MSc. Previous Year

Mathematics, MM - 02

REAL ANALYSIS

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल

MADHYA PRADESH BHOJ (OPEN) UNIVERSITY - BHOPAL

COURSE WRITERS

S L Gupta, Reader, Department of Mathematics, Shivaji College, University of Delhi
Nisha Rani, Lakshmibai College, University of Delhi
Units (1.5, 2.5-2.7, 2.13, 3.10)
Dr Bikas Chandra Bhui, Head, Department of Mathematics, Meghnand Saha Institute of Technology, Kolkata
Prof Dipak Chatterjee, Distinguished Professor, St Xavier's College, Kolkata
Units (2.0-2.3, 3.0-3.8, 3.11)
Manisha Pant, Former Lecturer, Department of Mathematics, H.N.B. Garhwal University Srinagar Garhwal (A Central University), U.K.
Units (1.0-1.4, 1.6-1.14, 2.4, 2.8-2.12, 2.14-2.21, 3.9, 3.12, 3.13, 3.14, 3.15-3.19, 4, 5)

Published by Registrar, MP Bhoj (open) University, Bhopal in 2020

VIKAS®

Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Real Analysis

Syllabi	Mapping in Book
UNIT - 1: Definition and Existence of Riemann-Stieltjes Integral, Properties of the Integral, Integration and Differentiation, The Fundamental Theorem of Calculus, Integration of Vector-Valued Functions, Rectifiable Curves. Rearrangements of Terms of a Series, Riemann's Theorem.	**Unit-1:** Riemann-Stieltjes Integral **(Pages 3-34)**
UNIT - 2: Sequences and Series of Functions, Pointwise and Uniform Convergence, Cauchy Criterion for Uniform Convergence, Weierstrass M-Test, Abel's and Dirichlet's Tests for Uniform Convergence, Uniform Convergence and Continuity, Uniform Convergence and Riemann-Stieltjes Integration, Uniform Convergence and Differentiation, Weierstrass Approximation Theorem, Power Series, Uniqueness Theorem for Power Series, Abel's and Tauber's Theorems.	**Unit-2:** Sequences and Series of Functions **(Pages 35-94)**
UNIT - 3: Functions of Several Variables, Linear Transformations, Derivatives in an open Subset or R^n, Chain Rule, Partial Derivatives, Interchange of the Order of Differentiation, Derivatives of Higher Order, Taylor's Theorem, Inverse Function Theorem, Implicit Function Theorem, Jacobians, Extremum Problems with Constraints, Lagrange's Multiplier Method, Differentiation of Integrals, Partitions of Unity, Differential Forms, Stokes' Theorem.	**Unit-3:** Functions of Several Variables and Higher Differentials **(Pages 95-153)**
UNIT - 4: Lebesgue Outer Measure, Measurable Sets, Regularity, Measurable Functions, Borel and Lebesgue Measurability, Non-Measurable Sets. Integration of Non-Negative Functions, The General Integral, Integration of Series, Riemann and Lebesgue Integrals. The Four Derivatives, Functions of Bounded Variation, Lebesgue Differentiation Theorem, Differentiation and Integration.	**Unit-4:** Integration of Non-Negative Functions **(Pages 155-197)**
UNIT - 5: Measures and Outer Measures, Extension of a Measure, Uniqueness of Extension, Completion of a Measure, Measure Spaces, Integration with respect to a Measure. The L^p-Spaces, Convex Functions, Jensen's Inequality, Hölder and Minkowski Inequalities, Completeness of L^p, Convergence in Measure, Almost Uniform Convergence.	**Unit-5:** Measures and Outer Measures **(Pages 199-240)**

CONTENTS

INTRODUCTION

NOTES

In mathematical analysis, the term 'Real Analysis' refers to the specific branch of mathematical evaluations and unique analysis that examines the typical behaviour of real numbers, sequences and series of real numbers, and the real functions. Some specific and distinctive properties of real valued sequences and functions used in the context of real analysis includes convergence, divergence, limits, continuity, smoothness, differentiability, integrability and measurability.

In real analysis, the theorems typically depend on the properties of the real number system, which should be determined, recognised and established. Characteristically, the real number system comprises of an uncountable set ($\mathbb{R}$), in addition to two binary operations denoted by '+' and '•', and an order which is denoted by '<'. The operations and analysis on the real numbers produce a field while with the order it produces an ordered field. Principally, the real number system is referred as the unique complete ordered field for the reason that any other complete ordered field is isomorphic to it. Instinctively, completeness implies that there are no 'Gaps' in the real numbers. This is the unique property of real numbers which distinguishes the real numbers from other ordered fields. Additionally, the properties of real numbers are critical and essentially significant for proving numerous key and basic properties of the functions that are analysed using the real numbers. The completeness property of the reals is often appropriately and conveniently stated and typically expressed as the Least Upper Bound (LUB) property. Furthermore, in real analysis, the order-theoretic properties produce a number of fundamental results or solutions typically based on the monotone convergence theorem, the intermediate value theorem, the mean value theorem, etc. Many of the theorems of real analysis are consequences of the topological properties of the real number line.

A sequence is defined as a function whose domain is considered as a countable and totally ordered set. Generally, the domain is defined to be the natural numbers, even though it is also occasionally appropriate to consider the bidirectional sequences indexed by means of the set of all integers, including negative indices. Generally, a limit is the value that a function or a sequence 'Approaches' as the input. This value can include the symbols '$\pm\infty$' while addressing the behaviour of a function or sequence as the variable increases or decreases without bound. The concept of a limit is fundamental to calculus and its conventional standard definition is specifically used in order to define notions like continuity, derivatives and integrals. For limits, the concept was introduced specifically for functions by Sir Isaac Newton and Gottfried Wilhelm (von) Leibniz, at the end of the 17th century, for developing the infinitesimal calculus. For sequences, the concept was introduced by Baron Augustin-Louis Cauchy and was later made rigorous and established at the end of the 19th century by Bernard Bolzano and Karl Theodor Wilhelm Weierstrass, who gave the modern ε–δ definition, which follows.

The term series validates and formalizes the imprecise notion of finding the sum of an endless sequence of numbers. In modern terminology, any ordered infinite sequence (a_1, a_2, a_3, …….) of terms, i.e., numbers, functions or anything

NOTES

that can be added, defines a series which refers to the operation of adding the a_i one after the other. To emphasize that there are an infinite number of terms, a series may be called an 'Infinite Series'. Series are classified not only by whether they converge or diverge, but also by the properties of the terms an absolute or conditional convergence; type of convergence of the series – pointwise or uniform; the class of the term a_n, i.e., whether it is a real number, arithmetic progression, trigonometric function, etc.

This book, *Real Analysis*, is divided into five units. The topics discussed include definition and existence of Riemann-Stieltjes integral, the fundamental theorem of calculus, integration of vector valued functions, rectifiable curves, rearrangements of terms of a series, Riemann's theorem, sequence and series of functions, pointwise and uniform convergence, Cauchy criterion for uniform convergence, Weierstrass's M test, Abel's and Dirichlet's tests for uniform convergence, uniform convergence and continuity, functions of several variables, derivatives in an open subset of R^n, partial derivatives, higher order differentials, Taylor's theorem, explicit and implicit functions, implicit function theorem and inverse function theorem, change of variables, extreme values of explicit and stationary values of implicit functions, Lagrange's multipliers method, Jacobian and its properties, Lebesgue outer measure, measurable sets, measurable functions, Borel and Lebesgue measurability, non-measurable sets, integration of non-negative functions, Reimann and Lebesgue integrals, functions of bounded variation, measures and outer measures, uniqueness of extension, the L^p-spaces, Jensen's inequality, Holder and Minkowski inequalities, completeness of L^p, and the almost uniform convergence.

The book follows the Self-Instructional Mode (SIM) format wherein each unit begins with an 'Introduction' to the topic. The 'Objectives' are then outlined before going on to the presentation of the detailed content in a simple and structured format. 'Check Your Progress' questions are provided at regular intervals to test the student's understanding of the subject. 'Answers to Check Your Progress Questions', a 'Summary', a list of 'Key Terms', and a set of 'Self-Assessment Questions and Exercises' are provided at the end of each unit for effective recapitulation.

UNIT 1 RIEMANN-STIELTJES INTEGRAL

NOTES

Structure

1.0 INTRODUCTION

In real analysis, the Riemann–Stieltjes integral is a generalization of the Riemann integral, named after Bernhard Riemann and Thomas Joannes Stieltjes. The definition of this integral was first published in 1894 by Stieltjes. It serves as an instructive and useful precursor of the Lebesgue integral, and an invaluable tool in unifying equivalent forms of statistical theorems that apply to discrete and continuous probability.

The Riemann–Stieltjes integral appears in the original formulation of F. Riesz's theorem which represents the dual space of the Banach space $C[a, b]$ of continuous functions in an interval $[a,b]$ as Riemann–Stieltjes integrals against functions of bounded variation. Later, the theorem was reformulated in terms of measures. The Riemann–Stieltjes integral also appears in the formulation of the spectral theorem for non-compact self-adjoint or more commonly as the normal operators in a Hilbert space. In this theorem, the integral is considered with respect to a spectral family of projections.

The best simple existence theorem states that, If f is continuous and g is of bounded variation on $[a, b]$, then the integral exists. A function g is of bounded variation if and only if it is the difference between two (bounded) monotone functions. If g is not of bounded variation, then there will be continuous functions which cannot be integrated with respect to g. Basically, the integral is not properly defined if f and g share any points of discontinuity, but there are other conditions also.

An important generalization is the Lebesgue–Stieltjes integral, which generalizes the Riemann–Stieltjes integral in a method analogous to how the Lebesgue integral generalizes the Riemann integral. If improper Riemann–Stieltjes integrals are allowed, then the Lebesgue integral is not strictly more general than the Riemann–Stieltjes integral.

In this unit, you will study about the definition and existence of Riemann-Stieltjes integral, properties of the integral, integration and differentiation, the fundamental theorem of calculus, integration of vector valued functions, rectifiable curves, rearrangements of terms of a series, and Riemann's theorem.

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the definition and existence of Riemann-Stieltjes integral
- Define the properties of the integral
- Elaborate on the integration and differentiation
- Analyse the fundamental theorem of calculus
- Explain the integration of vector valued functions and rectifiable curves
- Comprehend on the rearrangements of terms of a series
- Discuss the Riemann's theorem

1.2 DEFINITION AND EXISTENCE OF RIEMANN-STIELTJES INTEGRAL

Definition 1: Let $[a, b]$ be a given interval. A partition P of $[a, b]$ is a finite set of points $x_0, x_1, x_2, ..., x_n$ such that,

$$a = x_0 \le x_1 \le x_2 \le ....\le x_n = b.$$

Definition 2: Let α be a monotonically increasing function on $[a, b]$.

Corresponding to any partition P of $[a, b]$,

$$\alpha_i = \alpha(x_i) - \alpha(x_{i-1}), \qquad i = 1, 2, ..., n.$$

Then $\alpha_i \ge 0$.

Let f be a bounded real valued function on $[a, b]$.

Let $$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$

$$L(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$

Where $M_i = \sup\{ f(x) / x \in [x_{i-1}, x_i]\}$

And $m_i = \inf\{ f(x) / x \in [x_{i-1}, x_i]\}$

We define the upper Riemann-Stieltjes integral of f as,

$$\overline{\int_a^b} f d\alpha = \inf U(P, f, \alpha)$$

And the lower Riemann-Stieltjes integral of f as,

$$\underline{\int_a^b} f d\alpha = \sup L(P, f, \alpha),$$

Where the infimum and supremum are taken over all partitions P of $[a, b]$.

NOTES

$$\text{If } \underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha$$

Their common value is denoted by $\int_a^b f d\alpha$ or $\int_a^b f(x) d\alpha(x)$.

This is called the Riemann-Stieltjes integral of f with respect to α on $[a, b]$.

If $\int_a^b f d\alpha$ exists, then f is said to be integrable with respect to α on $[a, b]$.

It is written as $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

Definition 3: A partition P^* is said to be a refinement of P, if $P^* \supseteq P$.

Note: Given two partitions P_1 and P_2 of $[a, b]$, their common refinement is given by the notation $P^* = P_1 \cup P_2$.

Theorem 1.1: If P^* is a refinement of P, then

$$U(P^*, f, \alpha) \leq U(P, f, \alpha)$$

And $U(P^*, f, \alpha) \leq U(P, f, \alpha)$.

Proof: Assume that P^* contains just one point more than P.

Let this be c and $x_{i-1} < c < x_i$

Let $M_i' = \sup\{f(x) / x \in [x_{i-1}, c]\}$

And $M_i'' = \sup\{f(x) / x \in [c, x_i]\}$.

Then $M_i' \leq M_i$ and $M_i'' \leq M_i$.

Consider $U(p^*, f, \alpha) = \sum_{\substack{i=1 \\ k \neq 1}}^{n} M_k \Delta\alpha_k + M_i'[\alpha(c) - \alpha(x_{i-1})] + M_i''[\alpha(x_i) - \alpha(c)]$

$$\leq \sum_{\substack{i=1 \\ k \neq 1}}^{n} M_k \Delta\alpha_k + M_i[\alpha(c) - \alpha(x_{i-1})] + M_i[\alpha(x_i) - \alpha(c)]$$

$$\leq \sum_{\substack{i=1 \\ k \neq 1}}^{n} M_k \Delta\alpha_k + M_i[\alpha(x_i) - \alpha(x_{i-1})]$$

$$\leq U(P, f, \alpha).$$

Similarly we can prove that,

$$L(P^*, f, \alpha) \geq L(P, f, \alpha).$$

Hence the theorem is proved.

1.3 PROPERTIES OF THE INTEGRAL

NOTES

The significant properties of the integrals are discussed in this section.

Theorem 1.2: Consider the following statement:

$$\underline{\int_a^b} f d\alpha \le \overline{\int_a^b} f d\alpha .$$

Proof: Following is the proof of Theorem 1.2.

Let P_1 and P_2 be any partitions of $[a, b]$.

Let $P^*=P_1 \cup P_2$.

Then P^* is the common refinement of P_1 as well as P_2.

Therefore by Theorem 1.1,

$$U(P^*, f, \alpha) \le U(P_1, f, \alpha) \quad ....(1.1)$$

And $L(P^*, f, \alpha) \ge L(P_2, f, \alpha)$ ....(1.2)

Also we know that,

$$L(P^*, f, \alpha) \le U(P^*, f, \alpha) \quad ....(1.3)$$

From Equations (1.1), (1.2) and (1.3), we get

$$L(P_2, f, \alpha) \le L(P^*, f, \alpha) \le U(P^*, f, \alpha) \le U(P_1, f, \alpha)$$

Therefore for any two partitions P_1 and P_2 of $[a, b]$, we have

$$L(P_2, f, \alpha) \le U(P_1, f, \alpha).$$

Keeping P_2 fixed and varying P_1 over all partitions of $[a, b]$,

$$L(P_2, f, \alpha) \le \inf U(P, f, \alpha)$$

Now this is true for all partitions P_2 of $[a, b]$.

