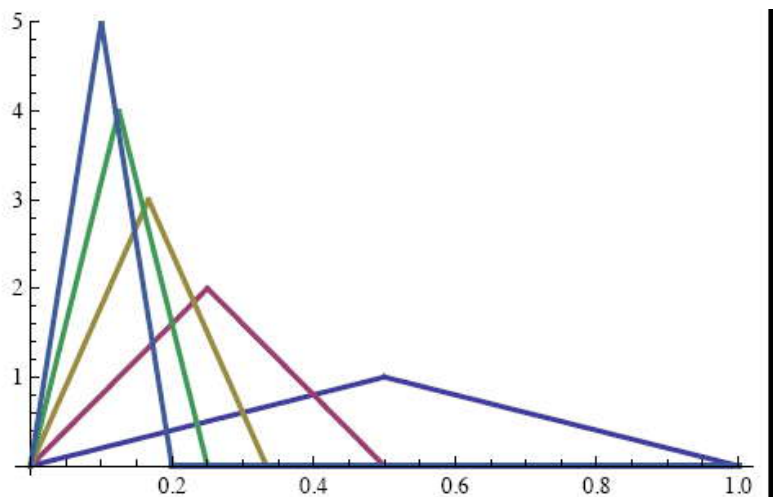
$$|f_n(x) - f(x)| = \left|f_N\left(\frac{1}{2N}\right) - f\left(\frac{1}{2N}\right)\right| = |1-0| \geq \frac{1}{2} = \varepsilon$$

Example: Let $I = [0,1]$ and $f_n(x) = \begin{pmatrix} 0, & x \in [1/n, 1] \\ n - n^2 x & x \in [0, 1/n] \end{pmatrix}$. Then f_x does not converge to any *f* pointwise nor uniformly.

Proof: Look at $x = 0$. Here $f_n(0) \to \infty$. Therefore, f_n does not converge pointwise to any *f*. Contrapositive tells us that f_n does not converge uniformly to any *f*.

Example: Let $I = [0,1]$ and f_n be from Figure 15.4

Figure 15.4: First 5 element of $\{f_n\}_{n=1}^{\infty}$ on [0, 1]. f_n starts at 0, has value of the peak equal of *n* at 1/2*n* and then returns to 0 at 1/*n*, value of the rest of 0.



Proof: Fix $x = 0$. Here $f_n(0) = 0 \to 0$. Now look at $x \in (0,1]$. Put $N = \left[\frac{1}{x}\right] + 1$. Then $\forall n \geq N : n \geq 1/x \Rightarrow x \geq 1/n$ and from image we can see that $f_n(x) \to 0$. Let us prove that f_n does not converge uniformly to the zero function. Take $\varepsilon = 1$. Then $\forall N \in \mathbb{N}$ choose *n* = *N* and *x* = 1/2*N*. We have

$$|f_n(x) - f(x)| = |f_N(1/2N) - f(1/2N)| = |N - 0| = N \geq 1 = \varepsilon.$$

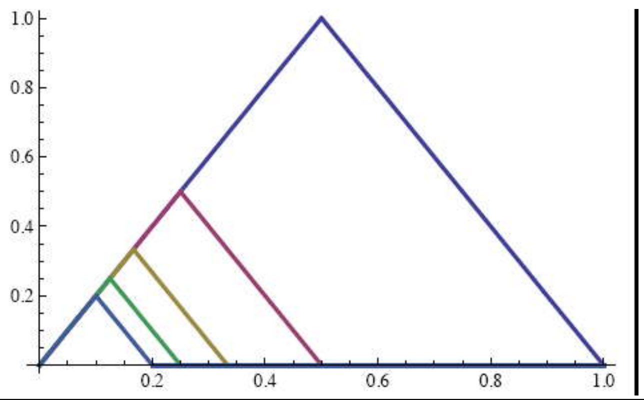
Example: Let *I* = [0, 1] and f_n to from Figure 15.5. Then f_n converges to the zero function pointwise and uniformly.

Proof: Pointwise convergence is clear since $\forall n > 1/x : f_n(x) = 0 \to 0$. For proof of uniform convergence of every $\varepsilon > 0$ choose $N = [1/\varepsilon] + 1$. Then $\forall n \geq N$ and $\forall n \in [0,1]$ we have

$$|f_n(x) - f(x)| = f_n \leq 1/n \leq 1/N \leq \varepsilon.$$

Notes

Figure 15.5: First 5 $\{f_n\}_{n=1}^{\infty}$ on [0, 1]. f_n is the same one as in Figure 15.4 but having value of the peak equal to 1/*n* instead of *n*.



Example: Let $I = (0, \infty)$ and $f_n(x) = \frac{1}{(x+n)^2}$. Then f_n converges to $f(x) = 0$ pointwise and uniformly.

Proof: Fix $x \in I$. We see that $f_n(x) \to 0$ as $n \to \infty$. Hence f_n converges pointwise. For proof of uniform convergence fix $\varepsilon > 0$, choose $N = [1/\varepsilon] + 1$. Then $\forall n \geq N$ and $\forall x \in I$ we have

$$|f_n(x) - f(x)| = \frac{1}{x+n} \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Uniform Convergence and Continuity

If all f_n are continuous and $f_n \to f$ pointwise, does *f* have to be continuous? The answer is no, it suffices to look at Example. Hence, pointwise convergence does not preserve continuity but on the other hand uniform does.

Theorem 2: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $[a,b]$. If all f_n are all continuous and $f_n \to f$ uniformly on $[a,b]$ then the limit function *f* is continuous.

Proof: We need to prove *f* is continuous at every $x \in |a,b|$. That is fix $x \in [a,b]$ and show

$$\forall \varepsilon > 0 \exists \delta > 0 \forall y \in [a,b] : |y - x| < \delta \Rightarrow |f(y) - f(x)| < \varepsilon$$

Let $\varepsilon > 0$. Since $f_n \to f$ uniformly

$$\exists N \in \mathbb{N} \forall n \geq \forall z \in [a,b] : |f_n(z) - f(z)| < \varepsilon/3.$$

In particular,

$$\forall z \in [a,b] : |f_N(z) - f(z)| < \varepsilon/3.$$

Since f_N is continuous at *x*,

$$\exists \tilde{\delta} > 0 \ \forall y \in [a,b] : |y - x| < \tilde{\delta} \Rightarrow |f_N(y) - f_N(x)| < \varepsilon/3$$

Therefore take $\delta = \tilde{\delta}$ and we get

$$|f(y) f(x)| = |f(y) + f_N(y) - f_N(y) + f_N(x) - f_N(x) - f(x)| \leq$$

Notes

$$|f(y)-f_N(y)|+|f_N(y)-f_N(x)|+|f_N(x)-f(x)| \le \varepsilon/3+\varepsilon/3+\varepsilon/3=\varepsilon.$$

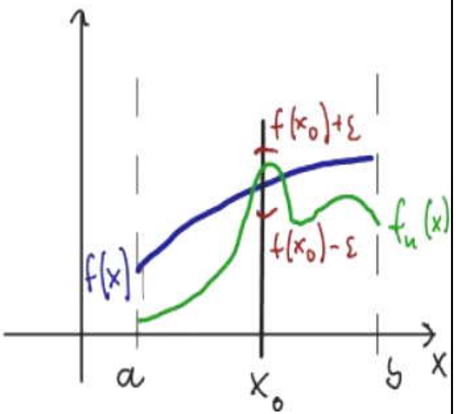
Remark: This theorem also work for f_n defined on any $I \subset \mathbb{R}$. Fix $x \in I$. This can be either isolated point in I or the limit point of I. In both cases we use exactly the same argument as for $[a, b]$ case.

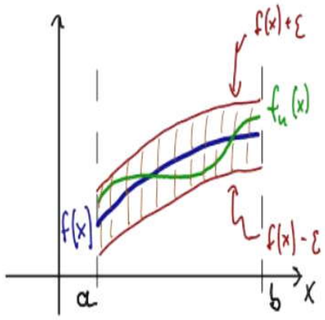
We should compare uniform with pointwise convergence:

- For pointwise convergence we could first fix a value for x and then choose N. Consequently, N depends on both and x.
- For uniform convergence $f_n(x)$ must be uniformly close to f(x) for all x in the domain. Thus N only depends on but not on x.

Let's illustrate the difference between pointwise and uniform convergence graphically:

Table 15.1

<table>
<tr><th>Pointwise Convergence</th><th>Uniform Convergence</th></tr>
<tr><td>For pointwise convergence we first fix a value x_0. Then we choose an arbitrary neighborhood around $f(x_0)$, which corresponds to a vertical interval centered at $f(x_0)$.


Finally we pick N so that $f_n(x_0)$ intersects the vertical line $x = x_0$ inside the interval $(f(x_0) - \varepsilon, f(x_0) + \varepsilon)$</td><td>For uniform convergence we draw an ε-neighborhood around the entire limit function f, which results in an "ε-strip" with f(x) in the middle.


Now we pick N so that $f_n(x)$ is completely inside that strip for all x in the domain.</td></tr>
</table>

Uniform convergence clearly implies pointwise convergence, but the converse is false as the above examples illustrate. Therefore uniform convergence is a more "difficult" concept. The good news is that uniform convergence preserves at least some properties of a sequence.