Therefore,

$$\sup L(P, f, a) \le \inf U(P, f, \alpha).$$

Consequently,

$$\underline{\int_a^b} f d\alpha \le \overline{\int_a^b} f d\alpha .$$

Hence the theorem is proved.

Theorem 1.3: Show that $f \in \mathbf{R}(\alpha)$ on $[a, b]$ if and only if there exists a partition P of $[a, b]$ such that,

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

Proof:

Let $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

Then $$\underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha , \quad ...(1.4)$$

Where $$\overline{\int_a^b} f d\alpha = \inf U(P, f, \alpha)$$

And $$\underline{\int_a^b} fd\alpha = \sup L(P, f, \alpha),$$

Therefore, by definition of infimum and supremum, for given $\varepsilon > 0$, there exists a partition P_1 of $[a, b]$ such that,

$$U(P_1, f, \alpha) < \overline{\int_a^b} fd\alpha + \varepsilon/2 \quad ...(1.5)$$

And a partition P_2 of $[a, b]$ such that,

$$L(P_2, f, \alpha) > \underline{\int_a^b} fd\alpha - \varepsilon/2 \quad ...(1.6)$$

Let $P = P_1 \bigcup P_2$.

Then by Theorem 1.1,

$$U(P, f, \alpha) \leq U(P_1, f, \alpha) \quad ...(1.7)$$

And $$L(P, f, \alpha) \geq L(P_2, f, \alpha) \quad ...(1.8)$$

Therefore, from Equations (1.4), (1.5), (1.6), (1.7) and (1.8), we get

$$\begin{aligned} U(P, f, \alpha) &\leq U(P_1, f, \alpha) \\ &< \overline{\int_a^b} fd\alpha + \varepsilon/2 \\ &< \overline{\int_a^b} fd\alpha + \varepsilon/2 \\ &< L(P_2, f, \alpha) + \varepsilon/2 + \varepsilon/2 \\ &< L(P, f, \alpha) + \varepsilon. \end{aligned}$$

Consequently, there exists a partition P of $[a, b]$ such that,

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

Conversely, assume that there exists a partition P of $[a, b]$ such that,

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon. \quad ...(1.9)$$

For every partition P of $[a, b]$, we have

$$L(P, f, \alpha) \leq \underline{\int_a^b} fd\alpha \leq \overline{\int_a^b} fd\alpha \leq U(P, f, \alpha) \quad ...(1.10)$$

From Equations (1.9) and (1.10), we know that

$$0 \leq \overline{\int_a^b} fd\alpha - \underline{\int_a^b} fd\alpha \leq U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon.$$

This is true for every $\varepsilon > 0$.

Hence, $$\overline{\int_a^b} fd\alpha - \underline{\int_a^b} fd\alpha = 0.$$

Therefore,

$$\overline{\int_a^b} fd\alpha = \underline{\int_a^b} fd\alpha.$$

Subsequently, $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

Hence the theorem is proved.

Theorem 1.4: If f is continuous on $[a, b]$, then $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

NOTES

Proof: Let $\varepsilon > 0$

Consider that $\eta > 0$ such that $[\alpha(b) - \alpha(a)]\,\eta < \varepsilon$.

Since f is continuous on $[a, b]$ and $[a, b]$ is compact, then f is uniformly continuous on $[a, b]$.

Therefore, for this $\eta > 0$, there exists a $\delta > 0$ such that,

$$|f(x) - f(t)| < \eta \text{ whenever } x, t \in [a, b] \text{ with } |x - t| < \delta. \qquad ...(1.11)$$

If P is any partition of $[a, b]$ such that $\Delta x_i < \delta$,

Then $M_i - m_i = \sup \{|f(x) - f(t)| \,/\, x, t \in [x_{i-1}, x_i]\} \leq \eta$,

$i = 1, 2, ..., n$.

$$\text{Therefore,}\quad U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_1 - \sum_{i=1}^{n} M_i \Delta\alpha_1$$

$$= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_1$$

$$\leq \eta \sum_{i=1}^{n} \Delta\alpha_i$$

$$\leq \eta[\alpha(b) - \alpha(a)]$$

$$< \varepsilon.$$

Therefore, $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

Hence the theorem is proved.

Theorem 1.5: If f is monotonic on $[a, b]$ and if α is continuous on $[a, b]$, then $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

Proof: Let α be increasing on $[a, b]$. Let $\varepsilon > 0$ be given.

Consider that n is large enough such that,

Assume that there is a partition P such that $\Delta\,\alpha_i = [\alpha(b) - \alpha(a)]/n$.

Let f be increasing on $[a, b]$.

Hence $f(x_{i-1}) \leq f(x) \leq f(x_i)$ whenever $x_{i-1} \leq x \leq x_i$.

Consequently, $M_i = f(x_i)$ and $m_i = f(x_{i-1})$, $i = 1, 2, ..., n$.

$$\text{Therefore,}\quad U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i - \sum_{i=1}^{n} M_i \Delta\alpha_i$$

$$= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i$$

$$= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))[\alpha(b) - \alpha(a))/n]$$

$$= [(\alpha(b) - \alpha(a))/n][f(b) - f(a)]$$
$$< \varepsilon$$

Therefore, $f \in \mathbf{R}(\alpha)$ on $[a, b]$. Hence the theorem is proved.

Theorem 1.6: If $f_1 \in \mathbf{R}(\alpha)$ and $f_2 \in \mathbf{R}(\alpha)$ on $[a, b]$, then $f_1 + f_2 \in \mathbf{R}(\alpha)$ and

$$\int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha .$$

Proof: Let $f = f_1 + f_2$ and P be any partition of $[a, b]$.

Consider,

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_k \Delta\alpha_k$$

Where $\quad M_k = \sup\{f(x)/\ x \in [x_{k-1}, x_k]\}$

$$= \sup\{f_1(x) + f_2(x)\ /\ x \in [x_{k-1}, x_k]\}$$
$$\leq \sup\{f_1(x)/x \in [x_{k-1}, x_k]\} + \sup\{f_2(x)/\ x \in [x_{k-1}, x_k]\}$$
$$\leq M_k' + M_k''$$

Hence,

$$\sum_{k=1}^{n} M_k \Delta\alpha_k \leq \sum_{k=1}^{n} M_k' \Delta\alpha_k + \sum_{k=1}^{n} M_k'' \Delta\alpha_k$$

Therefore,

$$U(P, f, \alpha) \leq U(P, f_1, \alpha) + U(P, f_2, \alpha) \qquad ...(1.12)$$

Similarly, $\quad L(P, f, \alpha) \geq L(P, f_1, \alpha) + L(P, f_2, \alpha) \qquad ...(1.13)$

Since $f_1 \in \mathbf{R}(\alpha)$ on $[a, b]$ and $f_2 \in \mathbf{R}(\alpha)$ on $[a, b]$,

for given $\varepsilon > 0$, there exists partitions P_1 and P_2 such that,

$$U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha) < \varepsilon/2 \qquad ...(1.14)$$

And $\quad U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha) < \varepsilon/2 \qquad ...(1.15)$

Let $P = P_1 \cup P_2$

This implies that,

$$U(P, f_1, \alpha) \leq U(P_1, f_1, \alpha)$$

And $\quad U(P, f_2, \alpha) \leq U(P_2, f_z, \alpha)$

And $\quad L(P, f_1, \alpha) \geq L(P_1, f_1, \alpha)$

And $\quad L(P, f_2, \alpha) \geq L(P_2, f_2, \alpha).$

Therefore,

$$U(P, f_1, \alpha) - L(P, f_1, \alpha) \leq U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha)$$
$$< \varepsilon/2 \qquad ...(1.16)$$

And $\quad U(P, f_2, \alpha) - L(P, f_2, \alpha) \leq U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha)$

$$< \varepsilon/2 \qquad ...(1.17)$$

NOTES

From Equations (1.12), (1.13), (1.16) and (1.17), we get

$$U(P, f, \alpha) - L(P, f, \alpha) \leq [U(P, f_1, \alpha) + U(P, f_2, \alpha)] - [L(P, f_1, \alpha) + L(P, f_2, \alpha)]$$
$$\leq [U(P, f_1, \alpha) - L(P, f_1, \alpha)] + [U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha)]$$
$$< \varepsilon/2 + \varepsilon/2$$
$$< \varepsilon.$$

Hence, $\quad f = f_1 + f_2 \in \mathbf{R}(\alpha)$ on $[a, b]$.

For the same partition P of $[a, b]$,

$$U(P, f_1, \alpha) < L(P, f_1, \alpha) + \varepsilon/2$$
$$< \sup\{L(P, f_1, a)\} + \varepsilon/2$$
$$< \underline{\int_a^b} f_1 d\alpha + \varepsilon/2$$
$$< \int_a^b f_1 d\alpha + \varepsilon/2$$

[Since $f_1 \in \mathbf{R}(\alpha)$ on $[a, b]$, $\int_a^b f_1 d\alpha = \overline{\int_a^b} f_1 d\alpha = \underline{\int_a^b} f_1 d\alpha$.]

Similarly, $\quad U(P, f_2, \alpha) < \int_a^b f_2 d\alpha + \varepsilon/2$

Therefore,

$$\int_a^b f d\alpha = \inf\{U(P, f, \alpha)\}$$
$$< U(P, f, \alpha)$$
$$< U(P, f_1, \alpha) + U(P, f_2, \alpha)$$
$$< \int_a^b f_1 d\alpha + \varepsilon/2 + \int_a^b f_2 d\alpha + \varepsilon/2$$
$$< \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha + \varepsilon$$

Hence, $$\int_a^b f d\alpha \leq \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha \qquad \text{....(1.18)}$$

Replacing f_1 by $-f_1$ and f_2 by $-f_2$, we get

$$-\int_a^b f d\alpha \leq -\int_a^b f_1 d\alpha - \int_a^b f_2 d\alpha$$

Multiplying both sides by (–1), we get

$$\int_a^b f d\alpha \geq \int_a^b f_1 d\alpha \int_a^b f_2 d\alpha \qquad \text{....(1.19)}$$

From Equations (1.18) and (1.19), we know that

$$\int_a^b f d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha$$

Hence the theorem is proved.

Theorem 1.7: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$ then $cf \in \mathbf{R}(\alpha)$ on $[a, b]$, for any constant c and

$$= \int_a^b cf d\alpha .$$

NOTES

Proof: If $c = 0$, then the result is true.

Assume that $c > 0$.

Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$,

for given $\varepsilon > 0$, then there exists a partition P of $[a, b]$ such that,

$$U(P, f, \alpha) - \mathrm{L}(P, f, \alpha) < \varepsilon/c.$$

Consider,

$$U(cP, f, \alpha) - L(cP, f, \alpha) = \sum_{k=1}^{n} M'_k \, \Delta\alpha_k - \sum_{k=1}^{n} m'_k \, \Delta\alpha_k$$

Where $\quad M_k' = \sup\{f(x) / x \in [x_{k-1}, x_k]\}$

$= c \sup\{(cf)(x) / x \in [x_{k-1}, x_k]\}$

$= cM_k$

Similarly, $\quad m_k' = cm_k$

Therefore,

$$U(cP, f, \alpha) = c\sum_{k=1}^{n} M_k \, \Delta\alpha_k = cU(P, f, \alpha)$$

And, $$L(cP, f, \alpha) = c\sum_{k=1}^{n} m_k \, \Delta\alpha_k = cL(P, f, \alpha)$$

Consequently, $U(cp,f,\alpha) - L(cP,f,\alpha) = c[U(P,f,\alpha) - L(P,f,\alpha)] < \varepsilon.$

Hence $cf \in \mathbf{R}(a)$ on $[a, b]$.

Therefore for the same P,

$$\begin{aligned} U(cp,f,\alpha) &< L(cP, f, \alpha) + \varepsilon \\ &\leq \sup L(cP, f, \alpha)] + \varepsilon \\ &\leq c \sup L(P, f, \alpha)] + \varepsilon \\ &\leq c\int_a^b f d\alpha + \varepsilon \end{aligned}$$

Subsequently,

$$\begin{aligned} \inf U(cP,f,\alpha) &\leq U(cP,f,\alpha) \\ &\leq c\int_a^b f d\alpha + \varepsilon \end{aligned}$$

Therefore,

$$\int_a^b cf d\alpha \leq c\int_a^b f d\alpha \qquad ...(1.20)$$

Replacing f by $-f$, we get

$$-\int_a^b cf d\alpha \leq c[-\int_a^b f d\alpha]$$

Multiplying both sides by (–1), we get

$$\int_a^b cfd\alpha \ \geq c\int_a^b fd\alpha \qquad ...(1.21)$$

From Equations (1.20) and (1.21), we have

$$\int_a^b cfd\alpha \ \leq c\int_a^b fd\alpha\,.$$

Hence the theorem is proved.

Theorem 1.8 If $f_1 \in \mathbf{R}(\alpha)$ on $[a, b]$, $f_2 \in \mathbf{R}(\alpha)$ on $[a, b]$ and $f_1(x) \leq f_2(x)$ on $[a, b]$, then $\int_a^b f_1 d\alpha \leq \int_a^b f_2 d\alpha$

Proof: Let P be any partition of $[a, b]$.

Since $f_1(x) \leq f_2(x)$,

$$\sup\{f_1(x) / x \in [x_{k-1}, x_k]\} \leq \sup\{f_2(x) / x \in [x_{k-1}, x_k]\}$$

Therefore,

$$U(P,f_1,\alpha) \leq U(P,f_2,\alpha)$$

Consequently,

$$\inf U(P,f_1,\alpha) \leq U(P,f_2,\alpha)$$

Hence,

$$\overline{\int_a^b f_1 d\alpha} \leq \overline{\int_a^b f_2 d\alpha}$$

Consequently,

$$\int_a^b f_1 d\alpha \leq \int_a^b f_2 d\alpha$$

Since $f_1 \in \mathbf{R}(\alpha)$ on $[a, b]$, and $f_2 \in \mathbf{R}(\alpha)$ on $[a, b]$

$$\int_a^b f_1 d\alpha = \overline{\int_a^b f_1 d\alpha} \text{ and } \int_a^b f_2 d\alpha = \overline{\int_a^b f_2 d\alpha}\,].$$

Hence the theorem is proved.

Theorem 1.9: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and if $a < c < b$, then $f \in \mathbf{R}(\alpha)$ on $[a, c]$ and $f \in \mathbf{R}(\alpha)$ on $[c, b]$, and $\int_a^b fd\alpha + \int_a^b fd\alpha = \int_a^b fd\alpha$.