15.3 Covers and Subcovers

Consider a collection of open intervals $\{I_\alpha\}_{\alpha \in A}$, where A is an index set:

1. Finite collection: $\{I_1, I_2, ..., I_m\}$. In this case $A = \{1, ..., m\}$.

Example: $I_1 = (0, 2)$, $I_2 = (4, 5)$.

Notes

2. Infinite collection indexed by $\mathbb{N}: \{I_1, I_2, ...\}$. In this case $A = \mathbb{N}$.

Example: $I_n = (n-1/3, n+1/3)$.

3. Infinite collection indexed by $\mathbb{R}$. In this case $A = \mathbb{R}$.

Example: $\{(x-1, x+1)\}_{x \in \mathbb{R}}$. This set contains all open intervals of length 2.

Definition 3: A collection of open intervals $\{I_\alpha\}_{\alpha \in A}$ is a **cover** of a set $S \subset \mathbb{R}$ if $S \subset \cup_{\alpha \in A} I_\alpha$.

Definition 4: Given a set S and its cover $\{I_\alpha\}_{\alpha \in A}$, a **subcover** of $\{I_\alpha\}_{\alpha \in A}$ is a subcollection of $\{I_\alpha\}_{\alpha \in A}$, which itself is a cover for S.

Example: Let $I_1 = (0, 2)$, $I_2 = (4, 5)$. A collection containing these two intervals covers (0, 1) and $\{1\} \cup (4, 4.5)$, but does not cover [0, 1).

Example: Let $I_n = (n-1/3, n+1/3)$, $n \in \mathbb{N}$. Then $\{I_n\}_{n \in \mathbb{N}}$ covers $\mathbb{N}$, but does not cover $\mathbb{Z}$ or $\{1/2\}$. Let $S = \{1\} \cup (3-1/4, 3+1/4)$. Then $\{I_n\}_{n \in \mathbb{N}}$ is a cover for S. Moreover, $\{I_1, I_3\}$ is a finite subcover. Consider another case where $\{I_n\}_{n \in \mathbb{N}}$ is a cover of $\mathbb{N}$. Here $\{I_n\}_{n \in \mathbb{N}}$ has no finite subcover.

Example: Let $I_n = (-1+1/n, 1-1/n)$, $n \in \mathbb{N} \setminus \{1\}$. We see that the collection {(–3/4, 3/4)} is a finite subcover for set $S = [-17/24, 17/24]$. Now, consider set $S = (-1, 1)$. Is $\{I_n\}_{n \geq 2}$ a cover for S? The answer is positive and $\{I_n\}_{n \geq 2}$ has no finite subcover.

Example: We can observe that given a set S and its cover $\{I_\alpha\}_{\alpha \in A}$ sometimes it's possible to extract a finite subcover and sometimes isn't.

Theorem 3: (Heine-Borel). Every cover of closed interval $[a, b]$ has a finite subcover.

Proof: Definite a set $B = \{x \in [a, b] : [a, x]$ has a finite subcover. We see that $a \in B$, since $[a, a]$ is a single point. We need to prove that $b \in B$. Define c = sub B. Now, we have to prove two claims:

Claim 1: $c = b$ (Figure 13.6, left).

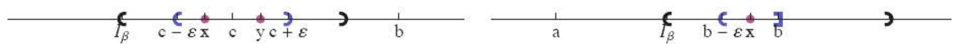
Assume that $c < b$. Then there is an interval I_β s.t. $c \in I_\beta$. Pick $\varepsilon > 0$ s.t. $(c-\varepsilon, c+\varepsilon) \subset I_\beta$ and $(c-\varepsilon, c+\varepsilon) \subset [a, b]$. Since c = sup B there is $x \in B$ s.t. $x \in (c-\varepsilon, c]$. Since $x \in B$, $[a, x]$ can be covered by finitely many intervals $\{I_{\alpha 1}, ..., I_{\alpha m}\}$. Pick $y \in (c, c+\varepsilon)$ and see that $[a, y]$ is covered by finite subcover $\{I_{\alpha 1}, ..., I_{\alpha m}, I_\beta\}$. Hence $y \in B$. But $y > c$ what implies that $c \neq \sup B$? What is contradiction? Therefore $c = b$.

Claim 2: $b \in B$ (Figure 15.6, right)

Pick I_β covering b. Take $\varepsilon > 0$ s.t. $(b-\varepsilon, b] \subset I_\beta$. Since $b = \sup B$ there exists $x \in B$ s.t. $x \in (b-\varepsilon, b]$. Since $x \in B$, $[a, x]$ can be covered by finitely many intervals $\{I_{\alpha 1}, ..., I_{\alpha m}\}$. Then $[a, b]$ is covered by finitely many intervals $\{I_{\alpha 1}, ..., I_{\alpha m}, I_\beta\}$.

Notes

Figure 15.6: Claim 1 (left) and Claim 2 (right) visualizations.



15.4 Dini's Theorem

Now, we are above to state in certain sense a converse to Theorem 1.

Theorem 4: (Dini's Theorem). Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $[a, b]$. If: (a) $f_n \to f$ pointwise, (b) all f_n are continuous, (c) f is continuous and (d) $\forall x \in [a, b] : \{f_n(x)\}_{n=1}^{\infty}$ is monotone then $f_n \to f$ uniformly.

Proof: Given $\varepsilon > 0$ we want

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \geq \ \forall x \in [a,b] : |f_n(x) - f(x)| < \varepsilon.$$

Let $x \in [a, b]$. Since $f_n \to f$ pointwise

$$\exists N(x) \in \mathbb{N} \ \forall n \geq N(x) : |f_n(x)| < \varepsilon / 2.$$

In particular,

$$|f_{N(x)}(x) - f(x)| < \varepsilon / 2.$$

Let $g(y) = f_{N(x)}(y) - f(y)$. Then $g(y)$ is continuous by (b) and (c). In particular $g(y)$ is continuous at x

$$\exists \delta(x) > 0 \ s.t. \ \forall y \in [a, b] : |y - x| < \delta(x) \Rightarrow |g(y) - g(x)| < \varepsilon / 2$$

which implies

$$|f_{N(x)}(y) - f(y)| = |g(y)| \leq |g(y) - g(x)| + |g(x)| =$$

$$= |g(y) - g(x)| + |f_{N(x)}(x) - f(x)| < \varepsilon / 2 + \varepsilon / 2 < \varepsilon.$$

Moreover,

$$\forall n \geq N(x) \ \forall y \in [a, b] : |y - x| < \delta(x) \Rightarrow |f_n(y) - f(y)| < \varepsilon$$

since, by (d)

$$|f_n(y) - f(y)| \leq |f_{N(x)}(y) - f(y)|.$$

Denote $I(x) = (x - \delta(x), x + \delta(x)) \ \forall x \in [a,b]$. We have shown

$$\exists N(x) \in \mathbb{N} \ \forall n \geq N(x) \ \forall y \in I(x) \cap [a, b] : |f_n(y) - f(y)| < \varepsilon.$$

Note that, $\{I(x)\}_{x \in [a, b]}$ is a cover for $[a, b]$, since $\forall x$ is covered at least by $I(x)$. By Heine-Borel Theorem 3 there is finite subcover $\{I(x_1), I(x_2), ..., I(x_m)\}$. Choose $N = \max\{N(x_1), ..., N(x_m)\}$ for any $\varepsilon > 0$. Let $n \geq N$ and $x \in [a, b]$. Let $I(x_i)$ be an interval converging x. Then $|x - x_i| < \delta(x_i)$. Since $n \geq N$ we have $n \geq N(x_i)$ and therefore $|f_n(x) - f(x)| < \varepsilon$.

Are all conditions of Dini's Theorem 4 important? The answer is yes, look at following examples.

Notes

Example: If (b) is not satisfied: let $I = [0, 1]$ and f_n is in the figure 15.7. We see that:

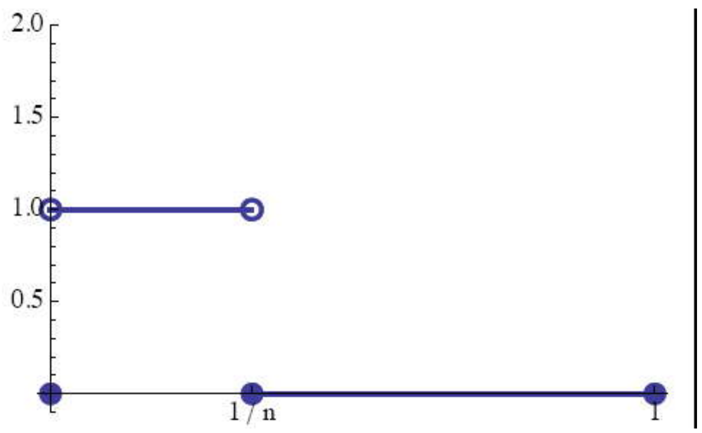
(a) $f_n \to 0$ pointwise.

(c) $f(x) = 0$ is continuous

(d) $\{f_n(x)\} = \{1, 1, 1, 0, 0, ...\}$ is decreasing

But f_n does not converge to $f(x) = 0$ uniformly.