Proof: Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$, for given $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that,

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

Let $P_1 = P \cap [a, c]$ and $P_2 = P \cap [c, b]$.

The P_1 is a partition of $[a, c]$ and P_2 is a partition of $[c, b]$.

Also on $[a, b]$,

$$U(P_1,f,\alpha) - L(P_1,f,\alpha) < U(P,f,\alpha) - L(P,f,\alpha)] < \varepsilon.$$

And on $[c, b]$,

$$U(P_2,f,\alpha) - L(P_2,f,\alpha) < U(P,f,\alpha) - L(P,f,\alpha)] < \varepsilon.$$

Therefore,

$$f \in \mathbf{R}(\alpha) \text{ on } [a, c] \text{ and } [c, b].$$

For any partition P of $[a, b]$,

Since $\quad P = P_1 \cup P_2,$

$$\begin{aligned} U(P,f,\alpha) &= U(P_1,f,\alpha) + U(P_2,f,\alpha)] \\ &\geq \inf U(P_1,f,\alpha) + \inf U(P_2,f,\alpha) \\ &= \overline{\int_a^c} f d\alpha + \overline{\int_c^b} f d\alpha \end{aligned}$$

Therefore,

$$\inf U(P,f,\alpha) \geq \overline{\int_a^c} f d\alpha + \overline{\int_c^b} f d\alpha$$

Consequently,

$$\overline{\int_a^b} f d\alpha \geq \overline{\int_a^c} f d\alpha + \overline{\int_c^b} f d\alpha$$

Subsequently,

$$\int_a^b f d\alpha \geq \int_a^c f d\alpha + \int_c^b f d\alpha \qquad ...(1.22)$$

Similarly, using lower sums,

$$\begin{aligned} L(P,f,\alpha) &= L(\mathrm{P}_1,f,\alpha) + L(\mathrm{P}_2,f,\alpha)] \\ &\leq \sup L(P_1,f,\alpha) + \sup L(P_2,f,\alpha) \\ &\leq \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha \end{aligned}$$

Therefore,

$$\sup L(P,f,\alpha) \leq \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha$$

Subsequently,

$$\underline{\int_a^b} f d\alpha \leq \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha$$

Hence,

$$\underline{\int_a^b} f d\alpha \leq \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha \qquad ...(1.23)$$

From Equations (1.22) and (1.23), we get

$$\underline{\int_a^b} f d\alpha = \underline{\int_a^c} f d\alpha + \underline{\int_c^b} f d\alpha$$

Hence the theorem is proved.

Theorem 1.10: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and if $|f(x)| \le M$ on $[a, b]$, then $\left|\int_a^b f d\alpha\right| \le M[\alpha(b) - \alpha(a)]$.

NOTES

Proof: Let P be any partition of $[a, b]$,

Since $|f(x)| \le M$ on $[a, b]$,

$$M_k = \sup\{(x)/\, x \in [x_{k-1}, x_k]\}$$
$$\le M, \text{ for all } k = 1,2,\ldots,n.$$

Because $f \in \mathbf{R}(\alpha)$ on $[a, b]$,

$$\int_a^b f d\alpha = \overline{\int_a^b f d\alpha} = \inf U(P,f,\alpha)$$
$$\le U(P,f,\alpha)$$
$$= \sum_{k=1}^{n} M_k \, \Delta\alpha_k$$
$$\le \mathrm{M} \sum_{k=1}^{n} \Delta\alpha_k$$
$$= M[\alpha(b) - \alpha(\mathrm{a})] \qquad \text{....(1.24)}$$

Replacing f by $-f$, we get

$$-\int_a^b f d\alpha \le M[\alpha(b) - \alpha(a)] \qquad \text{....(1.25)}$$

From Equations (1.24) and (1.25), we have

$$\left|\int_a^b f d\alpha\right| \le M[\alpha(b) - \alpha(a)]$$

Hence the theorem is proved.

Theorem 1.11: If $f \in \mathbf{R}(\alpha_1)$ on $[a, b]$ and $f \in \mathbf{R}(\alpha_2)$ on $[a, b]$, then $f \in \mathbf{R}(\alpha_1 + \alpha_2)$ on $[a, b]$ and

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2$$

Proof: Let P be any partition of $[a, b]$.

Let $\alpha = \alpha_1 + \alpha_2$.

Then, $\Delta\alpha_k = (\alpha_1 + \alpha_2)(x_k) - (\alpha_1 + \alpha_2)(x_{k-1})$

$$= \alpha_1(x_k) + \alpha_2(x_k) - [\alpha_1(x_{k-1}) + \alpha_2(x_{k-1})]$$
$$= \alpha_1(x_k) - \alpha_1(x_{k-1}) + \alpha_2(x_k) - \alpha_2(x_{k-1})$$
$$= \Delta(\alpha_1)_k + \Delta(\alpha_2)_k$$

Consider, $U(P,f,\alpha) = \sum_{k=1}^{n} M_k \, \Delta\alpha_k$

$$= \sum_{k=1}^{n} M_k \, [\Delta(\alpha_1)_k + \Delta(\alpha_2)_k]$$

$$= \sum_{k=1}^{n} M_k \ \Delta(\alpha_1)_k + \sum_{k=1}^{n} M_k \ \Delta(\alpha_2)_k$$

$$= U(P,f,\alpha_1) + U(P,f,\alpha_2)$$

$$\leq \inf U(P,f,\alpha_1) + \inf U(P,f,\alpha_2)$$

$$= \overline{\int_a^b fd\alpha_1} + \overline{\int_a^b fd\alpha_2}$$

NOTES

Therefore,

$$\inf U(P,f,\alpha) \leq \overline{\int_a^b fd\alpha_1} + \overline{\int_a^b fd\alpha_2}$$

Consequently,

$$\overline{\int_a^b fd\alpha} \leq \overline{\int_a^b fd\alpha_1} + \overline{\int_a^b fd\alpha_2}$$

Hence,

$$\int_a^b fd\alpha \leq \int_a^b fd\alpha_1 + \int_a^b fd\alpha_2 \qquad ....(1.26)$$

Similarly considering the lower sums, we can prove that,

$$\int_a^b fd\alpha \geq \int_a^b fd\alpha_1 + \int_a^b fd\alpha_2 \qquad ....(1.27)$$

From Equations (1.26) and (1.27), we have,

$$\int_a^b fd\alpha = \int_a^b fd\alpha_1 + \int_a^b fd\alpha_2 .$$

Therefore,

$$\int_a^b fd(\alpha_1 + \alpha_2) = \int_a^b fd\alpha_1 + \int_a^b fd\alpha_2 .$$

Hence the theorem is proved.

Theorem 1.12: If $f \in \mathbf{R}(\alpha\}$ on $[a, b]$ and c is a positive constant,

then $f \in \mathbf{R}(c\alpha)\}$ on $[a, b]$ and $\int_a^b fd(c\alpha) = c\int_a^b fd\alpha$.

Proof: The proof follows from theorem 1.11.

Theorem 1.13: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and $g \in \mathbf{R}(\alpha)$ on $[a, b]$, then

(*a*) $f^2 \in \mathbf{R}(\alpha)$ on $[a, b]$.

(*b*) $fg \in \mathbf{R}(\alpha)$ on $[a, b]$.

(*c*) $|f| \in \mathbf{R}(\alpha)$ on $[a, b]$ and $|\int_a^b fd\alpha| \leq \int_a^b |f| d\alpha$

Proof: Theorem 1.13 can be proved as follows:

(*a*) Let P be any partition of $[a, b]$.

And also, $M_k(f)$ denotes $\sup\{ f(x)/ x \in [x_{k-1}, x_k]\}$

NOTES

And $m_k(f)$ denotes $\inf\{f(x)/\,x\in[x_{k-1}, x_k]\}$

Then $$M_k(f^2) = \sup\,\{f^2(x)/\,x\in[x_{k-1}, x_k]\}$$
$$= [M_k(|f|)]^2$$
$$m_k(f^2) = [m_k(|f|)]^2$$

Since $f \in \mathbf{R}\,(\alpha)$ on $[a, b]$,

Then f is bounded on $[a, b]$.

Therefore there exists $M > 0$ such that,
$$|f(x)| \leq M.$$
Then $$M_k(f) \leq M$$
And $$M_k(f) \leq M, \text{ for all } k = 1, 2, ..., n.$$

Also for given $\varepsilon > 0$, there exists a partion P of $[a, b]$ such that,
$$U(P,f, \alpha) - L(P,f, \alpha) < \varepsilon$$

Therefore,
$$\begin{aligned} M_k(f^2) - m_k(f^2) &= [M_k(|f|)]^2 - [m_k(|f|)]^2 \\ &= [M_k(|f|) + m_k(|f|)][M_k(|f|) - m_k(|f|)] \\ &< 2M[M_k(|f|) - m_k(|f|)] \end{aligned}$$

Consequently,
$$\begin{aligned} U(P, f^2, \alpha) - L(P, f^2, \alpha) &= \sum_{k=1}^{n}[M_k(f^2) - m_k(f^2)]\,\Delta\,\alpha_k \\ &< 2M \sum_{k=1}^{n}[M_k(f) - m_k(f)]\Delta\alpha_k \\ &\leq 2M[U(P, f, \alpha) - L(P, f, \alpha)] \\ &< 2M(\varepsilon / 2M) \\ &< \varepsilon. \end{aligned}$$

Therefore, $f^2 \in \mathbf{R}(\alpha)$ on $[a, b]$.

Hence Part (a) of Theorem 1.13 is proved.

(*b*) Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and $g \in \mathbf{R}(\alpha)$ on $[a, b]$, then

by Theorems 1.6 and 1.7,

$f + g \in \mathbf{R}(\alpha)$ on $[a, b]$, and $f - g \in \mathbf{R}(\alpha)$ on $[a, b]$.

Therefore by Part (a) of Theorem 1.13 proved above,
$$(f + g)^2 \in \mathbf{R}(\alpha) \text{ on } [a, b]$$
And $$(f - g)^2 \in \mathbf{R}(\alpha) \text{ on } [a, b]$$

Therefore again by Theorems 1.6 and 1.7,
$$(1/4)[(f + g)^2 - (f - g)^2] \in \mathbf{R}(\alpha) \text{ on } [a, b]$$
$$\text{i.e., } fg \in \mathbf{R}(\alpha) \text{ on } [a, b].$$

Hence Part (*b*) of Theorem 1.13 is proved.

(*c*) Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$, for given $\varepsilon > 0$, there exists a partition P of $[a, b]$, such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Let P be any partition of $[a, b]$.

Since $\quad \| f(x) | - | f(y) \| \leq | f(x) - f(y) |$,

$$\begin{aligned} M_k(|f|) - m_k(|f|) &= \sup\{ \| f(x) | - | f(y) \| \,/\, x, y \in [x_{k-1}, x_k]\} \\ &\leq \sup\{ | f(x) - f(y) | \,/\, x, y \in [x_{k-1}, x_k]\} \\ &\leq M_k(f) - m_k(f) \end{aligned}$$

Therefore,

$$U(P, |f|, \alpha) - L(P, |f|, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha)$$

Hence, $|f| \in \mathbf{R}(\alpha)$ on $[a,b] < \varepsilon$

Now for all x, $\quad f(x) \leq | f(x) |$

And $\qquad -f(x) \leq | f(x)$.

Therefore by applying Theorem 1.8, we get

$$\int_a^b f d\alpha \leq \int_a^b | f | d\alpha \text{ and}$$

$$-\int_a^b f d\alpha \leq \int_a^b | f | d\alpha$$

Consequently, $\left| \int_a^b f d\alpha \right| \leq \int_a^b | f | d\alpha$

Therefore, Part (*c*) of Theorem 1.13 is proved.

Hence the theorem is proved.

Theorem 1.14: Suppose φ is a strictly increasing continuous function that maps an interval $[A, B]$ onto $[a, b]$. Suppose α is monotonically increasing on $[a, b]$ and $f \in \mathbf{R}(\alpha)$ on $[a, b]$.

Define β and g on $[A, B]$ by,

$$\beta(y) = \alpha(\varphi(y)), \qquad g(y) = f(\varphi(y)).$$

Show that $g \in \mathbf{R}(\beta)$ and $\qquad \int_A^B g d\beta \leq \int_a^b f d\alpha.$

Proof: To each paritition $P = \{x_0, x_1, ...x_n\}$ of $[a, b]$, there exists a partition $Q = \{y_0, y_1, ...,y_n\}$ of $[A, B]$ such that,

$$x_i = \varphi(y_i).$$

All partitions of $[A, B]$ can be obtained in this way.

Since $\qquad g(y) = f(\varphi(y))$ on $[A, B]$

NOTES

The values taken by g on $[y_{i-1}, y_i]$ are the same as those taken by f on $[x_{i-1}, x_i]$. Therefore,

$$U(Q, g, \beta) = U(P, f, \alpha)$$

And $$L(Q, g, \beta) = L(P, f, \alpha). \quad ..(1.28)$$

Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$,

for given $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that,

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Consequently,

$$U(Q, g, \beta) - L(Q, g, \beta) = U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Hence, $g \in \mathbf{R}(\beta)$ on $[A, B]$.

Moreover from Equation (1.28), we have

$$\inf U(Q, g, \beta) = \inf U(P, f, \alpha)$$

Subsequently,

$$\overline{\int_B^B gd\beta} = \overline{\int_b^b fd\alpha}.$$

Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and $g \in \mathbf{R}(\beta)$ on $[A, B]$,

$$\int_A^B gd\beta = \overline{\int_b^b gd\beta}$$

And $$\int_a^b fd\alpha = \overline{\int_b^b fd\alpha}.$$

Therefore,

$$\int_A^B gd\beta = \int_a^b fd\alpha.$$

Hence the theorem is proved.

1.4 INTEGRATION AND DIFFERENTIATION

Theorem 1.15: Let $f \in R$ on $[a, b]$(i.e., f is Riemann-integrable on $[a, b]$).

For $a \leq x \leq b$, define $F(x) = \int_a^x f(t)dt$.

Then F is continuous on $[a, b]$.

Furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0 and

$$F'(x_0) = f(x_0).$$

Proof: Since $f \in R$ on $[a, b]$, f is bounded.

Therefore, there exists $M > 0$ such that,

$$|f(t)| \leq M \quad \text{for } a \leq t \leq b$$

NOTES

Subsequently, if $a \le x \le y \le b$, then

$$
\begin{aligned}
|F(y) - F(x)| &= \left| \int_a^y f(t)dt - \int_a^x f(t)dt \right| \\
&= \left| \int_x^y f(t)dt \right| \\
&\le \int_x^y |f(t)|\,dt \\
&\le M\int_x^y dt = M(y-x)
\end{aligned}
$$

i.e., $\quad |F(y) - F(x)| \le M(y-x)$

Therefore,

$$|F(y) - F(x)| < \varepsilon \quad \text{provided that } |y - x) < \varepsilon/M.$$

Hence F is continuous on $[a, b]$.