Figure 15.7: Sequence of Real-valued Functions f_n being 1 in (0,1/n) and 0 elsewhere on [0, 1].



Example: If (c) is not satisfied: Let $I = [0,1]$ $f_n(x) = x^n$. We see that:

(a) $f_n \to f$. Where $f(x) = \begin{cases} 0, & x \neq 1 \\ 1, & x = 1 \end{cases}$.

(b) All f_n are continuous.

(d) $\{f_n\}$ is decreasing.

But f_n does not coverage to f uniformly.

Example: If (d) is not satisfied: Let $I = [0,1]$ and f_n here we see that:

(a) $f_n \to 0$ pointwise

(b) All f_n are continuous

(c) $f(x) = 0$ is continuous

(d) Put $x = 1/4$. Then $f_n(1/4) = \{a, b, c, ...\}$ where $a < b$ and $c < b$. Hence not satisfied.

But f_n does not converge to $f(x) = 0$ uniformly.

Notes

Example: If the interval is not closed. Let $I=(0,1)$ and $f_n(x)=x^n$.

(a) $f_n \to 0$ pointwise

(b) x^n is continuous.

(c) $f(x) = 0$ is continuous

(d) $\{f_n\}$ is monotonic $\forall x \in (0, 1)$.

But f_n does not converge to $f(x)=0$ uniformly.

Self Assessment

Fill in the blanks:

1. If there is no pointwise convergence, there is no ....
2. If f_n .... to some f test the uniform convergence to f.
3. A collection of open intervals $\{I_\alpha\}_{\alpha \in A}$ is a **cover** of a set ....
4. Given a set S and its cover $\{I_\alpha\}_{\alpha\in A}$, a .... is a subcollection of $\{I_\alpha\}_{\alpha\in A}$, which itself is a cover for S.

15.5 Summary

- Let f_1 f_2 f_3 … be a sequence of functions from one metric space into another, such that for any x in the domain, the images $f_1(x)$ $f_2(x)$ $f_3(x)$ … form a convergent sequence. Let g(x) be the limit of this sequence. Thus the function g is the limit of the functions f_1 f_2 f_3 etc.
- This sequence of functions is uniformly convergent throughout a region R if, for every Σ there is n such that fj(x) is within ε(g(x), for every x in R and for every j ≥ n. The functions all approach g(R) together, one n fits all. This is similar to uniform continuity, where one δ fits all.
- If the range space is complex, or a real vector space, the sequence of functions is uniformly convergent iff. All the component functions are uniformly convergent. Given an ε, find f_n that is close to g, and the components of f_n must be close to the components of g, for all x. Conversely, if the components are within ε then the n dimensional function is within nε, for all x.
- Without uniform convergence, g is rather unpredictable. Let the domain be the closed interval [0, 1], and let $f_n = x_n$. Note that the sequence f approaches a function g that is identically 0, except for g(1) = 1. Each function in the sequence is uniformly continuous, yet the limit function isn't even continuous.

15.6 Keywords

Heine-Borel: Every cover of closed interval $[a, b]$ has a finite subcover.

Dini's Theorem: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $[a, b]$. If: (a) $f_n \to f$ pointwise, (b) all f_n are continuous, (c) f is continuous and (d) $\forall x \in [a, b] : \{f_n(x)\}_{n=1}^{\infty}$ is monotone then $f_n \to f$ uniformly.

Notes

15.7 Review Questions

1. Let I = [1, 2] and $f_n(x) = \begin{cases} 1, & x \in [0, 1/n] \\ 0, & x \in [1, 2/n, 1] \end{cases}$. f_n converges pointwise to

   $f(x)$ = $f_n(x) = \begin{cases} 1, & x \in [0, 1/n] \\ 0, & x \in [1, 2/n, 1] \end{cases}$. f_n but does not converge uniformly to f.

2. Let $I = [0, 1]$ and f_n to from Figure 15.5. Then f_n converges to the zero function pointwise and uniformly.

3. Let $I = (0, \infty)$ and $f_n(x) = \dfrac{1}{(x+n)^2}$. Then f_n converges to $f(x) = 0$ pointwise and uniformly.

4. $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $[a, b]$. If all f_n are all continuous and $f_n \to f$ uniformly on $[a, b]$ then the limit function f is continuous.

5. Let $I_n = (n - 1/3, n + 1/3), n \in \mathbb{N}$. Then $\{I_n\}_{n \in \mathbb{N}}$ covers $\mathbb{N}$, but does not cover $\mathbb{Z}$ or {1/2}. Let $S = \{1\} \cup (3 - 1/4, 3 + 1/4)$. Then $\{I_n\}_{n \in \mathbb{N}}$ is a cover for S. Moreover, $\{I_1, I_3\}$ is a finite subcover. Consider another case where $\{I_n\}_{n \in \mathbb{N}}$ is a cover of ¥. Here $\{I_n\}_{n \in \mathbb{N}}$ has no finite subcover.

6. Let $I_n = (-1 + 1/n, 1 - 1/n), n \in \mathbb{N} \setminus \{1\}$. We see that the collection {(–3/4, 3/4)} is a finite subcover for set $S = [-17/24, 17/24]$. Now, consider set $S = (-1, 1)$. Is $\{I_n\}_{n \geq 2}$ a cover for S? The answer is positive and $\{I_n\}_{n \geq 2}$ has no finite subcover.

Answers: Self Assessment

1. uniform convergence
2. converges pointwise
3. $S \subset \mathbb{R}$ if $S \subset \cup_{\alpha \in A} I_\alpha$
4. subcover of $\{I_\alpha\}_{\alpha \in A}$

15.8 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7 (7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10, Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Notes

Unit 16: Uniform Convergence and Continuity

CONTENTS

Objectives

Introduction

16.1 Uniform Convergence Preserves Continuity

16.2 Uniform Convergence and Supremum Norm

16.3 Uniform Convergence and Integrability

16.4 Convergence almost Everywhere

16.5 Lebesgue's Bounded Convergence Theorem

16.6 Summary

16.7 Keywords

16.8 Review Questions

16.9 Further Readings

Objectives

After studying this unit, you will be able to:

- Define Uniform Convergence preserves Continuity
- Explain the Supremum Norm
- Discuss Sup-norm and Uniform Convergence

Introduction

In earlier unit as you all studied about uniform convergence. Uniform convergence clearly implies pointwise convergence, but the converse is false. Therefore uniform convergence is a more "difficult" concept. The good news is that uniform convergence preserves at least some properties of a sequence. This unit will explain Uniform Convergence preserves Continuity.

16.1 Uniform Convergence Preserves Continuity

If a sequence of functions $f_n(x)$ defined on D converges uniformly to a function f(x), and if each $f_n(x)$ is continuous on D, then the limit function f(x) is also continuous on D.

All ingredients will be needed, that f_n converges uniformly and that each f_n is continuous. We want to prove that f is continuous on D. Thus, we need to pick an x_0 and show that

$$|f(x_0) - f(x)| < \varepsilon \text{ if } |x_0 - x| < \delta$$

Let's start with an arbitrary $\varepsilon > 0$. Because of uniform convergence we can find an N such that

$$|f_n(x) - f(x)| < \varepsilon/3 \text{ if } n \geq N$$

for all $x \in D$. Because all f_n are continuous, we can find in particular a $\delta > 0$ such that

$$|f_N(x_0) - f_N(x)| < \varepsilon/3 \text{ if } |x_0 - x| < \delta$$

But then we have:

$$|f(x_0) - f(x)| \leq |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| \leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

as long as $|x_0 - x| < \delta$. But that means that f is continuous at x_0.

Before we continue, we will introduce a new concept that will somewhat simplify our discussion of uniform convergence, at least in terms of notation: we will use the supremum of a function to define a 'norm' of f

16.2 Uniform Convergence and Supremum Norm

Definition 1: The supremum norm of a function $f: I \to \mathbb{R}$ is

$$\|f\|_{\sup} = \|f\|_{\infty} = \sup_{x \in I} |f(x)|.$$

Example: Let $I = \mathbb{R}$ and $f(x) = \sin(x)$. Then $\|f\|_{\sup} = 1.$

Example: Let $I = [0,1]$ and $f(x) = -2x$. Then $\|f\|_{\sup} = 2.$ The norm stays the same even if we change the interval [0, 1] to (0, 1).

Theorem 1: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on *I*. Then $f_n \to f$ uniformly if and only if $\|f_n - f\|_{\sup} \to 0.$ Note that $\|f_n - f\|_{\sup}$ is just a sequence of number.

Proof: $f_n \to f$ uniformly $\Leftrightarrow \forall \varepsilon\ \exists N\ \forall n \geq N\ \forall x \in I : |f_n(x) - f(x)| \leq \varepsilon \Leftrightarrow \forall \varepsilon\ \exists N\ \forall n \geq N : \sup_{x \in I}$ $|f_n(x) - f(x)| \leq \varepsilon \Leftrightarrow \forall \varepsilon\ \exists N\ \forall n \geq N : \|f_n - f\| \leq \varepsilon$

Example: Let $f_n(x) = x^n$ on (0, 1). We can observe that $\|f_n - f\|_{\sup} = \sup_{x \in (0,1)} |x^n - 0|$ = $1 \to 0.$ As $f_n \to 0$ uniformly.

Example: Let $f_n(x) = \begin{cases} 1, & x \geq n \\ 0, & x < n \end{cases}$ on $\mathbb{R}$. Then $\|f\|_{\sup} = \sup_{x \in \mathbb{R}} |f_n(x) - 0| = 1 \to 0.$ So $f_n \to 0$ uniformly.

Using this proposition it can be easy to show uniform convergence of a function sequence, especially if the sequence is bounded. Still, even with this idea of sup-norm uniform convergence can not improve its properties: it preserves continuity but has a hard time with differentiability.

Example: Consider the sequence $f_n(x) = 1/n \sin(nx)$:

- Show that the sequence converges uniformly to a differentiable limit function for all x.
- Show that the sequence of derivatives f_n' does not converge to the derivative of the limit function.

This example is ready-made for our sup-norm because $|\sin(x)| < 1$ for all x. As for our proof: the sequence converges uniformly to zero because:

$$||f_n - f||_D = ||1/n \sin(n x) - 0||_D \leq 1/n \to 0$$

Notes

The sequence of derivatives is

$$f'(x) = \cos(nx)$$

Which does not converge (take for example $x = \pi$).

Example: Find a sequence of differentiable functions that converges uniformly to a continuous limit function but the limit function is not differentiable

we found a sequence of differentiable functions that converged point wise to the continuous, non-differentiable function $f(x) = |x|$. Recall:

$$f_n(x) = \begin{cases} -x - \dfrac{1}{2n} & \text{if } -1 \le x \le -\dfrac{1}{n} \\ \dfrac{n}{2}x^2 & \text{if } -\dfrac{1}{n} < x < \dfrac{1}{n} \\ x - \dfrac{1}{2n} & \text{if } \dfrac{1}{n} \le x \le 1 \end{cases}$$

That same sequence also converges uniformly, which we will see by looking at '$\| f_n - f \|_D$. We will find the sup in three steps:

If $1 \le x - 1/n$:

$f_n(x) - f(x)| = |-x - {}^1/_{2n} + x| = {}^1/_{2n}$

If $-1/n < x < 1/n$:

$|f_n(x) - f(x)| \le |{}^n/_2\, x^2| + |x| \le {}^n/_2\, {}^1/_n{}^2 + {}^1/_n = {}^3/_{2n}$

If $1/n \le x \le 1$:

$|f_n(x) - f(x)| = |x - {}^1/_{2n} - x| = {}^1/_{2n}$

Thus, $\| f_n - f \|_D < {}^3/_{2n}$ which implies that f_n converges uniformly to f. Note that all f_n are continuous so that the limit function must also be continuous (which it is). But clearly $f(x) = |x|$ is not differentiable at $x = 0$.

16.3 Uniform Convergence and Integrability

Theorem 2: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $[a, b]$. If all f_n are Riemann-integrable and $f_n \to f$ uniformly then *f* is Riemann-integrable and

$$\int_a^b f_n(x)dx \to \int_a^b f(x)dx.$$

Proof: Recall partition $P: a = t_0 < t_1 < ... < t_n = b.$ The upper Darboux sum $U(f, p) = \sum_{i=1}^{n} \sup_{[t_{i-1},\, t_i]}$ $f(x)(t_i - t_{i-1})$ and the lower Darboux sum $L(f, P) = \sum_{i=1}^{n} \inf_{[t_{i-1},\, t_i]} f(x)\,(t_i - t_{i-1})$. f is Riemann-integrable on $[a, b]$ if and only if $\forall \varepsilon > 0\; \exists P U(f, P) - L(f, P) < \varepsilon.$

Claim 1: f is Riemann-integrable.

Let $\varepsilon > 0.$ Since $f_n \to f$ uniformly $\exists n$ s.t. $\|f_n - f\|_{\sup} < \dfrac{\varepsilon}{4(b-a)}.$ Since f_n is R-integrable $\exists P$ s.t. $U(f_n, P) - L(f_n, P) < \varepsilon / 2.$ We have

Notes

$$\|f_n - f\|\sup < \frac{\varepsilon}{4(b-a)} \Rightarrow f_n(x)\frac{\varepsilon}{4(b-a)} \le f(x) \le f_n(x) + \frac{\varepsilon}{4(b-a)}, \forall x$$

i.e.

$$\sup_{[t_{i-1}, t_i]} f(x) \le \sup_{[t_{i-1}, t_i]} f_n(x) + \frac{\varepsilon}{4(b-a)}.$$

Then

$$U(f, P) = \sum_{i=1}^{n} \sup_{[t_{i-1}, t_i]} f(x)(t_i - t_{i-1}) \le \sum_{i=1}^{n}\left(\sup_{[t_{i-1}, t_i]} f_n(x) + \frac{\varepsilon}{4(b-a)}\right)(t_i - t_{i-1}) =$$

$$= U(f_n, P) + \frac{\varepsilon}{4(b-a)}\sum_{i=1}^{n}(t_i - t_{i-1}) = U(f_n, P) + \frac{\varepsilon}{4}.$$

Similarly,

$$L(f, P) \ge L(f_n, P) - \frac{\varepsilon}{4}.$$

So

$$U(f, P) - L(f, P) \ge U(f_n, P) + \frac{\varepsilon}{4} - L(f_n, P) + \frac{\varepsilon}{4} \le \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon$$

Claim 2: $\int_a^b f_n(x)dx \to \int_a^b f(x)dx.$

Since $\|f_n - f\|_{\sup} \to 0$, given $\varepsilon > 0$ there exists $N \in \mathbb{R}$ s.t. $\forall n \ge N : \|f_n - f\|_{\sup} < \varepsilon/(b-a)$. Therefore, $\forall n \ge N$:

$$\left|\int_a^b f_n(x)dx - \int_a^b f(x)dx\right| = \left|\int_a^b f_n(x)dx\right| \le \int_a^b |f_n(x) - f(x)|dx \le$$

$$\int_a^b \underbrace{\|f_n - f\|_{\sup}}_{\text{number}} dx = \|f_n - f\|_{\sup}(b-a) < \frac{\varepsilon}{(b-a)}(b-a) = \varepsilon.$$

Much more can be said about convergence and integration if we consider the Lebesgue integral instead of the Riemann integral. To focus on Lebesgue integration, for example, we would first define the concept of "convergence almost everywhere":

16.4 Convergence almost Everywhere

A sequence f_n defined on a set D converges (pointwise or uniformly) almost everywhere if there is a set S with Lebesgue measure zero such that f_n converges (pointwise or uniformly) on D\S. We say that f_n converges (point wise or uniformly) to f a.e.

In other words, convergence a.e. means that a sequence converges everywhere except on a set with measure zero. Since the Lebesgue integral ignores sets of measure zero, convergence a.e. is ready-made for that type of integration.

Notes

Example: Let r_n be the (countable) set of rational numbers inside the interval [0, 1], ordered in some way, and define the functions

$$\begin{cases} 1 & \text{if} \quad x = r_1, r_2, r_3, \ldots r_n \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad g(x) \begin{cases} 1 & \text{if} \quad x = \text{rational} \\ 0 & \text{if} \quad x = \text{irrational} \end{cases}$$

Show the following:

- The sequence g_n converges pointwise to g but the sequence of Riemann integrals of g_n does not converge to the Riemann integral of g.
- The sequence g_n converges a.e. to zero and so does the sequence of Lebesgue integrals of g_n.

Solution:

Each g_n is continuous except for finitely many points of discontinuity. But then each g_n is integrable and it is easy to see that

$$\int_{gn(x)} dx = 0$$

But the limit function is not Riemann-integrable and hence the sequence of Riemann integrals does not converge to the Riemann integral of the limit function.

Please note that while each g_n is continuous except for finitely many points, the limit function g is discontinuous everywhere

On the other hand, each g_n is zero except on a set of measure zero, and so is the limit function. Thus, using Lebesgue integration we have that all integrals evaluate to zero. But then, in particular, the sequence of Lebesgue integrals of g_n converge to the Lebesgue integral of g.

There are many theorems relating convergence almost everywhere to the theory of Lebesgue integration. They are too involved to prove at our level but they would certainly be on the agenda in a graduate course on Real Analysis. For us we will be content stating, without proof, one of the major theorems.