Suppose if f is continuous at x_0. Then for given $\varepsilon > 0$, there exists a $\delta > 0$ such that,

$$|f(t) - f(x_0)| < \varepsilon \text{ whenever } |t - x_0| < \delta.$$

Hence, if $x_0 - \delta < s \le x_0 \le t < x_0 + \delta$ and $a \le s < t \le b$,

$$
\begin{aligned}
\left| \frac{F(t) - F(s)}{t-s} - f(x_0) \right| &= \left| \frac{1}{t-s} \int_s^t f(t)dt - f(x_0) \right| \\
&= \left| \frac{1}{t-s} \int_s^t [f(t) - f(x_0)]dt \right| \\
&\le \frac{1}{t-s} \int_s^t |f(t) - f(x_0)|\,dt \\
&< \frac{\varepsilon}{t-s} \int_s^t dt \\
&< \frac{\varepsilon}{t-s} [t-s] \\
&< \varepsilon
\end{aligned}
$$

Consequently,

$$\left| \frac{F(t) - F(s)}{t-s} - f(x_0) \right| < \varepsilon \quad \text{whenever } x_0 - \delta < s \le x_0 \le t < x_0 + \delta$$

Therefore, $\quad F'(x_0) = f(x_0)$

Hence the theorem is proved.

1.5 THE FUNDAMENTAL THEOREM OF CALCULUS

NOTES

The fundamental theorem of calculus states that this theorem is specifically used to link or connect the concept or theory of differentiating a function or calculating the gradient with the theory of integrating a function or calculating the area under the curve.

Basically, the integration and the differentiation are the closely related operations, and each is essentially considered as the inverse of the other.

Theorem 1.16: If f is bounded and integrable on $[a, b]$ and there exists a function F such that $F' = f$ on $[a, b]$, then

$$\int_a^b f\,dx = F(b) - F(a).$$

Proof: Let p be any partition of $[a, b]$ then by Mean Value Theorem (MVT) on every $\delta_r \exists \quad \xi_r \in (x_{r-1}, x_r)$ such that,

$$F(x_r) - F(x_{r-1}) = f(\xi_r)\delta_r.$$

On summing for $r = 1, 2, ..., n$ this gives,

$$F(b) - F(a) = \sum_{r=1}^{n} f(\xi r)\delta_r.$$

Since f is bounded and integrable on $[a, b]$, therefore, when $\|P\| \to 0$, then we get

$$F(b) - F(a) = \int_a^b f dx, \quad \text{i.e., } \int_a^b f dx = F(b) - F(a).$$

Note that F' may differ from f at a set of points whose set of limit points is finite.

Theorem 1.17: If f is continuous on (a, b) and $c \in (a, b)$, then function F defined by $F(x) = \int_c^x f(t)\,dt$, which is derivable and $F'(x) = f(x)$ on (a, b).

Proof: If $x \in (a, b)$, let $(x + h) \in (a, b)$. Then,

$$F(x + h) - F(x) = \int_x^{x+h} f(t)\,dt$$

$$= hf(x + \theta h), \text{ for some } \theta \in (0, 1).$$

By continuity of f at x, $\lim_{h\to 0} f(x) + Qh) = f(x)$. Therefore,

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = f(x), \text{ i.e., } F'(x) = f(x) \;\forall\; x \in (a, b).$$

Corollary: If f is continuous on $[a, b]$ and $c \in [a, b]$ then the function F can be defined by,

$$F(x) = \int_c^x f(t)dt \;\forall\; x \in [a, b], \text{ is derivable and } F'(x) = f(x) \text{ on } [a, b].$$

It should be noted that the condition of continuity in the above theorem and in its corollary cannot be totally removed.

NOTES

Example 1.1: If f is defined on $[-1, 1]$ by

$$f(x) = 1 \text{ when } 1 \geq x \geq 0,$$
$$= 0 \text{ when } -1 \leq x < 0,$$

Then $$F(x) = \int_c^x f(t)dt = x \text{ when } 1 \geq x \geq 0.$$

$$= 0 \text{ when } -1 \leq x < 0.$$

The function F is not derivable at 0, and so the conclusions of the theorem and the corollary on $(-1, 1)$ and $[-1, 1]$ respectively, do not hold.

Example 1.2: For $\sin^{-1} x$, which denotes the inverse of the function $\sin x$ in $[0, \pi/2]$, note that

$$(sin^{-1}x)' = \frac{1}{\sqrt{1-x^2}}$$

Hence, $$\int_0^x \frac{dt}{\sqrt{1-t^2}} = \sin^{-1}x, \;\forall\; x \in [0, 1]$$

This gives another way of introducing the trigonometrical functions, through $\sin x$ defined as the inverse function of $\sin^{-1} x$ and $\sin^{-1} 1 = \pi/2$.

Besides continuity and derivability of the functions defined by means of integrals we can examine various other properties, such as uniform convergence of functional sequences defined by means of integrals.

Example 1.3: The sequence given below converges uniformly to 0 on $[0, a]$, where $a > 0$.

$$\int_0^x \frac{t}{1+n^2t}dt$$

Solution: Since $\forall\; x \in [0, a], a > 0,$

$$0 \leq \int_0^x \frac{t}{1+n^2t}dt \leq \int_0^x \frac{1}{n^2}dt < \frac{a}{n^2} \to 0 \text{ as } n \to \infty,$$

Therefore, for $\varepsilon > 0 \;\exists\; m \in N$ such that,

$$\left|\int_0^x \frac{t}{1+n^2t}dt\right| < \varepsilon \;\forall\; n \geq m\left(> \sqrt{\frac{a}{\varepsilon}}\right) \text{ and } \forall\; x \in [0, a].$$

Hence, $\int_0^x \frac{t}{1+n^2t}dt$ converges uniformly to 0, on $[0, a]$ where $a > 0$.

1.6 INTEGRATION OF VECTOR VALUED FUNCTIONS

NOTES

Definition: Let $f_1, f_2, \ldots\ldots, f_k$ be real valued functions on $[a, b]$.

Let $f = (f_1, f_2, \ldots\ldots, f_k)$ be the corresponding mapping of $[a, b]$ onto R^k.

If α increases monotonically on $[a, b]$ and if $f_j \in \mathbf{R}(\alpha)$ for $j = 1, 2, \ldots, k$, then we say that $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and define $\int_a^b f d\alpha$ as,

$$\int_a^b f d\alpha = \left(\int_a^b f_1 d\alpha,\ \int_a^b f_2 d\alpha\ ,\ldots\ldots,\ \int_a^b f_k d\alpha\right)$$

i.e., $\int_a^b f d\alpha$ is the point in R^k whose jth coordinate is $\int_a^b f_j d\alpha$.

Let $f = (f_1, f_2, \ldots\ldots\ldots, f_k)$ and $g = (g_1, g_2, \ldots\ldots\ldots\ldots, g_k)$ be vector valued functions on $[a, b]$.

Then by the method in which we have defined $\int_a^b f d\alpha$, we get the following results.

Theorem 1.18: If $f \in \mathbf{R}(\alpha)$ and $g \in \mathbf{R}(\alpha)$ on $[a, b]$, then $f + g \in \mathbf{R}(\alpha)$ and

$$\int_a^b (f+g) d\alpha = \int_a^b f d\alpha + \int_a^b g d\alpha\,,$$

Proof: Since $f \in \mathbf{R}(\alpha)$ and $g \in \mathbf{R}(\alpha)$ on $[a, b]$

$$f_j,\ g_j \in \mathbf{R}(\alpha) \text{ on } [a, b], \text{ for } j = 1, 2, \ldots., k.$$

Hence, $\quad f_j + g_j \in \mathbf{R}(\alpha)$ on $[a, b]$, for $j = 1, 2, \ldots\ldots, k$.

Therefore,

$$(f+g) \in \mathbf{R}(\alpha) \text{ on } [a, b]$$

And $\int_a^b (f+g) d\alpha = \left(\int_a^b (f_1 + g_1) d\alpha, \int_a^b (f_2 + g_2) d\alpha, \ldots, \int_a^b (f_k + g_k) d\alpha\right)$

$$= \left(\int_a^b f_1 d\alpha + \int_a^b g_1 d\alpha, \int_a^b f_2 d\alpha + \int_a^b g_2 d\alpha, \ldots\ldots, \int_a^b f_k d\alpha + \int_a^b g_k d\alpha\right)$$

$$= \left(\int_a^b f_1 d\alpha, \int_a^b f_2 d\alpha, \ldots.\ \int_a^b f_k d\alpha + \int_a^b g_1 d\alpha, \int_a^b g_2 d\alpha, \ldots, \int_a^b g_k d\alpha\right)$$

$$= \int_a^b f d\alpha + \int_a^b g d\alpha$$

Hence the theorem is proved.

In the similar way we can prove the following results.

Theorem 1.19: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$, then $cf \in \mathbf{R}(\alpha)$ on $[a, b]$, for any constant c and

$$\int_a^b cf d\alpha = c\int_a^b f d\alpha\,.$$

Theorem 1.20: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and if $a < c < b$, then $f \in \mathbf{R}(\alpha)$ on $[a, c]$ and on $[c, b]$, and

$$\int_a^c fd\alpha + \int_c^b fd\alpha = \int_a^b fd\alpha .$$

Theorem 1.21: If $f \in \mathbf{R}(\alpha_1)$ on $[a, b]$ and $f \in \mathbf{R}(\alpha_2)$ on $[a, b]$, then $f \in \mathbf{R}(\alpha_1 + \alpha_2)\}$ on $[a, b]$ and

NOTES

$$\int_a^b fd(\alpha_1 + \alpha_2) = \int_a^b fd\alpha_1 + \int_a^b fd\alpha_2 .$$

Theorem 1.22: If $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and c is positive constant, then $f \in \mathbf{R}(c\alpha)$ on $[a, b]$ and $\int_a^b fd(c\alpha) = c\int_a^b fd\alpha$.

Theorem 1.23: Let $f \in \mathbf{R}$ on $[a, b]$(i.e., f is Riemann-integrable on $[a,b]$).

For $a \le x \le b$, define $F(x) = \int_a^b f(t)dt$.

Then F is continuous on $[a, b]$.

Furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Theorem 1.24: If f and F maps $[a, b]$ onto $\mathbf{R}^k$, if $f \in \mathbf{R}$ on $[a, b]$ and if $F' = f$, then

$$\int_a^b f(t)dt = F(b) - F(a).$$

Theorem 1.25: If f maps $[a,b]$ onto $\mathbf{R}^k$, and if $f \in \mathbf{R}(\alpha)$ for some monotonically increasing function α on $[a, b]$,

Then $|f| \in \mathbf{R}(\alpha)$ on $[a, b]$

And $$\left|\int_a^b fd\alpha\right| \le \int_a^b |f|\, d\alpha.$$

Proof: If $f_1, f_2, \ldots\ldots, f_k$ are the components of f, then

$$|f| = (f_1^2 + f_2^2 + \ldots + f_k^2)^{1/2}$$

Since $f \in \mathbf{R}(\alpha)$ on $[a, b]$,

By definition,

each $f_j \in \mathbf{R}(\alpha\}$ for $j = 1, 2, \ldots, k$ and

By Theorem 1.13,

$f_j^2 \in \mathbf{R}(\alpha)$ for $j = 1, 2, \ldots, k$.

Consequently, by Theorem 1.6,

$$f_1^2 + f_2^2 + f_k^2 \in \mathbf{R}(\alpha).$$

Hence,

$$|f| = (f_1^2 + f_2^2 + \ldots + f_k^2)^{1/2} \in \mathbf{R}(\alpha).$$

Since square root of a continuous function is continuous on $[0, M]$, for every real M.

To prove that $\left|\int_a^b fd\alpha\right| \le \int_a^b |f|\, d\alpha$.

NOTES

Let $y = (y_1, y_2, ..., y_k)$, where $y_j = \int_a^b f_j d\alpha$

Then, $$y = \int_a^b f d\alpha$$

And $$|y|^2 = \sum_{j=1}^{k} y_j^2$$

$$= \sum_{j=1}^{k} y_j \int_a^b f_j d\alpha$$

$$= \int_a^b (\sum_{j=1}^{n} y_j f_j) d\alpha$$

From Schwarz inequality,

$$\sum_{j=1}^{k} y_j f_j \; | y || f(t) | \quad (a \le b)$$

Consequently,

$$| y |^2 \le | y | \int_a^b | f | d\alpha$$

Therefore, if $y \to 0$, dividing this inequality by $| y |$, we get

$$| y | \le \int_a^b | f | d\alpha$$

Subsequently,

$$| \int_a^b f d\alpha | \le \int_a^b | f | d\alpha$$

Hence the theorem is proved.

1.7 RECTIFIABLE CURVES

Definition 1: A continuous mapping γ of an interval $[a, b]$ into R^k is called a curve in R^k or γ is a curve on $[a, b]$.

If γ is one-to-one, then γ is called an arc.

If $\gamma(a) = \gamma(b)$, then γ is called a closed curve.

Definition 2: To each partition $P = \{x_0, x_1, x_2, ..., x_n\}$ of $[a, b]$ and to each curve γ on $[a, b]$, we associate a number $\Lambda(P, \gamma) = \sum_{i=1}^{n} | \gamma(x_i) - \gamma(x_{i-1}) |$.

Where $| \gamma(x_i) - \gamma(x_{i-1}) |$ = Distance between the points $\gamma(x_{i-1})$ and $\gamma(x_i)$

$\Lambda(P, \gamma)$ = Length of a polygonal path with vertices at $\gamma(x_0), \gamma(x_1), ..., \gamma(x_n)$.

As the partition P becomes finer and finer, the polygon approaches γ more and more closely.

The lenght of Λ is defined as,

$$\Lambda(\gamma) = \sup \Lambda(P, \gamma),$$

Where the supremum is taken over all partitions of $[a, b]$.

If $\Lambda(\gamma) < \infty$, then γ is said to be **rectifiable.**

NOTES

Theorem 1.26: If γ' is continuous on $[a, b]$, then γ is rectifiable and $\Lambda(\gamma)=\int_a^b |\gamma'(t)|\,dt$.

Proof:

If $a \le x_{i-1} < x_i \le b$, then

$$|\gamma(x_i) - \gamma(x_{i-1})| = \left|\int_{x_{i-1}}^{x_i} \gamma'(t)dt\right|$$

$$\le \int_{x_{i-1}}^{x_i} \gamma'(t)dt$$

Hence,

$$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|$$

$$\le \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt$$

$$\le \int_a^b |\gamma'(t)|\,dt$$

for every partition P of $[a, b]$.