16.5 Lebesgue's Bounded Convergence Theorem

Let $\{f_n\}$ be a sequence of (Lebesgue) integrable functions that converges almost everywhere to a measurable function f. If $|f_n(x)| \leq g(x)$ almost everywhere and g is (Lebesgue) integrable, then f is also (Lebesgue) integrable and:

$$\lim_{n\to\infty} \int |f_n - f| dm = 0$$

Self Assessment

Fill in the blanks:

1. Let $I = \mathbb{R}$ and $f(x) = \sin(x)$. Then ....
2. Let $I = [0,1]$ and $f(x) = -2x$. Then .... The norm stays the same even if we change the interval [0, 1] to (0, 1).
3. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on *I*. Then $f_n \to f$ uniformly if and only if .... Note that $\|f_n - f\|_{\text{sup}}$ is just a sequence of number.
4. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $[a, b]$. If all f_n are Riemann-integrable and $f_n \to f$ uniformly then *f* is .... and

$$\int_a^b f_n(x)dx \to \int_a^b f(x)dx.$$

Notes

16.6 Summary

- If a sequence of functions $f_n(x)$ defined on D converges uniformly to a function f(x), and if each $f_n(x)$ is continuous on D, then the limit function f(x) is also continuous on D.
- All ingredients will be needed, that f_n converges uniformly and that each f_n is continuous. We want to prove that f is continuous on D.
- Before we continue, we will introduce a new concept that will somewhat simplify our discussion of uniform convergence, at least in terms of notation: we will use the supremum of a function to define a 'norm' of f.
- A sequence f_n defined on a set D converges (point wise or uniformly) almost everywhere if there is a set S with Lebesgue measure zero such that f_n converges (pointwise or uniformly) on D\S. We say that f_n converges (pointwise or uniformly) to f a.e.
- Much more can be said about convergence and integration if we consider the Lebesgue integral instead of the Riemann integral. To focus on Lebesgue integration, for example, we would first define the concept of "convergence almost everywhere".
- In other words, convergence a.e. means that a sequence converges everywhere except on a set with measure zero. Since the Lebesgue integral ignores sets of measure zero, convergence a.e. is ready-made for that type of integration.
- Let $\{ f_n \}$ be a sequence of (Lebesgue) integrable functions that converges almost everywhere to a measurable function f. If $| f_n(x) | \le g(x)$ almost everywhere and g is (Lebesgue) integrable, then f is also (Lebesgue) integrable and:

$$\lim_{n\to\infty}\int|f_n - f|dm = 0$$

16.7 Keywords

Convergence almost Everywhere: A sequence f_n defined on a set D converges (pointwise or uniformly) almost everywhere if there is a set S with Lebesgue measure zero such that f_n converges (point wise or uniformly) on D\S. We say that f_n converges (pointwise or uniformly) to f a.e.

Uniform Convergence Preserves Continuity: If a sequence of functions $f_n(x)$ defined on D converges uniformly to a function f(x), and if each $f_n(x)$ is continuous on D, then the limit function f(x) is also continuous on D.

Lebesgue's Bounded Convergence Theorem: Let $\{f_n\}$ be a sequence of (Lebesgue) integrable functions that converges almost everywhere to a measurable function f. If $| f_n(x) | \le g(x)$ almost everywhere and g is (Lebesgue) integrable, then f is also (Lebesgue) integrable and:

$$\lim_{n\to\infty}\int|f_n - f|dm = 0$$

16.8 Review Questions

1. Let $f_n(x) = x^n$ on (0, 1). We can observe that $\|f_n - f\|_{\sup} = \sup_{x\in(0,1)}|x^n - 0| = 1 \to 0$. By Theorem $f_n \to 0$ uniformly.

2. Let $f_n(x) = \begin{cases} 1, & x \ge n \\ 0, & x < n \end{cases}$ on $\mathbb{R}$. Then $\|f\|_{\sup} = \sup_{x\in\mathbb{R}}|f_n(x) - 0| = 1 \to 0$. So $f_n \to 0$ uniformly.

Notes

Answers: Self Assessment

1. $\|f\|_{sup} = 1$
2. $\|f\|_{sup} = 2$
3. $\|f_n - f\|_{sup} \to 0$
4. Riemann-integrable

16.9 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.

Unit 17: Uniform Convergence and Differentiability

Notes

CONTENTS

Objectives

Introduction

17.1 Uniform Convergence and Differentiability

17.2 Series of Functions

17.3 Central Principle of Uniform Convergence

17.4 Power Series and Uniform Convergence

17.5 Continuous but Nowhere Differentiable Function

17.6 Summary

17.7 Keywords

17.8 Review Questions

17.9 Further Readings

Objectives

After studying this unit, you will be able to:

- Discuss the uniform convergence and differentiability
- Explain the series of the functions
- Describe the central principle of uniform convergence

Introduction

In last unit you have studied about uniform convergence and continuity. Simple or pointwise convergence is not enough to preserve differentiability, and neither is uniform convergence by itself. However, if we combine pointwise with uniform convergence we can indeed preserve differentiability and also switch the limit process with the process of differentiation. It remains to clarify the connection between uniform convergence and differentiability.

17.1 Uniform Convergence and Differentiability

Is it true that if all f_n are differentiable and $f_n \to f$ uniformly then f is differentiable and $f_n' \to f'$? The answer is no. Look at the following examples.

Example: Let $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = \frac{1}{n}\sin(n^2 x)$.

We see that all f_n are differentiable and that $f_n \to f$ uniformly $\left(\|f_n - f\|_{\sup} = \left\|\frac{1}{n}\sin(n^2 x)\right\| = 1/n \to 0\right)$. But $f_n'(x) = n\cos(n^2 x) \to 0$ neither uniformly nor pointwise, although the zero function is differentiable.

Notes

Example: Let $f_n(x) = |x|^{1+1/n}$ be defined on [–1, 1]. See Figure 17.1.

We can observe that all f_n are differentiable, f_n converges uniformly to $f(x) = |x|$ by Dini's Theorem ($f_n(x) \to |x|$ pointwise, $f_n(x)$ are continuous, $f(x)$ is continuous, $\forall x$ in [–1, 1], $\{f_n(x)\}_{n=1}^{\infty}$ is increasing). But $|x|$ is not even differentiable.

The reason why this theorem cannot hold is that the uniform convergence, in fact, does not tell anything about slopes of f_n. See Figure 17.2.

Figure 17.1: First 4 terms of $\{|x|^{1+1/n}\}_{n=1}^{\infty}$

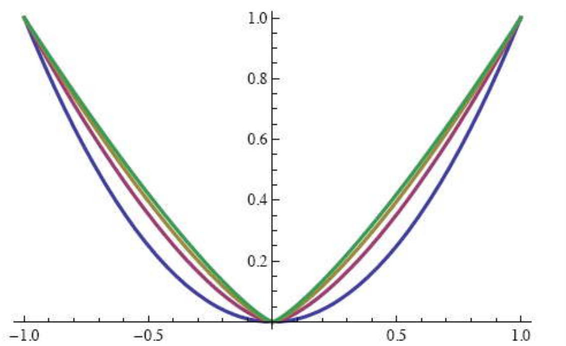
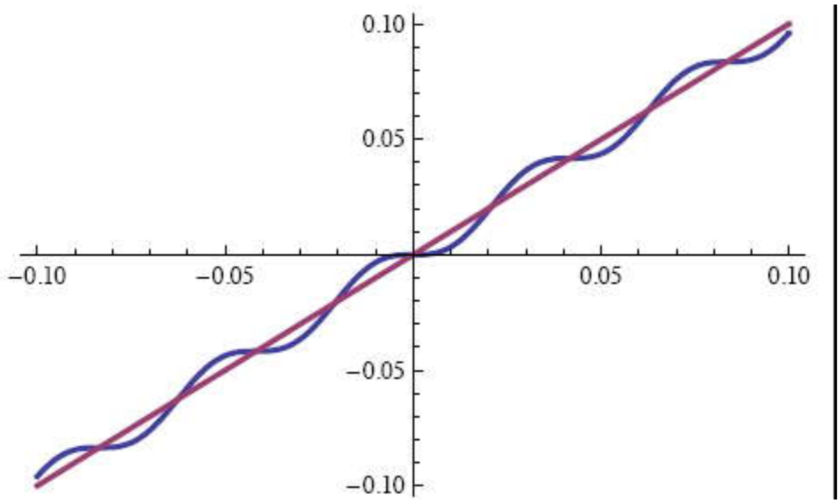


Figure 17.2: Slopes are Independent of Uniform Convergence



Theorem 1: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued function on $[a, b]$. If (a) all f_n are differentiable, (b) all f_n' are continuous, (c) $f_n' \to h$ uniformly, for some function $h:[a, b] \to \mathbb{R}$, (d) $\exists c \in [a, b]$ s.t. $f_n(c)$ converges then f_n converges uniformly to some $f:[a, b] \to \mathbb{R}$, in addition, this uniform limit *f* is differentiable and $f'(x) = h(x)$.

Proof: Firstly, we define a function $f(x)$ which satisfies $f'(x) = h(x)$ in Part 1 and then we show that $f_n \to f$ uniformly in Part 2.

Part 1: Define $f(c) := \lim_{n\to\infty} f_n(c)$ and $f(x) := f(c) + \int_c^x h(t)dt.$

Note that $h(t)$ is R-integrable since it is a uniform limit of continuous functions $f_n^{'}$ (Theorem 1). Therefore, according to the definition of $f(x)$ and by the fundamental theorem of calculus we can see that $f'(x) = \frac{d}{dx}(f(c) + \int_c^x h(t)dt) = \frac{d}{dx}(f(c) + \int_a^x h(t)dt + \int_c^a h(t)dt) = \frac{d}{dx}\int_a^x h(t) = h(x), \forall x \in [a, b].$

Part 2: We want to show that $f_n \to f$ uniformly on $[a, b]$. By the fundamental theorem of calculus:

$$f_n(x) = f_n(c) + \int_c^x f_n^{'}(t)dt.$$

Put $\varepsilon > 0$. Since $f_n(c) \to f(c)$ we have $|f_n(c) - f(c)| < \varepsilon/2$ for all $n \geq N_1$. Also $f_n^{'} \to h$ uniformly, so we have $\|f_n^{'} - h\|_{\sup} < \varepsilon/(2(b-a))$ for all $n \geq N_2$. Therefore put $N = \max\{N_1, N_2\}$. Then $\forall n \geq N\ \forall x \in [a, b]$ we get

$$|f_n(x) - f(x)| = \left|f_n(c) - f(c) + \int_c^x f_n^{'}(t)dt - \int_c^x h(t)dt\right| \leq$$

$$\leq |f_n(c) - f(c)| + \left|\int_c^x f_n^{'}(t) - h(t)dt\right| \leq$$

$$\leq |f_n(c) - f(c)| + \left|\int_c^x |f_n^{'}(t) - h(t)|dt\right| \leq$$

$$\leq |f_n(c) - f(c)| + \|f_n^{'} - h\|_{\sup}\left|\int_c^x 1dt\right| =$$

$$= |f_n(c) - f(c)| + \|f_n^{'} - h\|_{\sup}|x - c| \leq$$

$$\leq |f_n(c) - f(c)| + \|f_n^{'} - h\|_{\sup}(b - a) \leq$$

$$< \varepsilon/2 + \frac{\varepsilon}{2(b-a)}(b-a) = \varepsilon$$

Example: Theorem 1 is not true if we replace $[a, b]$ by $\mathbb{R}$. Look at $f_n(x) = \sin(x/n)$ on $\mathbb{R}$. Then $f_n^{'}(x) = \frac{1}{n}\cos(x/n) \to 0$ uniformly $\left\|\frac{1}{n}\cos(x/n)\right\|_{\sup} = 1/n \to 0\}$. Conditions (a) – (d) of Theorem 1 are satisfied but $f_n \to 0$ uniformly.

17.2 Series of Functions

Definition 1: Let $I \subset \mathbb{R}$ and $\{g_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on I. Then

$$\sum_{n=1}^{\infty} g_n$$

Notes

is a series of functions. Partial sums: $f_n(x) = \sum_{i=1}^{n} g_i(x)$. We say that $\sum_{n=1}^{\infty} g_n$ converges pointwise/uniformly to $f: I \to \mathbb{R}$ if $f_n \to f$ pointwise/unformly.

Example: Consider a series $\sum_{n=0}^{\infty} x^n$ on [–1/2, 1/2] and on (–1, 1). We want to investigate convergence of this series, whether it is pointwise of n uniform and eventually, what is the limit. Partial sums are in this case $f_n(x) = \sum_{i=0}^{n} x^i = 1 + x + x^2 + ... + x^n = \frac{x^{n+1}-1}{x-1}$. If we fix $x \in (-1, 1)$ we see that $f_n(x) = \frac{x^{n+1}-1}{x-1} \to \frac{1}{1-x}$ pointwise on (–1, 1) and therefore also on [–1/2, 1/2]. Does not series converge to the $\frac{1}{1-x}$ also uniformly on the same intervals? Look at the modulus

$$|f_n(x) - f(x)| = \left|\frac{x^{n+1}-1}{x-1} - \frac{1}{1-x}\right| = \left|\frac{x^{n+1}}{1-x}\right|$$

On [–1/2, 1/2] our series converges uniformly to *f* since partial sums do

$$\|f_n - f\|_{\sup} = \sup_{x \in [-1/2, 1/2]} \left|\frac{x^{n+1}}{1-x}\right| \le \frac{1}{2^{n+1}} 2 = \frac{1}{2^n} \to 0.$$

On (–1, 1) the series does not converge to *f* since partial sums do not.

$$\|f_n - f\|_{\sup} = \sup_{x \in (-1, 1)} \left|\frac{x^{n+1}}{1-x}\right| = \infty \text{ if } x \to 1.$$

Let us fix a sequence of real-valued functions $\{g_n\}_{n=1}^{\infty}$ on $[a, b]$.

Theorem 2: Continuity for Series

If (a) all g_n are continuous and (b) $\sum_{n=1}^{\infty} g_n$ converges uniformly then $\sum_{n=1}^{\infty} g_n$ is continuous.

Proof: Consider partial sum $f_n = g_1 + g_2 + ... + g_n$. We see that f_n is continuous, since it is a sum of continuous functions. Also, by (b), $f_n \to \sum_{n=1}^{\infty} g_n$ uniformly. Therefore, the continuity of the uniform limit $\sum_{n=1}^{\infty} g_n$ is Riemann-integrable and we can integrate the series term by term

$$\int_a^b \sum_{n=1}^{\infty} g_n dx = \sum_{n=1}^{\infty} \int_a^b g_n dx.$$

Proof: Since sum of R-integrable functions is R-integrable, we see that the partial sum $\sum_{i=1}^{n} g_i = g_1 + g_2 + ... + g_n$ is R-integrable and by (b) $\sum_{i=1}^{n} g_i \to \sum_{i=1}^{\infty} g_i$

$$\sum_{i=1}^{\infty} \int_a^b g_i dx = \lim_{n\to\infty} \sum_{i=1}^{n} \int_a^b g_i dx = \lim_{n\to\infty} g_i dx = \int_a^b \sum_{i=1}^{\infty} g_i dx$$

Theorem 3: Differentiability for Series

If (a) all g_n are differentiable, (b) all $g_n^{'}$ are continuous, (c) $\sum_{n=1}^{\infty} g_n^{'}$ converges uniformly and (d) $\exists c \in [a,b]$ s.t. $\sum_{n=1}^{\infty} g_n(c) < \infty$ then $\sum_{n=1}^{\infty} g_n$ converges uniformly and

$$\left(\sum_{n=1}^{\infty} gn\right)' = \sum_{n=1}^{\infty} g_n^{'}.$$

Proof: Since sum of differentiable functions is differentiable we observe that the partial sum $f_n = g_1 + g_2 + ... + g_n$ is differentiable. Similarly, all $f_n^{'} = g_1^{'} + g_2^{'} + ... + g_n^{'}$ are continuous. By (c) $f_n^{'} \to \sum_{n=1}^{\infty} g_n^{'}$ uniformly and by (d) $\exists c \in [a,b]$ s.t. $\sum_{n=1}^{\infty} g_n(c)\ \infty$, i.e. $\lim_{n\to\infty} f_n(c) < \infty$. Therefore we can apply Theorem to f_n and observe that $f_n \to \sum_{n=1}^{\infty} g_n$ uniformly and $\left(\sum_{n=1}^{\infty} g_n\right)' = \sum_{n=1}^{\infty} g_n^{'}$

17.3 Central Principle of Uniform Convergence

Definition 2: Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $I \subset \mathbb{R}$. Then $\{f_n\}_{n=1}^{\infty}$ is called a uniform Cauchy sequence if:

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N : \|f_n - f_m\|_{\sup} < \varepsilon.$$

Theorem 4: Central Principle of Uniform Convergence, CPUC

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $I \subset \mathbb{R}$. Then $\{f_n\}_{n=1}^{\infty}$ converges uniformly of I if and only if $\{f_n\}_{n=1}^{\infty}$ is a uniform Cauchy sequence on *I*.

Proof: '$\Rightarrow$' : Suppose f_n converges uniformly to some *f*. Let $\varepsilon > 0$, since $f_n \to f$ uniformly we have

$$\exists N \in \mathbb{N}\ \forall n \geq \forall x \in I : |f_n(x) - f(x)| < \varepsilon / 4.$$

Then $\forall_n, m \geq \forall x \in I$:

$$|f_n(x) - f_m(x)| = |f_n(x) - f(x) + f_m(x) + f(x)| \leq$$

$$\leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \varepsilon/4 + \varepsilon/4 = \varepsilon/2$$

Therefore,

$$\|f_n - f_m\|_{\sup} = \sup_{x \in I} |f_n(x) - f_m(x)| \leq \varepsilon/2 < \varepsilon.$$

'$\Rightarrow$' : Let $\{f_n\}$ be a uniform Cauchy sequence i.e.

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N : \|f_n - f_m(x)\|_{\sup} < \varepsilon/2.$$

In particular $|f_n(x) - f_m(x)| < \varepsilon/2$ for any $x \in I$. Look at the sequence of numbers $\{f_n(x)\}_{n=1}^{\infty}$ which is usual sequence of number and hence converges. Denote its limit $f(x)$. Now let $m \to \infty$ and get

Notes

$$\forall n \geq N \ \forall x \in I : |f_n(x) - f(x)| \leq \varepsilon/2 < \varepsilon$$

what by definition means that $f_n \to f$ uniformly.

Theorem 5: Weierstrass M-test

Let $\{g_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $I \subset \mathbb{R}$. Let $\{M_n\}_{n=1}^{\infty}$ be a sequence of number s.t. $\sum_{n=1}^{\infty} M_n < \infty$. If $|g_n(x)| \leq M_n, \forall x \in I \ \forall n \in \mathbb{N}$ then $\sum_{n=1}^{\infty} g_n$ converges uniformly.