Therefore,

$$\Lambda(\gamma) \le \int_a^b |\gamma'(t)|\,dt \qquad ....(1.29)$$

To prove the opposite inequality, let $\varepsilon > 0$ be given.

Since γ' is uniformly continuous on $[a, b]$, there exists a $\delta > 0$ such that,

$$|\gamma'(s) - \gamma'(t)| < \varepsilon \quad \text{whenever } |s - t| < \delta.$$

Let $P = \{x_0, x_1, ..., x_n\}$ be a partition of $[a\ b]$, with $\Delta x_i < \delta$ for all i.

Therefore, if $x_{i-1} \le t \le x_i$,

$$|\gamma'(t) - \gamma'(x_i)| < \varepsilon$$

Consequently,

$$|\gamma'(t)| - |\gamma'(x_i)| \le |\gamma'(t) - \gamma'(x_i)| < \varepsilon$$

Subsequently,

$$|\gamma'(t)| \le |\gamma'(x_i)| + \varepsilon.$$

Hence,

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt \le \int_{x_{i-1}}^{x_i} (|\gamma'(x_i)| + \varepsilon)dt$$

$$\le |\gamma'(t)|\,\Delta x_i + \varepsilon\,\Delta x_i$$

$$\le \int_{x_{i-1}}^{x_i} \gamma'(x_i)dt\,| + \varepsilon\,\Delta x_i$$

$$\le \left|\int_{x_{i-1}}^{x_i} [\gamma'(t) + \gamma'(x_i) - \gamma'(t)]dt + \varepsilon\,\Delta x_i\right.$$

NOTES

$$\leq \left| \int_{x_{i-1}}^{x_i} [\gamma'(t)dt \right| + \left| \int_{x_{i-1}}^{x_i} [\gamma'(x_i) - \gamma'(t)]dt \right| + \varepsilon \, \Delta x_i$$

$$\leq |\gamma(x_i) - \gamma(x_{i-1})| + \varepsilon \, \Delta x_i + \varepsilon \, \Delta x_i$$

$$\leq |\gamma(x_i) - \gamma(x_{i-1})| + 2\varepsilon \, \Delta x_i$$

Therefore,

$$\int_a^b |\gamma'(t)|\, dt = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |\gamma'(t)|\, dt$$

$$\leq \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})| + 2\varepsilon \sum_{i=1}^{n} \Delta x_i$$

$$\leq \Lambda\,(P, \gamma) + 2\varepsilon(b-a)$$

$$\leq \Lambda\,(\gamma) + 2\varepsilon(b-a).$$

Since $\varepsilon > 0$ was arbitrary, therefore we have,

$$\int_a^b |\gamma'(t)|\, dt \leq \Lambda\,(\gamma) \qquad ...(1.30)$$

From Equations (1.29) and (1.30), we get

$$\Lambda\,(\gamma) = \int_a^b |\gamma'(t)|\, dt$$

Hence the theorem is proved.

1.8 REARRANGEMENTS OF TERMS OF A SERIES

Riemann rearrangement theorem, named after 19th-century German mathematician Bernhard Riemann, says that if an infinite series of real numbers is conditionally convergent, then its terms can be arranged in a permutation so that the new series converges to an arbitrary real number or diverges.

As an example, the series 1 – 1 + 1/2 – 1/2 + 1/3 – 1/3 + ... converges to 0 for a sufficiently large number of terms, the partial sum gets arbitrarily near to 0; but replacing all terms with their absolute values gives 1 + 1 + 1/2 + 1/2 + 1/3 + 1/3 + ... , which sums to infinity. Thus the original series is conditionally convergent, and can be rearranged by taking the first two positive terms followed by the first negative term, followed by the next two positive terms and then the next negative term, etc. to give a series that converges to a different sum: 1 + 1/2 – 1 + 1/3 + 1/4 – 1/2 + ... = ln 2. More generally, using this procedure with p positives followed by q negatives gives the sum ln(p/q). Other rearrangements give other finite sums or do not converge to any sum.

Existence of a Rearrangement that Sums to Any Positive Real *M*

For simplicity, this proof assumes first that $a_n \neq 0$ for every n. The general case requires a simple modification, given below. Recall that a conditionally convergent

NOTES

series of real terms has both infinitely many negative terms and infinitely many positive terms. First, define two quantities a_n^+ and a_n^- by,

$$a_n^+ = \frac{a_n + |a_n|}{2}, \quad a_n^- = \frac{a_n - |a_n|}{2}.$$

That is, the series $\sum_{n=1}^{\infty} a_n^+$ includes all a_n positive, with all negative terms replaced by zeroes, and the series $\sum_{n=1}^{\infty} a_n^-$ includes all an negative, with all positive terms replaced by zeroes. Since $\sum_{n=1}^{\infty} a_n$ is conditionally convergent, then both the positive and the negative series diverge. Let M be a positive real number. Taking now sufficient positive terms so that their sum exceeds M. Suppose p terms, are required, then the following statement is considered true:

$$\sum_{n=1}^{p-1} a_n^+ \leq M < \sum_{n=1}^{p} a_n^+.$$

This is possible for any $M > 0$ because the partial sums of a^+n tend to $+\infty$ Discarding the zero terms one may write,

$$\sum_{n=1}^{p} a_n^+ = a_{\sigma(1)} + \cdots + a_{\sigma(m_1)}, \quad a_{\sigma(j)} > 0, \;\; \sigma(1) < \ldots < \sigma(m_1) = p.$$

Now adding the sufficient negative terms a^-n say q of them, so that the resulting sum is less than M. This is always possible because the partial sums of a^-n tend to $-\infty$ Now we have,

$$\sum_{n=1}^{p} a_n^+ + \sum_{n=1}^{q} a_n^- < M \leq \sum_{n=1}^{p} a_n^+ + \sum_{n=1}^{q-1} a_n^-.$$

Again, one can write,

$$\sum_{n=1}^{p} a_n^+ + \sum_{n=1}^{q} a_n^- = a_{\sigma(1)} + \cdots + a_{\sigma(m_1)} + a_{\sigma(m_1+1)} + \cdots + a_{\sigma(n_1)},$$

With,

$$\sigma(m_1 + 1) < \ldots < \sigma(n_1) = q.$$

The map σ is injective, and 1 belongs to the range of σ, either as image of 1 (if $a_1 > 0$) or as image of $m_1 + 1$ (if $a_1 < 0$). Now repeat the process of adding sufficient positive terms to exceed M, starting with $n = p + 1$, and then adding sufficient negative terms to be less than M, starting with $n = q + 1$. Extend σ in an injective manner so that all terms selected so far must be covered, and observe that a_2 must have been selected previously or now thus 2 belongs to the range of this extension. The process infinitely includes various such "Changes of Direction".

Existence of a Rearrangement that Diverges to Infinity

Let $\sum_{i=1}^{\infty} a_i$ be a conditionally convergent series. The following is a proof that there exists a rearrangement of this series that tends to ∞ a similar argument can be used to show that –∞ can also be attained.

Let $p_1 < p_2 < p_3 < \cdots$ be the sequence of indexes such that each a_{p_i} is positive, and define $n_1 < n_2 < n_3 < \cdots$ to be the indexes such that each a_{n_i} is negative (assuming that a_i is never 0). Each natural number will appear in exactly one of the sequences (p_i) and (n_i).

Let b_1 be the smallest natural number such that,

$$\sum_{i=1}^{b_1} a_{p_i} \geq |a_{n_1}| + 1.$$

Such a specific value must exist since (a_{p_i}), the subsequence of positive terms of (a_i) which diverges. Similarly, let b_2 be the smallest natural number such that,

$$\sum_{i=b_1+1}^{b_2} a_{p_i} \geq |a_{n_2}| + 1,$$

And so on. This leads to the following permutation:

$$(\sigma(1), \sigma(2), \sigma(3), \ldots) = (p_1, p_2, \ldots, p_{b_1}, n_1, p_{b_1+1}, p_{b_1+2}, \ldots, p_{b_2}, n_2, \ldots).$$

And the obtained rearranged series, $\sum_{i=1}^{\infty} a_{\sigma(i)}$ then diverges to ∞.

Existence of a Rearrangement that Fails to Approach Any Limit, Finite or Infinite

In fact, if $\sum_{n=1}^{\infty} a_n$ is conditionally convergent, then there is such a rearrangement of it that the partial sums of the rearranged series form a dense subset of $\mathbb{R}$.

1.9 RIEMANN'S THEOREM

Definition: A series $\sum_{n=1}^{\infty} a_n$ converges if there exists a value ℓ such that the sequence of the partial sums,

$$(S_1, S_2, S_3, \ldots), \quad S_n = \sum_{k=1}^{n} a_k,$$

Converges to ℓ. That is, for any $\varepsilon > 0$, there exists an integer N such that if $n \geq N$, then

NOTES

$$|S_n - \ell| \leq \epsilon.$$

Theorem 1.27: A series converges conditionally if the series $\sum_{n=1}^{\infty} a_n$ converges but the series $\sum_{n=1}^{\infty} |a_n|$ diverges.

A permutation is simply a bijection from the set of positive integers to itself. This specifies that if σ is a permutation, then for any positive integer b there exists exactly one positive integer a such that $\sigma(a) = b$. In particular, if $x \neq y$, then $\sigma(x) \neq \sigma(y)$.

Suppose that $(a_1, a_2, a_3, \ldots)$ is a sequence of real numbers and that $\sum_{n=1}^{\infty} a_n$ is conditionally convergent. Let M be a real number, then there exists a permutation σ such that,

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = \infty.$$

The sum can also be rearranged to diverge to $-\infty$ or to fail to approach any limit, finite or infinite.

In Riemann's theorem, the permutation used for rearranging a conditionally convergent series to obtain a given value in $\mathbf{R} \cup \{\infty, -\infty\}$ may have arbitrarily many non-fixed points, i.e., all the indexes of the terms of the series may be rearranged. It is possible to rearrange only the indexes in a smaller set so that a conditionally convergent series converges to an arbitrarily chosen real number or diverges to (positive or negative) infinity. The answer of this question is positive, Sierpiński proved that is sufficient to rearrange only some strictly positive terms or only some strictly negative terms.

Check Your Progress

1. Define the Riemann-Stieltjes integral.
2. State the integration and differentiation theorems.
3. What is the fundamental theorem of calculus?
4. Define on the integration of vector valued functions.
5. What do you understand by the rectifiable curves?
6. Define the rearrangements of terms of a series.
7. Give the definition of Riemann's theorem.

1.10 ANSWERS TO 'CHECK YOUR PROGRESS'

NOTES

1. Let $[a, b]$ be a given interval. A partition P of $[a, b]$ is a finite set of points $x_0, x_1, x_2, ..., x_n$ such that,

   $a = x_0 \leq x_1 \leq x_2 \leq ....\leq x_n = b$.

2. Let $f \in R$ on $[a, b]$(i.e., f is Riemann-integrable on $[a, b]$).

   For $a \leq x \leq b$, define $F(x) = \int_a^x f(t)dt$.

   Then F is continuous on $[a, b]$.

3. If f is bounded and integrable on $[a, b]$ and there exists a function F such that $F' = f$ on $[a, b]$, then

   $$\int_a^b f\,dx = F(b) - F(a).$$

4. Let $f_1, f_2, ..., f_k$ be real valued functions on $[a, b]$.

   Let $f = (f_1, f_2, ..., f_k)$ be the corresponding mapping of $[a, b]$ onto R^k.

   If α increases monotonically on $[a, b]$ and if $f_j \in \mathbf{R}(\alpha)$ for $j = 1, 2, ..., k$, then we say that $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and define $\int_a^b f d\alpha$ as

   $$\int_a^b f d\alpha = (\int_a^b f_1 d\alpha,\ \int_a^b f_2 d\alpha\ ,........, \int_a^b f_k d\alpha)$$

   i.e., $\int_a^b f d\alpha$ is the point in R^k whose jth coordinate is $\int_a^b f_j d\alpha$.

5. A continuous mapping γ of an interval $[a, b]$ into R^k is called a curve in R^k or γ is a curve on $[a, b]$.

   If γ is one-to-one, then γ is called an arc.

   If $\gamma(a) = \gamma(b)$, then γ is called a closed curve.

6. Riemann rearrangement theorem, named after 19th-century German mathematician Bernhard Riemann, says that if an infinite series of real numbers is conditionally convergent, then its terms can be arranged in a permutation so that the new series converges to an arbitrary real number or diverges.

7. A series $\sum_{n=1}^{\infty} a_n$ converges if there exists a value ℓ such that the sequence of the partial sums,

   $$(S_1, S_2, S_3, \ldots), \quad S_n = \sum_{k=1}^{n} a_k,$$

   Converges to ℓ. That is, for any $\varepsilon > 0$, there exists an integer N such that if $n \geq N$, then

   $$|S_n - \ell| \leq \epsilon.$$

1.11 SUMMARY

NOTES

- Let $[a, b]$ be a given interval. A partition P of $[a, b]$ is a finite set of points $x_0, x_1, x_2, ..., x_n$ such that,

  $a = x_0 \leq x_1 \leq x_2 \leq ....\leq x_n = b.$
- Let α be a monotonically increasing function on $[a, b]$.

  Corresponding to any partition P of $[a, b]$,

  $\alpha_i = \alpha(x_i) - \alpha(x_{i-1}), i = 1, 2, ..., n.$

  Then $\alpha_i \geq 0$.
- A partition P^* is said to be a refinement of P, if $P^* \supseteq P$.
  Given two partitions P_1 and P_2 of $[a, b]$, their common refinement is $P^* = P_1 \cup P_2$.
- Let $f \in R$ on $[a, b]$(i.e., f is Riemann-integrable on $[a, b]$).

  For $a \leq x \leq b$, define $F(x) = \int_a^x f(t)dt$.

  Then F is continuous on $[a, b]$.
- Furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0 and

  $F'(x_0) = f(x_0)$.
- If f is bounded and integrable on $[a, b]$ and there exists a function F such that $F' = f$ on $[a, b]$, then

  $$\int_a^b f\,dx = F(b) - F(a).$$
- If f is continuous on $[a, b]$ and $c \in [a, b]$ then the function F defined by

  $F(x) = \int_c^x f(t)dt \ \forall\ x \in [a, b],$ is derivable and $F'(x) = f(x)$ on $[a, b]$.
- Let $f_1, f_2, ..., f_k$ be real valued functions on $[a, b]$.

  Let $f = (f_1, f_2, ..., f_k)$ be the corresponding mapping of $[a, b]$ onto R^k.