Proof: $\sum_{i=1}^{\infty} M_i < \infty$ means $\left\{\sum_{i=1}^{n} M_i\right\}_{n=1}^{\infty}$ converges and therefore is a Cauchy sequence i.e. given $\varepsilon > 0$ we can find $N \in \mathbb{N}$ s.t. without loss of generality

$$\forall n, m \in \mathbb{N}, n > m \geq N : \left|\sum_{i=1}^{n} M_i - \sum_{i=1}^{m} M_i\right| = \left|\sum_{i=m+1}^{n} M_i\right| = \sum_{i=m+1}^{n} M_i < \varepsilon/2.$$

Let's prove that $\left\{\sum_{i=1}^{n} g_i\right\}_{n=1}^{\infty}$ is a uniform Cauchy sequence. Given $\varepsilon > 0$, pick N as above. Then $\forall n > m \geq N \in \mathbb{N} \ \forall x \in I$

$$\left|\sum_{i=1}^{n} g_i(x) - \sum_{i=1}^{m} g_i(x)\right| = \left|\sum_{i=m+1}^{n} g_i(x)\right| \leq$$

$$\leq \sum_{i=m+1}^{n} |g_i(x)| \leq \sum_{i=m+1}^{n} M_i < \varepsilon/2.$$

We get

$$\left\|\sum_{i=1}^{n} g_i - \sum_{i=1}^{m} g_i\right\|_{\sup} \leq \varepsilon/2 < \varepsilon.$$

Therefore $\left\{\sum_{i=1}^{n} g_i\right\}_{n=1}^{\infty}$ is a uniform Cauchy sequence and converges uniformly and so does $\sum_{i=1}^{\infty} g_i$.

Example: Consider series $\sum_{i=1}^{\infty} \frac{\sin(nx)}{2^n}$ on $\mathbb{R}$. By Weierstrass M-test, we see that this series converges uniformly on $\mathbb{R}$ since

$$\left|\frac{\sin(nx)}{2^n}\right| \leq \frac{1}{2^n} \text{ and } \sum_{n=1}^{\infty} \frac{1}{2^n} < \infty$$

Example: Consider series $\sum_{n=1}^{\infty} \frac{1}{n^2 + x}$ on $[0, \infty)$. By Weierstrass M-test, we can obtain uniform convergence of this series on $[0, \infty)$ since

$$\left|\frac{1}{n^2 + x}\right| \leq \frac{1}{n^2} \text{ and } \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

Notes

Example: Look at the uniform convergence of series $\sum_{n=1}^{\infty} x^n$ both on $[-r, r], 0 < r < 1$ and (–1, 1). In the first case we see that the series converges uniformly by Weierstrass M-test Since

$$|x^n| \le r^n \text{ and } \sum_{n=1}^{\infty} r^n < \infty$$

In the second one we will try to show that there is no uniform convergence. Look at the partial sums f_n. If we can prove that $\{f_n\}$ is not a uniform Cauchy sequence then $\{f_n\}$ is not uniformly convergent and therefore the series will not converge uniformly. Often it suffices to look at $\|f_{n+1} - f_n\|_{\sup}$ and show that it does not converge to 0.

$$\|f_{n+1} - f_n\|_{\sup} = \sup_{x \in (-1,1)} \left| \sum_{i=1}^{n+1} x_i - \sum_{i=1}^{n} x^i \right| = \sup_{x \in (-1,1)} |x^{n+1}| = 1.$$

Therefore, take $\varepsilon = 1/2$, and $\forall N \, \varepsilon \, \mathbb{N}$ put $n = N+1$ and $m = N+1$. We see that $\|f_n - f_m\|_{\sup} = 1 > 1/2 = \varepsilon$.

In conclusion, use M-test to prove uniform convergence.

17.4 Power Series and Uniform Convergence

Recall, from Analysis 2, that a power series is the series of functions of the form $\sum_{n=1}^{\infty} a_n x^n$, where a_n is sequence of real numbers. We define a radius of convergence R of the series such that $\sum_{n=1}^{\infty} a_n x^n$ converges absolutely on $(-R, R)$ and diverges for $|x| > R$.

Example: Consider $\sum_{n=0}^{\infty} x^n$. The series converges pointwise on (–1, 1), but this convergence is not uniform, whereas on $[-r, r]$ converges uniformly.

Theorem 6: Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with a radius of convergence R. Then for any $0 \le r < R$ the series converges uniformly on $[-r, r]$.

Proof: Fix $r \in (-R, R)$ and define a sequence $M_n = |a_n| r^n$. $\sum_{n=0}^{\infty} M_n$ converges absolutely by our choice of r and we get

$$\forall x \in [-r, r] : |a_n x^n| \le |a_n| r^n = M_n \text{ and } \sum_{n=0}^{\infty} M_n < \infty.$$

Therefore by Weierstrass M-test, the power series $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on [–r, r].

Notes

17.5 Continuous but Nowhere Differentiable Function

Theorem 7: There is a function f : ℝ → ℝ which is continuous but nowhere differentiable.

Proof: The idea of a proof is to find a function with a kind of fractal behaviour. Let g(x) = |x| on [–1, 1] extended by 2-periodicity on ℝ and let

$$f(x) = \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n g(4^n x)$$

Figure 17.3

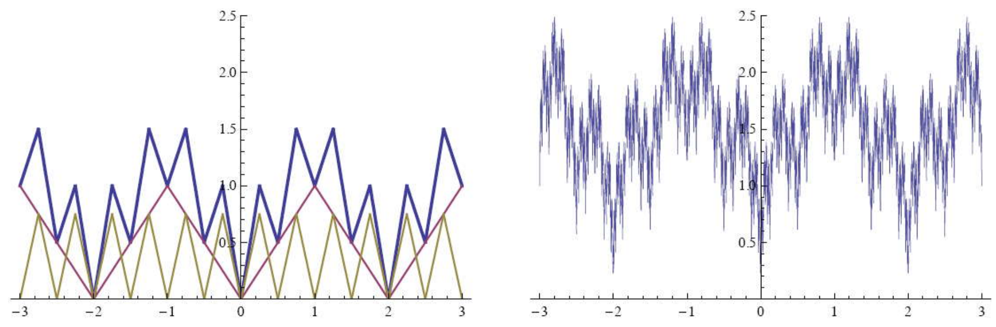


Figure 17.3 denotes the partial sums of f(x) by $s_n(x) = \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^i g(4^i x)$. On the left-hand side, we start with the red $s_0(x) = g(x)$. Then refine g(x) to the yellow $\frac{3}{4}g(4x)$. The iteration is obtained by adding these two together into the blue one, that is $s_1(x) = g(x) + \frac{3}{4}g(4x)$. $s_2(x)$ is obtained by adding refinement of $\frac{3}{4}g(4x)$ which is $\frac{9}{16}g(16x)$. Repeat this process at infinitum and get the limit function f(x) visualised on the right-hand side.

Now, we prove that the series is convergent and that the limit function f is continuous, but not differentiable.

Claim 1: The series $\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n g(4^n x)$ converges uniformly on ℝ.

Since

$$\left|\left(\frac{3}{4}\right)^n g(4^n x)\right| \le \left(\frac{3}{4}\right)^n \text{ and } \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n < \infty$$

Claim 2: The limit function f is continuous on ℝ.

Firstly, we prove that f is continuous on arbitrary interval [–M, M]. Each practical sum $s_n(x) = \sum_{k=1}^{n}\left(\frac{3}{4}\right)^k g(4^k x)$ is continuous on [–M, M] and $s_n \to f$ uniformly. Then we see, that the limit function f is continuous on [–M, M]. So for any x ∈ ℝ take sufficiently large s.t. x ∈ (–M, M). Continuity of f on [–M, M] implies continuity in x. Therefore, f is continuous on ℝ.

Claim 3: The limit function f is not differentiable.

Let $x \in \mathbb{R}$. Let us show that f is not differentiable at x. We will construct a sequence h_m, such that $h_m \to 0$ and

$$\left|\frac{f(x+h_m)-f(x)}{h_m}\right| \to \infty \text{ as } m \to \infty$$

Consider interval $(4^m x - \frac{1}{2}, 4^m x + \frac{1}{2}]$. Clearly, it is a half-closed interval of length 1 and therefore, can contain, only 1 integer. Define

$$h_m = \begin{cases} +\frac{1}{2}4^{-m}, \text{if there is no integer in } (4^m x, \ 4^m x + \frac{1}{2}) \\ -\frac{1}{2}4^{-m}, \text{if there is no integer in } (4^m x, \ -\frac{1}{2}, 4^m x) \end{cases}$$

We see that $h_m \to 0$ as $m \to \infty$. Let us define a_n as

$$a_n = \left(\frac{3}{4}\right)^n \frac{g(4^n(x+h_m)) - g(4^n x)}{h_m}.$$

Then we can rewrite the derivative as

$$\left|\frac{f(x+h_m)-f(x)}{h_m}\right| = \left|\sum_{n=0}^{\infty} a_n\right|.$$

Note that $|g(x) - g(y)| = |x - y|$, if $x, y \in [k, k+1]$ for $k \in \mathbb{Z}$ and $g|(x) - g(y)| \le |x - y|$, otherwise see Figure 17.4.