  If α increases monotonically on $[a, b]$ and if $f_j \in \mathbf{R}(\alpha)$ for $j = 1, 2, ..., k$, then we say that $f \in \mathbf{R}(\alpha)$ on $[a, b]$ and define $\int_a^b f d\alpha$ as,

  $$\int_a^b f d\alpha = (\int_a^b f_1 d\alpha,\ \int_a^b f_2 d\alpha\ ,........, \int_a^b f_k d\alpha)$$

  i.e., $\int_a^b f d\alpha$ is the point in R^k whose jth coordinate is $\int_a^b f_j d\alpha$.
- A continuous mapping γ of an interval $[a, b]$ into R^k is called a curve in R^k or γ is a curve on $[a, b]$.

  If γ is one-to-one, then γ is called an arc.

  If $\gamma(a) = \gamma(b)$, then γ is called a closed curve.

NOTES

- Riemann rearrangement theorem, named after 19th-century German mathematician Bernhard Riemann, says that if an infinite series of real numbers is conditionally convergent, then its terms can be arranged in a permutation so that the new series converges to an arbitrary real number or diverges.
- Let $\sum_{i=1}^{\infty} a_i$ be a conditionally convergent series. If there exists a rearrangement of this series that tends to $-\infty$ then a similar argument can be used to show that $-\infty$ can also be attained.
- Let $p_1 < p_2 < p_3 < \cdots$ be the sequence of indexes such that each a_{p_i} is positive, and define $n_1 < n_2 < n_3 < \cdots$ to be the indexes such that each a_{n_i} is negative (assuming that a_i is never 0). Each natural number will appear in exactly one of the sequences (p_i) and (n_i).
- In fact, if $\sum_{n=1}^{\infty} a_n$ is conditionally convergent, then there is such a rearrangement of it that the partial sums of the rearranged series form a dense subset of $\mathbb{R}$.
- A series $\sum_{n=1}^{\infty} a_n$ converges if there exists a value ℓ such that the sequence of the partial sums

  $$(S_1, S_2, S_3, \ldots), \quad S_n = \sum_{k=1}^{n} a_k,$$

  Converges to ℓ. That is, for any $\varepsilon > 0$, there exists an integer N such that if $n \geq N$, then

  $$|S_n - \ell| \leq \epsilon.$$

1.12 KEY TERMS

- **Riemann-Stieltjes integral:** Let $[a, b]$ be a given interval. A partition P of $[a, b]$ is a finite set of points $x_0, x_1, x_2, \ldots, x_n$ such that,

  $a = x_0 \leq x_1 \leq x_2 \leq \ldots\ldots\ldots \leq x_n = b.$
- **Integration and differentiation theorem:** Let $f \in R$ on $[a, b]$(i.e., f is Riemann-integrable on $[a, b]$).

  For $a \leq x \leq b$, define $F(x) = \int_a^x f(t)dt$.

  Then F is continuous on $[a, b]$.
- **Fundamental theorem of calculus:** If f is bounded and integrable on $[a, b]$ and there exists a function F such that $F' = f$ on $[a, b]$, then

  $$\int_a^b f\,dx = F(b) - F(a).$$

- **Integration of vector valued functions:** Let $f_1, f_2, ..., f_k$ be real valued functions on $[a, b]$.

  Let $f = (f_1, f_2, ..., f_k)$ be the corresponding mapping of $[a, b]$ onto R^k.

- **Rectifiable curves:** A continuous mapping g of an interval $[a, b]$ into R^k is called a curve in R^k or g is a curve on $[a, b]$.

  If γ is one-to-one, then γ is called an arc.

  If $\gamma(a) = \gamma(b)$, then γ is called a closed curve.

- **Riemann's theorem:** Riemann rearrangement theorem, named after 19th-century German mathematician Bernhard Riemann, says that if an infinite series of real numbers is conditionally convergent, then its terms can be arranged in a permutation so that the new series converges to an arbitrary real number or diverges.

1.13 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Why is the Riemann-Stieltjes integral used?
2. Define the integration and differentiation theorems.
3. State the fundamental theorem of calculus.
4. Define the integration of vector valued functions.
5. What are on the rectifiable curves?
6. Define the rearrangements of terms of a series.
7. State the Riemann's theorem.

Long-Answer Questions

1. Briefly discuss the Riemann-Stieltjes integral giving appropriate examples.
2. Discuss the significance of the integration and differentiation theorems in real analysis.
3. Explain in detail about the fundamental theorem of calculus with the help of theorems and proofs.
4. Discuss the integration of vector valued functions with the help of theorems and proofs.
5. Explain the rectifiable curves with the help of theorems and examples.
6. Briefly explain the concept of rearrangements of terms of a series giving relevant theorems, proofs and exampels.
7. State and prove the Riemann's theorem.
8. Evaluate the following using the fundamental theorem of culculus:

   $$\int_{-2}^{2}\left(t^2 - 4\right)dt$$

1.14 FURTHER READING

NOTES

Rudin, Walter. 2017. *Real and Complex Analysis*, Third Edition. Noida: McGraw-Hill Education.

Gupta, S. L. and Nisha Rani. 2004. *Fundamental Real Analysis*, Fourth Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, First Edition. Cambridge (U.K.): Cambridge University Press.

Bartle, Robert G. and Donald R. Sherbert. 2014. *Introduction to Real Analysis*, Fourth Edition. New York: Wiley.

Trench, William F. 2002. *Introduction to Real Analysis*, London: Pearson.

Loeb, Peter A. 2016. *Real Analysis*, Basel (Switzerland): Birkhäuser.

Royden, Halsey. 2015. *Real Analysis*, Fourth Edition. Noida: Pearson Education India.

UNIT 2 SEQUENCES AND SERIES OF FUNCTIONS

NOTES

Structure

2.0 INTRODUCTION

In real analysis, a sequence is an enumerated collection of objects in which repetitions are allowed and order matters. Like a set, it contains members, also called elements, or terms. The number of elements (possibly infinite) is called the length of the sequence. Unlike a set, the same elements can appear multiple times at different positions in a sequence, and unlike a set, the order does matter. Formally, a sequence can be defined as a function from natural numbers, the positions of elements in the sequence, to the elements at each position. The notion of a sequence can be generalized to an indexed family typically defined as a function from an index set that may not be numbers to another set of elements.

A sequence can be thought of as a list of elements with a particular order. Sequences are considered significant in a number of mathematical disciplines for studying functions, spaces, and other mathematical structures using the convergence properties of sequences. In particular, sequences are the basis for series, which are important in differential equations and analysis. Multiple sequences can be considered simultaneously using different variables. sometimes the elements of the sequence are naturally related to a sequence are naturally related to a sequence of integers whose pattern can be inferred easily.

NOTES

In real analysis, a series is, roughly speaking, a description of the operation of adding infinitely many quantities, one after the other, to a given starting quantity. The study of series is a key part of calculus and its generalization for mathematical analysis. Series are used in most areas of mathematics, even for studying finite structures (such as, in combinatorics) through generating functions. In addition to their ubiquity in mathematics, infinite series are also widely used in other quantitative disciplines, such as physics, computer science, statistics and finance.

As for sequences of functions, and unlike for series of numbers, there exist many types of convergence for a function series, such as uniform convergence, point wise convergence, almost everywhere convergence, etc. The Weierstrass M-test is a useful result in studying convergence of function series.

In this unit, you will study about the sequence and series of functions, point-wise and uniform convergence, Cauchy criterion for uniform convergence, Weierstrass's *M* test, Abel's and Dirichlet's tests for uniform convergence, uniform convergence and continuity, uniform convergence and Riemann-Stieltjes Integration, uniform convergence and differentiation, Weierstrass approximation theorem, power series, uniqueness theorem for power series, Abel's and Tauber's theorems.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Define sequence and series
- Understand pointwise and uniform convergence
- Explain Cauchy criterion for uniform convergence
- Analyse the Weierstrass's *M*-test
- Discuss about the Abel's test and Dirichlet's test for uniform convergence
- Explain uniform convergence in context with continuity
- Know the significance of Riemann-Stieltjes integration and differentiation
- State Weierstrass approximation theorem
- Define power series
- Discuss uniqueness theorem for power series
- Explain Abel's theorem and Tauber's theorem

2.2 SEQUENCE

A **sequence** is a function whose domain is the set of natural numbers. If the codomain is the set $\mathbb{R}$ of real numbers, it is called a real sequence; if it is the set $\mathbb{C}$ of complex numbers, it is called a complex sequence and likewise if it is a set of polynomials, it is a sequence of polynomials.

The image of the numbers 1, 2, 3, ... are called the first, second, third terms of the sequence, respectively.

NOTES

Thus a real sequence can be conceived as a collection of numbers so that to every natural number there is a unique member of that collection. If the natural number is n, the corresponding number is denoted by x_n or y_n or z_n or u_n etc., and is called the nth term of the sequence. The sequence is denoted by $\{x_n\}$.

Thus $x_n = \frac{1}{n}$ is a sequence whose 1st, 2nd, 3rd terms are respectively 1, $\frac{1}{2}, \frac{1}{3}$. This sequence is called the *harmonic sequence*.

Another example of a sequence is $y_n = (-1)^n$. The first few terms of the sequence are $\{-1, 1, -1, 1, ...\}$.

The sequence $Z_n = 5$ is also a sequence, each of its term being 5. Such a sequence is called a *constant sequence*.

Bounded and Unbounded Sequences

A sequence $\{x_n\}$ is said to be *bounded above* if all its terms are less than or equal to a real number, i.e., there exists $K \in \mathbb{R}$ such that $x_n \le K$ for all $n \in \mathbb{N}$.

As for example, the sequence $\left\{\frac{1}{n}\right\}$ is bounded above since $\frac{1}{n} \le 1$ for all $n \in \mathbb{N}$, the sequence $\left\{\frac{5n+1}{2n+2}\right\}$ is bounded above since $\frac{5n+1}{3n+2} \le 3$ for all n, but the sequence $\le \{n^2\}$ is not bounded above since there exists no such real number K so that $n^2 \le K$ for all n. In fact it is easy to observe that for every real number K there is an n such that $n^2 > K$. Such a sequence as above is called an unbounded sequence.

A sequence $\{x_n\}$ is said to be *bounded below* if all its terms are greater than or equal to a real number, i.e., there exists $K \in \mathbb{R}$ such that $x_n \ge k$ for all $n \in \mathbb{N}$. The sequence $\left\{\frac{1}{n}\right\}$ is bounded below since $\frac{1}{n} \ge 0$ for all n. The sequence $\left\{\frac{5n+1}{3n+2}\right\}$ is also bounded below since $\frac{5n+1}{3n+2} \ge 0$ for all n. The sequence $\{(-1)^n 5\}$ is bounded below since $(-1)^n 5 \ge -5$ for all $n \in \mathbb{N}$, but the sequence $\{(-2)^n\}$ is not bounded below since there is no such real number k for which $k \le (-2)^n$. Indeed, if K is a negative real number, there always exists, an (odd) integer n such that $(-2)^n < k$.

A sequence is said to be *bounded* if it is bounded both above and below, i.e., if there exist $K, k \in \mathbb{R}$ such that $k \le x_n \le K$ for all $n \in N$.

The numbers K and k are called, respectively, an upper bound and a lower bound of the sequence $\{x_n\}$. Note that if a sequence $\{x_n\}$ has an upper bound, it has many upper bounds; similarly if a sequence $\{x_n\}$ has a lower bound, it has many lower bounds. For example, for the sequence $\left\{\left(1+\frac{1}{n}\right)^n\right\}$, just as 3 is an upper bound, any real number greater than 3 is also an upper bound.

NOTES

Monotone Sequence

A sequence $\{x_n\}$ is said to be monotone increasing if $x_n \le x_{n+1}$ for every $n \in \mathbb{N}$; the sequence is called *strictly increasing* if $x_n < x_{n+1}$ for every $n \in \mathbb{N}$. Clearly the sequence $\{n^2\}$ is monotone (strictly) increasing since $n^2 \le (n+1)^2$ always. The sequence $\{(-2)^n\}$ is not monotone increasing since $(-2)^2 \not\le (-2)^3$.

A sequence $\{x_n\}$ is said to be *monotone decreasing* if $x_{n+1} \le x_n$ for every $n \in \mathbb{N}$; the sequence is called *strictly decreasing* if $x_{n+1} < x_n$ for every $n \in \mathbb{N}$.

The sequence $\left\{\frac{1}{n^2+1}\right\}$ is monotone (strictly) decreasing as $\frac{1}{(n+1)^2+1} \le \frac{1}{n^2+1}$ for every n. The sequence $\{-n^3\}$ is strictly decreasing as $-(n+1)^3 < -n^3$ but the sequence $\left(-\frac{1}{2}\right)^n$ is not monotone or strictly decreasing as $\left(-\frac{1}{2}\right)^4 \not< \left(-\frac{1}{2}\right)^3$.

Convergent Sequence

A very natural inquiry about a sequence $\{x_n\}$ is whether the terms x_n come close to any real number when n is very very large. This is what is known as the convergence of a sequence.

Definition: A sequence $\{x_n\}$ is said to *converge* to a real number l if for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that,

$$|x_n - l| < \varepsilon \quad \text{for every } n \ge n_0$$

The number l is called *limit* of the sequence $\{x_n\}$.

The fact that $\{x_n\}$ converges to l is expressed symbolically by $\lim_{n\to\infty} x_n = l$.

A sequence $\{x_n\}$ is called *convergent* if it converges to a limit l.

A sequence which converges to zero is called a *null sequence.*

The following facts follow readily from the definition:

Fact 1 : A sequence may or may not converge.

Fact 2 : If a sequence is convergent, it converges to a unique limit, i.e., it cannot converge to two different limits.

Fact 3 : Every convergent sequence is always bounded, but not conversely.

Proof: Let $\{x_n\}$ be a convergent sequence with limit l. Then for a given $\varepsilon\,(>0) = l$, say, there exists a positive integer n_0 such that,

$$|x_n - l| < l \qquad \text{for all} \quad n \ge n_0$$

i.e., $$l - 1 < x_n < l + 1 \qquad \text{for all} \quad n \ge n_0$$

Fact 4 : A monotone increasing sequence bounded above is always convergent and converges to its Least Upper Bound (LUB).

Fact 5 : A monotone decreasing sequence bounded below is always convergent and converges to its Greatest Lower Bound (GLB).

Fact 6 : Every constant sequence is convergent.

Let $\quad L = \min\{x_1, x_2, ..., x_{n_0}, |l| - 1\} \in \mathbb{R}$

And $\quad U = \max\{x_1, x_2, ..., x_{n_0}, |l| + 1\} \in \mathbb{R}$

Then $\quad L < x_n < U \quad$ for all n.

Hence, $\{x_n\}$ is a bounded sequence.

But the converse of this theorem is not true.

For example, the sequence $\{1 + (-1)^n\}$ is bounded but it does not converges to any finite limit. If the sequence is $\{0, 2, 0, 2, ...\}$ then its lower bound is 0 and upper bound is 2.

Cauchy's Criterion of Convergence

Since proof of convergence of a sequence requires determination of the limit, proving convergence is not always easy. Cauchy therefore provided an alternative way to prove convergence of a sequence, called Cauchy's criterion which avoids the determination of the limit. This may be stated as follows:

A sequence $\{x_n\}$ is convergent iff, for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$, usually depending on ε, such that

$$|x_m - x_n| < \varepsilon \quad \text{for all } m, n \geq n_0.$$

Or equivalently, $|x_{n+p} - x_n| < \varepsilon \quad$ for all $n \geq n_0$, $p = 0, 1, 2, 3, ...$

The sequence $\left\{\frac{1}{n}\right\}$ is convergent since,

$$\left|\frac{1}{n+p} - \frac{1}{n}\right| < \varepsilon$$

If $\quad \dfrac{p}{n(n+p)} < \varepsilon$

i.e., if $\frac{1}{n} < \varepsilon$, i.e., if $n > \frac{1}{\varepsilon}$, i.e., if $n \geq n_0 = \left[\frac{1}{\varepsilon}\right] + 1 \in \mathbb{N}$

Examine and prove that, $n \geq \left[\frac{1}{\varepsilon}\right] + 1 \Rightarrow n > \frac{1}{\varepsilon} \Rightarrow \frac{1}{n} < \varepsilon \Rightarrow \frac{p}{n(n+p)} < \varepsilon \Rightarrow$

$$\left|\frac{1}{n+p} - \frac{1}{n}\right| < \varepsilon$$

Example 2.1: Show that the sequence $\{x_n\}$ is convergent when,

$$x_n = 1 + \frac{1}{2!} + \frac{1}{3!} + ... + \frac{1}{n!}$$

Solution: Examine and prove that,

$$\frac{1}{n!} = \frac{1}{1.2.3.\cdots.n} < \frac{1}{2.2.\cdots.2} = \frac{1}{2^{n-1}}$$

For $m > n$

$$|x_m - x_n| = \left|\frac{1}{(n+1)!} + \frac{1}{(n+2)!} + ... + \frac{1}{m!}\right| < \frac{1}{2^n} + \frac{1}{2^{n+1}} + ... + \frac{1}{2^{m-1}}$$

$$= \frac{1}{2^n}\left(1 + \frac{1}{2} + ... + \frac{1}{2^{m-n}}\right)$$

$$= \frac{1}{2^n} \cdot \frac{1-\left(\frac{1}{2}\right)^{m+n-1}}{1-\frac{1}{2}} < \frac{1}{2^n} \cdot 2 = \frac{1}{2^{n-1}} < \varepsilon$$

NOTES

$$\to 0 \text{ as } n \to \infty$$

Hence $\{x_n\}$ is convergent.

Algebra of Limits

The followng result is of immense importance in evaluation of limits.

Theorem 2.1: If $\lim\limits_{n\to\infty} x_n = l$ and $\lim\limits_{n\to\infty} y_n = m$, then

(*i*) $\lim\limits_{n\to\infty} \{x_n + y_n\} = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n = l + m.$

(*ii*) $\lim\limits_{n\to\infty} \{x_n - y_n\} = \lim\limits_{n\to\infty} x_n - \lim\limits_{n\to\infty} y_n = l - m.$

(*iii*) $\lim\limits_{n\to\infty} \{x_n\, y_n\} = \lim\limits_{n\to\infty} x_n \lim\limits_{n\to\infty} y_n = l \,.\, m$

(*iv*) $\lim\limits_{n\to\infty} \dfrac{x_n}{y_n} = \dfrac{\lim\limits_{n\to\infty} x_n}{\lim\limits_{n\to\infty} y_n} = \dfrac{l}{m}$ if $m \neq 0$, provided the above limits exist.

Another result plays a dominant role in many situations. This is the so called *sandwich theorem* stated as follows:

Theorem 2.2: (*a*) If $x_n < y_n$ for all $n \in \mathbb{N}$, then $\lim\limits_{x\to\infty} x_n \leq \lim\limits_{x\to\infty} y_n$.

(*b*) If $x_n < y_n < z_n$ and $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} z_n = l$, then $\lim\limits_{n\to\infty} y_n = l$.

The proofs of the above theorems are outside the scope of this text.

Example 2.2: Show that the sequence $\left\{\dfrac{2n+3}{3n-2}\right\}$ is convergent.

Solution: Since $-4 < 9, \quad 6n - 4 < 6n + 9.$

Or $\quad 2(3n-2) < 3(2n+3)$

Or $\quad \dfrac{2n+3}{3n-2} > \dfrac{2}{3}$

Hence, the sequence $\left\{\dfrac{2n+3}{3n-2}\right\}$ is bounded below.

Further taking, $x_n = \dfrac{2n+3}{3n-2}$, we observe

$$x_n - x_{n+1} = \frac{2n+3}{3n-2} - \frac{2(n+1)+3}{3(n+1)-2}$$

$$= \frac{(2n+3)(3n+1) - (2n+5)(3n-2)}{(3n-2)(3n+1)}$$

$$= \frac{6n^2 + 11n + 3 - 6n^2 - 11n + 10}{(3n-2)(3n+1)} = \frac{13}{(3n-2)(3n+1)} \geq 0$$

for all n

i.e., $x_{n+1} \le x_n$ for all n.

Thus, $\left\{\dfrac{2n+3}{3n-2}\right\}$ being monotone decreasing and bounded below is convergent.

Divergent and Oscillatory Sequences

A sequence may be such that its terms become successively larger and larger, ultimately exceeding any big number. Such a sequence is said to diverge to $+\infty$. On the other hand, a sequence may have decreasing terms so that ultimately it becomes smaller than any negative but numerically large real number. Such a sequence is said to diverge to $-\infty$. Such sequences are also possible the terms of which do not approach any definite real number nor do exceed any large positive real number or recede any arbitrary negative number. These are nothing but oscillatory sequences. The formal definitions go as follows:

Definition: A sequence $\{x_n\}$ is said to *diverge* to $+\infty$ if for every large $G > 0$, there exists $n_0 \in \mathbb{N}$ such that,

$$x_n \ge G \quad \text{for all } n \ge n_0.$$

The fact $\{x_n\}$ diverges to ∞ is expressed symbolically by $\lim\limits_{n\to\infty} x_n = \infty$.

A sequence $\{x_n\}$ is said to diverge to $-\infty$ if for every large $G > 0$, there exists $n_0 \in \mathbb{N}$ such that,

$$x_n \le -G \quad \text{for all } n \ge n_0.$$

This is expressed symbolically by $\lim\limits_{n\to\infty} x_n = -\infty$.

A non-constant sequence which is bounded and not convergent is a finitely oscillatory sequence and a non-constant sequence which is unbounded and not convergent is an infinitely oscillatory sequence.

For example, the sequence $x_n = 5 - (-1)^n 2$ is a finitely oscillatory sequence but the sequence $y_n = (-2)^n$ is an infinitely oscillatory sequence.

Theorem 2.3: If $\{x_n\}$ be a sequence such that $\lim\limits_{n\to\infty}\left|\dfrac{x_{n+1}}{x_n}\right| = l$ where $0 \le l < 1$, then the sequence $\{x_n\}$ is a null sequence, i.e., $\lim\limits_{n\to\infty} x_n = 0$.

Proof: Beyond the scope of this book.

Example 2.3: Prove that $\lim\limits_{n\to\infty} \dfrac{x^n}{\lfloor n} = 0$ for every real value of x.

Solution: Here, $x_n = \dfrac{x^n}{\lfloor n}$ and $x_{n+1} = \dfrac{(x)^{n+1}}{\lfloor n+1}$

$$\because \qquad \left|\frac{x_{n+1}}{x_n}\right| = \left|\frac{x^{n+1}}{\lfloor n+1} \times \frac{\lfloor n}{x^n}\right| = \left|\frac{x}{n+1}\right| = \frac{|x|}{n+1} \to 0$$

As $n \to \infty$ for all real value of x.

$$\therefore \qquad \lim_{n\to\infty} \left|\frac{x_{n+1}}{x_n}\right| = 0$$

NOTES

Hence, $$\lim_{n\to\infty} \frac{x^n}{\lfloor n} = 0 .$$

Example 2.4: Prove that $\lim_{n\to\infty} \frac{x^n}{n} = 0$ if $|x| < 1$.

Solution: Here, $x_n = \frac{x^n}{n}$ and $x_{n+1} = \frac{x^{n+1}}{n+1}$

$$\because \qquad \left|\frac{x_{n+1}}{x_n}\right| = \left|\frac{x^{n+1}}{n+1} \times \frac{n}{x^n}\right| = \left|\frac{nx}{n+1}\right| = \frac{n}{n+1}|x|$$

$$= \frac{1}{1+\frac{1}{n}}|x| \to |x| \quad \text{as} \quad n \to \infty$$

$$\therefore \qquad \lim_{n\to\infty} \frac{x^n}{n} = 0 \quad \text{if} \quad |x| < 1.$$

When $x = 1$, the given sequence is a harmonic sequence which converges to zero as $n \to \infty$ and when $x = -1$, the given sequence is $\frac{(-1)^n}{n}$ which converges to zero as $n \to \infty$.

Hence, $$\lim_{n\to\infty} \frac{x^n}{n} = 0 \text{ for } |x| < 1.$$

2.3 SERIES

An expression of the form,

$$u_1 + u_2 + u_3 + ... + u_n + ...$$

in which every term is followed by another according to some definite rule is called a **series**. If it contains finite number of terms, then it is called a *finite series*. If the number of terms is not finite, it is called an *infinite series*. Such a series is conveniently denoted by,

$$\sum_{n=1}^{\infty} u_n \text{ or simply by } \sum u_n .$$

The sum of the first n terms of this series is denoted by S_n where $S_n = u_1 + u_2 + ... + u_n$) and is called the nth partial sum of the series. Now, we consider the following cases:

(*i*) If $S_n \to S$ (a finite value) as $n \to \infty$, then the series Σu_n is said to be convergent and S is called its sum.

(*ii*) If $S_n \to \pm\infty$ as $n \to \infty$, then the series Σu_n is called a divergent series.

(*iii*) If S_n oscillates (finitely or infinitely) as $n \to \infty$, the series Σu_n is said to be oscillatory.

A divergent or oscillatory series is called non-convergent.

NOTES

The infinite series Σu_n is said to converge to S, if for every arbitrary small positive number ε, there exists a positive integer m depending on ε such that,

$$| S_n - S | < \varepsilon \text{ for all } n \geq m.$$

(*i*) The series $\frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + ... + \frac{1}{n(n+1)} + ...$ is convergent.

$$\text{Since } S_n = \frac{1}{1.2} + \frac{1}{2.3} + \frac{1}{3.4} + ... + \frac{1}{n(n+1)}$$

$$= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + ... + \left(\frac{1}{n} - \frac{1}{n+1}\right)$$

$$= \left(1 - \frac{1}{n+1}\right) \to 1 \text{ as } n \to \infty$$

Hence $S = 1$.

(*ii*) The series $1 + 2 + 3 + 4 + ... + n + ...$ diverges to $+\infty$.

(*iii*) The series $1 - 1 + 1 - 1 + 1 - 1 + ...$, oscillates finitely and the series $1 - 2 + 3 - 4 + ...$ oscillates infinitely.

Note that the nature of a series is determined by the nature of the sequence of its nth partial sum.

Two Important Series

1. **The Geometric Series:** The infinite geometric series,

$$a + ax + ax^2 + ax^3 + ... + ax^{n-1} + ... \ (a > 0) \text{ is}$$

(*i*) Convergent if the common ratio x lies between -1 and 1 (i.e., $-1 < x < 1$), and the sum of the series is $\frac{a}{1-x}$.

(*ii*) Divergent (to $+\infty$) if $x \geq 1$.

(*iii*) Oscillates finitely if $x = -1$ and oscillates infinitely if $x < -1$.

Proof: The nth partial sum of the given series is,

$$S_n = a + ax + ax^2 + ... + ax^{n-1}$$

$$= a\frac{x^n - 1}{x - 1} \ (x \neq 1)$$

(*i*) If $| x | < 1$, then $x^n \to 0$ as $n \to \infty$

$$\therefore \ S_n \to \frac{a}{(1-x)} \text{ as } n \to \infty.$$

(*ii*) If $x > 1$, then $x^n \to \infty$ as $n \to \infty$ and then $S_n \to \infty$ as $n \to \infty$. Hence, the geometric series diverges to $+\infty$ if $x > 1$.

When $x = 1$, the series becomes $a + a + a + a + ...$

And $S_n = a + a + ... + a = na \to \infty$ as $n \to \infty$.

$\therefore$ The series diverges to $+\infty$.

NOTES

(*iii*) If $x < -1$, x^n oscillates infinitely between $-\infty$ and $+\infty$.

If $x = -1$, then the series becomes $a - a + a - a + a - a$ ...

And $S_n = \begin{cases} a \text{ if } n \text{ is odd} \\ o \text{ if } n \text{ is even} \end{cases}$

$\therefore$ The series oscillates finitely.

2. **The *p*-Series:** The infinite series,

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + ... + \frac{1}{n^p} + ...$$

(*i*) Converges if $p > 1$.

(*ii*) Diverges if $p \leq 1$.

Proof: (*i*) $p > 1$. We consider the partial sum of order $2^n - 1$ where n is a positive integer.

$$S_{2^n-1} = 1 + \left(\frac{1}{2^p} + \frac{1}{3^p}\right) + \left(\frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p}\right) + \left(\frac{1}{8^p} + \frac{1}{9^p} + ... + \frac{1}{15^p}\right)$$

$$+ ... + \left\{\frac{1}{(2^{n-1})^p} + \frac{1}{(2^{n-1}+1)^p} + ... + \frac{1}{(2^n-1)^p}\right\} \leq 1$$

$$+ \left(\frac{1}{2^p} + \frac{1}{2^p}\right) + \left(\frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p} + \frac{1}{4^p}\right) + \left(\frac{1}{8^p} + \frac{1}{8^p} + ... + \frac{1}{8^p}\right)$$

$$+ ... + \left\{\frac{1}{(2^{n-1})^p} + \frac{1}{(2^{n-1})^p} + ... + \frac{1}{(2^{n-1})^p}\right\}$$

$$= 1 + \frac{2}{2^p} + 4.\frac{1}{4^p} + 8.\frac{1}{8^p} + ... + 2^{n-1}\frac{1}{(2^{n-1})^p}$$

$$= 1 + \frac{1}{2^{p-1}} + \frac{1}{(2^{p-1})^2} + \frac{1}{(2^{p-1})^3} + ... + \frac{1}{(2^{p-1})^{n-1}}$$

$$= \frac{1 - \left(\frac{1}{2^{p-1}}\right)^n}{1 - \frac{1}{2^{p-1}}} < \frac{1}{1 - \frac{1}{2^{p-1}}} < k \text{ (Constant) for all } n \quad \text{(say)}$$

Now for any positive integer m, there exists a positive integer n such that $2^n - 1 > m$.