Figure 17.4

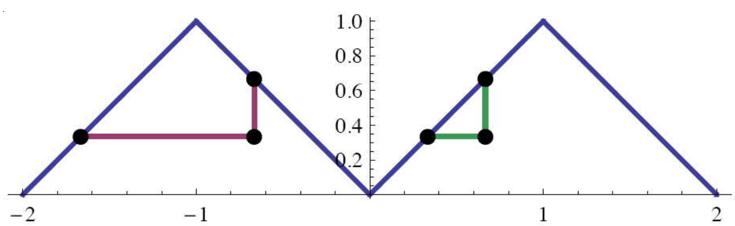


Figure that $|g(x) - g(x)| \le |x - y|$ with equality when $x, y \in [k, k+1]$, for some $k \in \mathbb{Z}$.

Let us prove the following three points,

(a) $a_n = 0$, if $n > m$

$g(4^n(x + h_m)) - g(4^n x \pm \underbrace{\tfrac{1}{2}4^{n-m}}_{\text{div. by 2}}) - g(4^n x) = 0$, due to 2-periodicity.

(b) $|a_n| = 3^m = 3^n$, if $n = m$

$g(4^n(x + h_m)) - g(4^n x \pm \underbrace{\tfrac{1}{2}4^{n-m}}_{=1} - g(4^n x) = 0,$

$$= |g(4^m x \pm \frac{1}{2}) - g(4^m x)|.$$

According to the definition of h_m, we see that interval with endpoints $4^m x \pm \frac{1}{2}$ and $4^m x$ does not contain any integer so by Figure 17.4 we obtain

Notes

$$\left|g(4^m x \pm \frac{1}{2}) - g(4^m x)\right| = \left|4^m x \pm \frac{1}{2} - 4^m x\right| = \frac{1}{2}$$

and finally

$$|a_n| = \left(\frac{3}{4}\right)^m \frac{1}{2} \frac{1}{\frac{1}{2}4^m} = 3^m = 3^n$$

(c) $|a_n| \le 3^n$, if $n < m$.

$$|g(4^n(x + h_m)) - g(4^n x)| = |g(4^n x \pm \frac{1}{2}4^{n-m}) - g(4^n x)| \le |4^m x \pm \frac{1}{2}4^{n-m} - 4^m x| = \frac{1}{2}4^{n-m},$$

and therefore

$$|a_n| \le \left(\frac{3}{4}\right)^n 4^{m-n} \frac{1}{2} \frac{1}{\frac{1}{2}4^m} = 3^n$$

Putting all things together we get

$$\left|\frac{f(x+h_m) - f(x)}{h_m}\right| = \left|\sum_{n=0}^{\infty} a_n\right| = \left|a_1 + \ldots + a_m \underbrace{+a_{m+1} + a_{m+2} + \ldots}_{=0 \text{ by (a)}}\right| =$$

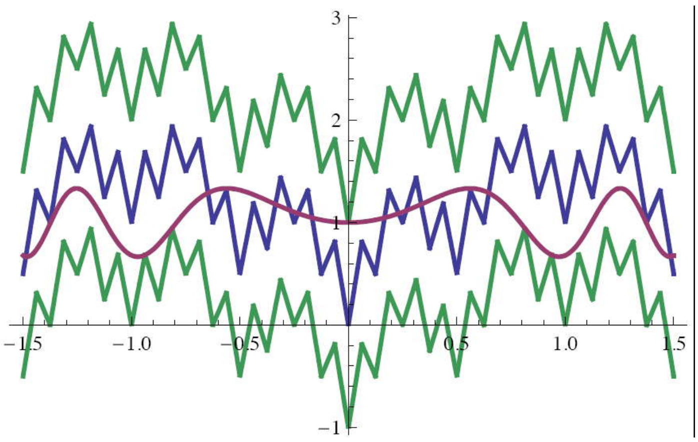
$$= ||a_1 + a_2 + \ldots + a_m \ge |a_m| - |a_1 + \ldots + a_{m-1}| \ge$$

$$\ldots \ge |a_m| - |a_1| - |a_2| - \ldots - |a_{m-1}| \overset{\text{by (b)}}{=} 3^m - |a_1| - |a_2| - \ldots - |a_{m-1}| \ge$$

$$\overset{\text{by (c)}}{\ge} 3^m - 3^1 - 3^2 - \ldots - 3^{m-1} = 3^m - \frac{3^m - 3}{3-1} = \frac{3}{2}(3^{m-1} - 1) \to \infty.$$

Remark: In original constructive proof of this theorem in 1872, Karl Weierstrab used f(x) = $\sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$ with a a ∈ (0, 1) and with positive odd integer b both satisfying ab > 1 + 3/4π. Interesting is, that despite of the differentiability of cosine, the limit function will not be differentiable anywhere.

Figure 17.5



Notes

In Figure 17.5, uniform approximation of continuous function f (in blue) by polynomial (in red) in e-tube around f (in green).

Self Assessment

Fill in the blanks:

1. Let $I \subset \mathbb{R}$ and $\{g_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on I. Then $\sum_{n=1}^{\infty} g_n$ is a ....

2. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $I \subset \mathbb{R}$. Then .... is called a uniform Cauchy sequence if:

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N : \|f_n - f_m\|_{\sup} < \varepsilon.$$

3. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on $I \subset \mathbb{R}$. Then $\{f_n\}_{n=1}^{\infty}$ .... of I if and only if $\{f_n\}_{n=1}^{\infty}$ is a uniform Cauchy sequence on *I*.

4. Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with a .... *R*. Then for any $0 \leq r < R$ the series converges uniformly on $[-r, r]$.

5. There is a function f : $\mathbb{R} \to \mathbb{R}$ which is continuous but nowhere ....

17.6 Summary

- Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued function on $[a, b]$. If (a) all f_n are differentiable, (b)′ all f_n' are continuous, (c) $f_n' \to h$ uniformly, for some function $h:[a, b] \to \mathbb{R}$, (d) $\exists c \in [a, b]$ s.t. $f_n(c)$ converges then f_n converges uniformly to some $f:[a, b] \to \mathbb{R}$, in addition, this uniform limit *f* is differentiable and $f'(x) = h(x)$.
- Let $I \subset \mathbb{R}$ and $\{g_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on I. Then

$$\sum_{n=1}^{\infty} g_n$$

  is a series of functions.
- Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with a radius of convergence *R*. Then for any $0 \leq r < R$ the series converges uniformly on $[-r, r]$.
- There is a function f : $\mathbb{R} \to \mathbb{R}$ which is continuous but nowhere differentiable.

17.7 Keywords

Series of Functions: Let $I \subset \mathbb{R}$ and $\{g_n\}_{n=1}^{\infty}$ be a sequence of real-valued functions on I. Then

$$\sum_{n=1}^{\infty} g_n$$

is a series of functions.

Continuity for Series: If (a) all g_n are are continuous and (b) $\sum_{n=1}^{\infty} g_n$ converges uniformly then $\sum_{n=1}^{\infty} g_n$ is continuous.

Notes

Differentiability for Series: If (a) all g_n are differentiable, (b) all g_n' are continuous, (c) $\sum_{n=1}^{\infty} g_n'$ converges uniformly and (d) $\exists c \in [a,b]$ s.t. $\sum_{n=1}^{\infty} g_n(c) < \infty$ then $\sum_{n=1}^{\infty} g_n$ converges uniformly and $\left(\sum_{n=1}^{\infty} gn\right)' = \sum_{n=1}^{\infty} g_n'$.

17.8 Review Questions

1. Prove that $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = \frac{1}{n}\sin(n^2 x)$.

2. Consider a series $\sum_{n=0}^{\infty} x^n$ on [–1/2, 1/2] and on (–1, 1). We want to investigate convergence of this series, whether it is pointwise of n uniform and eventually, what is the limit.

3. Consider series $\sum_{i=1}^{\infty} \frac{\sin(nx)}{2^n}$ on $\mathbb{R}$. By Weierstrass M-test, we see that this series converges uniformly on $\mathbb{R}$ since

$$\left|\frac{\sin(nx)}{2^n}\right| \le \frac{1}{2^n} \text{ and } \sum_{n=1}^{\infty} \frac{1}{2^n} < \infty$$

4. Consider series $\sum_{n=1}^{\infty} \frac{1}{n^2 + x}$ on $[0, \infty)$. By Weierstrass M-test, we can obtain uniform convergence of this series on $[0, \infty)$ since

$$\left|\frac{1}{n^2 + x}\right| \le \frac{1}{n^2} \text{ and } \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

Answers: Self Assessment

1. series of functions
2. $\{f_n\}_{n=1}^{\infty}$
3. converges uniformly
4. radius of convergence
5. differentiable

17.9 Further Readings

Books

Walter Rudin: Principles of Mathematical Analysis (3rd edition), Ch. 2, Ch. 3. (3.1-3.12), Ch. 6 (6.1 - 6.22), Ch.7(7.1 - 7.27), Ch. 8 (8.1- 8.5, 8.17 - 8.22).

G.F. Simmons: Introduction to Topology and Modern Analysis, Ch. 2(9-13), Appendix 1, p. 337-338.

Shanti Narayan: A Course of Mathematical Analysis, 4.81-4.86, 9.1-9.9, Ch.10,Ch.14, Ch.15(15.2, 15.3, 15.4)

T.M. Apostol: Mathematical Analysis, (2nd Edition) 7.30 and 7.31.

S.C. Malik: Mathematical Analysis.

H.L. Royden: Real Analysis, Ch. 3, 4.