$\therefore$ $\{S_m\}$ is clearly monotonically increasing ($\because$ All Terms are Positive)

$\therefore$ $S_m < S_{2^n-1} < k \quad \forall\ m$

$\therefore$ $\{S_m\}$ is monotonically increasing and bounded above.

Hence the series $\sum_{m=1}^{\infty} \frac{1}{m^p}$ is convergent.

(*ii*) $p \leq 1$. Here, we prove that $\sum_{m=1}^{\infty} \frac{1}{m^p}$ diverges.

Now when $p \leq 1$, $n^p \leq n$ where n is positive integer.

We consider the partial sum of order 2^n.

$$\therefore \quad S_{2^n} = 1+\frac{1}{2^p}+\left(\frac{1}{3^p}+\frac{1}{4^p}\right)+\left(\frac{1}{5^p}+\frac{1}{6^p}+\frac{1}{7^p}+\frac{1}{8^p}\right)+\ldots$$

$$+\left\{\frac{1}{(2^{n-1}+1)^p}+\frac{1}{(2^{n-1}+2)^p}+\ldots+\frac{1}{(2^n)^p}\right\} \geq \left(1+\frac{1}{2}\right)$$

$$+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)+\ldots+$$

$$\left(\frac{1}{2^{n-1}+1}+\frac{1}{2^{n-1}+2}+\ldots+\frac{1}{2^n}\right) \quad \left(\because \frac{1}{n^p} \geq \frac{1}{n}\right)$$

$$\geq \left(1+\frac{1}{2}\right)+\left(\frac{1}{4}+\frac{1}{4}\right)+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)+\ldots$$

$$+\left(\frac{1}{2^n}+\frac{1}{2^n}+\ldots+\frac{1}{2^n}\right)$$

$$= 1+\frac{1}{2}+2.\frac{1}{4}+4.\frac{1}{8}+\ldots+2^{n-1}.\frac{1}{2^n}$$

$$= 1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\ldots+\frac{1}{2} = 1+\frac{n}{2}$$

$$\therefore \quad S_{2^n} \geq 1+\frac{n}{2}$$

$\therefore$ Thus, for an arbitrary $G > 0$, $S_{2^n} > G$ whenever $1+\frac{n}{2} > G$

i.e., $n > 2G - 2$.

Thus, the partial sums are monotonically increasing.

If $m > 2^n$, then $S_m > S_{2^n} > G$ for all $m > 2^{2G-2}$

$\therefore$ The sequence of partial sums $\{S_m\}$ is monotonically increasing and unbounded above and hence converges to $+\infty$ as $n \to \infty$.

Hence, the series $\sum \frac{1}{m^p}$ diverges to $+\infty$ when $p \leq 1$.

Cauchy's General Principle of Convergence

Statement: A necessary and sufficient condition for the convergence of an infinite series $\sum_{n=1}^{\infty} u_n$ is that for every positive number ε, however small, there exists a positive integer n_0, which depends on ε, such that,

$$| u_{n+1} + u_{n+2} + \ldots + u_m | < \varepsilon \quad \text{for all } m \geq n \geq n_0$$

Note: If $\sum_{n=1}^{\infty} u_n$ converges then $\lim_{n\to\infty} u_n = 0$.

Applications of Cauchy's Principle

Example 2.5: Prove, by using Cauchy's criterion that the series,

$$\sum_{n=1}^{\infty}\frac{1}{n!} = 1+\frac{1}{\lfloor 1}+\frac{1}{\lfloor 2}+\frac{1}{\lfloor 3}+\frac{1}{\lfloor 4}+\ldots+\frac{1}{\lfloor n}+\ldots \text{ converges.}$$

Solution: Here $\left|\sum_{k=n+1}^{n+p} \frac{1}{\lfloor k}\right| = \left|\frac{1}{\lfloor n+1} + \frac{1}{\lfloor n+2} + \frac{1}{\lfloor n+3} + ... + \frac{1}{\lfloor n+p}\right|$ for $p \geq 1$.

NOTES

$$< \frac{1}{2^n} + \frac{1}{2^{n+1}} + ... + \frac{1}{2^{n+p-1}}$$

$$\left(\because \frac{1}{(n+1)\,(n)\,(n-1)...4.3.2.1} < \frac{1}{2.2...:2.2} = \frac{1}{2^n}\right)$$

$$= \frac{1}{2^n}\left(1 + \frac{1}{2} + \frac{1}{2^2} + ... + \frac{1}{2^{p-1}}\right)$$

$$= \frac{1}{2^n} \frac{1-\left(\frac{1}{2}\right)^p}{1-\frac{1}{2}} < \frac{1}{2^n\left(1-\frac{1}{2}\right)} \text{ for all } p \geq 1$$

$$\therefore \quad \left|\sum_{k=n+1}^{n+p} \frac{1}{\lfloor k}\right| < 2 \,.\, \frac{1}{2^n} \quad \text{for all } p \geq 1.$$

$$\because \quad \left(\frac{1}{2}\right)^n \to 0 \text{ as } n \to \infty,$$

Then for each $\varepsilon > 0$, there exists a positive integer n_0 which depends on ε for which $2\left(\frac{1}{2}\right)^n < \varepsilon$ if $n > n_0$.

$$\therefore \quad \left|\sum_{k=n+1}^{n+p} \frac{1}{\lfloor k}\right| < \varepsilon \text{ for } n > n_0 \text{ and } p \geq 1$$

Hence, by the Cauchy's principle, the given series is convergent.

Tests of Convergence and Divergence

Result (1): If Σu_n be a convergent series of positive terms, then it necessarily follows that $\lim\limits_{n\to\infty} u_n = 0$.

Result (2): If Σu_n is a convergent series of positive and decreasing terms, then it necessarily follows that $\lim\limits_{n\to\infty} nu_n = 0$.

1. Comparison Test

Let Σu_n and Σv_n be two infinite series of positive terms. If $\lim\limits_{n\to\infty} \frac{u_n}{v_n} = k$, a non zero finite quantity, then the series are both convergent or both divergent. If $\lim\limits_{n\to\infty} \frac{u_n}{v_n}$ = non zero finite quantity, then if Σv_n is convergent, Σu_n is convergent and if Σv_n is divergent, Σu_n is divergent.

Example 2.6: Test the convergence of the series whose nth term is $\sqrt{n^2+1} - n$.

Solution: Here, $u_n = \sqrt{n^2+1} - n$ and we take $v_n = \frac{1}{n}$.

$$\therefore \quad \frac{u_n}{v_n} = n\left(\sqrt{n^2+1}-n\right) = \frac{n\left(n^2+1-n^2\right)}{\sqrt{n^2+1}+n} = \frac{1}{\left(\sqrt{1+\frac{1}{n^2}}\right)+1}$$

$$\to \frac{1}{2} \text{ as } n \to \infty$$

NOTES

Since, $\lim\limits_{n\to\infty} \frac{u_n}{v_n} = \frac{1}{2}$ and $\Sigma v_n = \sum \frac{1}{n}$ is divergent, thus by the comparison test Σu_n is divergent.

2. Cauchy's Root Test

Let Σu_n be a series of positive terms and,

$$\lim_{n\to\infty} (u_n)^{\frac{1}{n}} = l$$

(*i*) If $l < 1$, then Σu_n is convergent.

(*ii*) If $l > 1$, then Σu_n is divergent.

(*iii*) If $l = 1$, then the test fails.

Example 2.7: Test the convergence of the series,

$$\frac{1}{3} + \left(\frac{2}{5}\right)^2 + \left(\frac{3}{7}\right)^3 + \ldots + \left(\frac{n}{2n+1}\right)^n + \ldots$$

Solution: Here, $u_n = \left(\frac{n}{2n+1}\right)^n$

$$\therefore \quad (u_n)^{\frac{1}{n}} = \frac{n}{2n+1}$$

$$\therefore \quad \lim_{n\to\infty} (u_n)^{\frac{1}{n}} = \lim_{n\to\infty} \frac{n}{2n+1} = \lim_{n\to\infty} \frac{1}{2+\frac{1}{n}} = \frac{1}{2} < 1$$

Hence, by the Cauchy's root test the given series is convergent.

3. D'Alembert's Ratio Test

Let Σu_n be a series of positive terms and $\lim\limits_{n\to\infty} \frac{u_{n+1}}{u_n} = l$

(*i*) If $l < 1$, then Σu_n is convergent.

(*ii*) If $l > 1$, then Σu_n is divergent.

(*iii*) If $l = 1$, then the test fails.

Example 2.8: Prove that the series $x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots (x \geq 0)$ is convergent if $0 \leq x < 1$ and divergent if $x > 1$.

Solution: Here, $u_n = \frac{x^n}{n}$ and $u_{n+1} = \frac{x^{n+1}}{n+1}$

$$\therefore \quad \frac{u_{n+1}}{u_n} = \frac{x^{n+1}}{n+1} \times \frac{n}{x^n} = \frac{x}{\left(1+\frac{1}{n}\right)} \to x \text{ as } n \to \infty$$

If $0 < x < 1$, then $\lim\limits_{n\to\infty} \dfrac{u_{n+1}}{u_n} < 1$, and the series is convergent.

NOTES

If $x > 1$, then $\lim\limits_{n\to\infty} \dfrac{u_{n+1}}{u_n} > 1$, and the series is divergent.

And if $x = 1$, then the D'Alembert's ratio test does not give any definite conclusion.

But when $x = 1$, the given series becomes a divergent harmonic series,

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n} + \ldots$$

$\therefore$ The series is convergent if $0 \le x < 1$ and divergent if $x \ge 1$.

4. Raabe's Test

Let Σu_n be a series of positive terms and,

$$\lim_{n\to\infty}\left\{n\left(\frac{u_n}{u_{n+1}} - 1\right)\right\} = l \text{ or } \lim_{n\to\infty} R_n = l \text{ where } R_n = n\left(\frac{u_n}{u_{n+1}} - 1\right)$$

If $l < 1$, then Σu_n is divergent.

If $l > 1$, then Σu_n is convergent.

If $l = 1$, then the test fails.

Remember that when D'Alembert's ratio test fails, we normally apply Raabe's test.

Example 2.9: Prove that the series,

$$1 + \frac{1}{2}\cdot\frac{1}{3} + \frac{1.3}{2.4}\cdot\frac{1}{5} + \frac{1.3.5}{2.4.6}\cdot\frac{1}{7} + \ldots \text{ to } \infty, \text{ converges.}$$

Solution: Here, $u_n = \dfrac{1.3.5\ldots(2n-1)}{2.4.6\ldots 2n} \; \dfrac{1}{2n+1}$

$$u_{n+1} = \frac{1.3.5\ldots(2n-1)\,(2n+1)}{2.4.6\ldots(2n)\,(2n+2)} \; \frac{1}{2n+3}$$

Now, $\dfrac{u_{n+1}}{u_n} = \dfrac{(2n+1)\,(2n+1)}{2(n+1)\,(2n+3)} = \dfrac{\left(1+\dfrac{1}{2n}\right)\left(1+\dfrac{1}{2n}\right)}{\left(1+\dfrac{1}{n}\right)\left(1+\dfrac{3}{2n}\right)} \to 1 \text{ as } n \to \infty.$

Hence, D'Alembert's ratio test fails. So we consider the Raabe's sequence $\{R_n\}$, where

$$R_n = n\left(\frac{u_n}{u_{n+1}} - 1\right) = n\left\{\frac{2(n+1)\,(2n+3)}{(2n+1)\,(2n+1)} - 1\right\}$$

$$= n\left(\frac{4n^2 + 10n + 6 - 4n^2 - 4n - 1}{(2n+1)^2}\right) = n\;\frac{6n+5}{(2n+1)^2}$$

$$= \frac{6 + \dfrac{5}{n}}{\left(2 + \dfrac{1}{n}\right)^2}$$

$$\therefore \lim_{n\to\infty} R_n = \lim_{n\to\infty} \frac{6+\frac{5}{n}}{\left(2+\frac{1}{n}\right)^2} = \frac{6}{4} = \frac{3}{2} > 1$$

Hence, by the Raabe's test the given series is convergent.

5. Logarithmic Test

Let Σu_n be a series of positive terms and,

$$\lim_{n\to\infty} \left\{ n \log\left(\frac{u_n}{u_{n+1}}\right)\right\} = l.$$

If $l < 1$, then Σu_n is divergent.

If $l > 1$, then Σu_n is convergent.

If $l = 1$, then the test fails.

Example 2.10: Test the convergence of the series,

$$x + \frac{2^2\, x^2}{\lfloor 2} + \frac{3^3\, x^3}{\lfloor 3} + \frac{4^4\, x^4}{\lfloor 4} + \ldots$$

Solution: Here, $u_n = \dfrac{n^n\, x^n}{\lfloor n}$ and $u_{n+1} = \dfrac{(n+1)^{n+1}\, x^{n+1}}{\lfloor n+1}$

$$\therefore \frac{u_{n+1}}{u_n} = \frac{(n+1)^{n+1}\, x^{n+1}}{\lfloor n+1} \times \frac{\lfloor n}{n^n\, x^n} = \frac{(n+1)^{n+1}}{(n+1)\lfloor n} \times \frac{\lfloor n}{n^n}\, x$$

$$= \left(1+\frac{1}{n}\right)^n x$$

$$\therefore \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n x = ex$$

Hence, by the D'Alembert's ratio test, the series is convergent if $ex < 1$, i.e., $x < \dfrac{1}{e}$ and divergent if $ex > 1$, i.e., $x > \dfrac{1}{e}$ and the test fails when $x = \dfrac{1}{e}$. We apply the logarithmic test for $x = \dfrac{1}{e}$.

$$\therefore \log\left(\frac{u_n}{u_{n+1}}\right) = \log e - \log\left(1+\frac{1}{n}\right)^n$$

$$= \log e - n\log\left(1+\frac{1}{n}\right)$$

$$= 1 - n\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \ldots\right)$$

$$= \left(\frac{1}{2n} - \frac{1}{3n^2} + \ldots\right)$$

$$\therefore n\log\left(\frac{u_n}{u_{n+1}}\right) = \frac{1}{2} - \frac{1}{3n} + \ldots$$

$$\therefore \lim_{n\to\infty} n \log\left(\frac{u_n}{u_{n+1}}\right) = \frac{1}{2} < 1$$

NOTES

Thus, the series is divergent when $x = \frac{1}{e}$.

Hence, the series is convergent if $x < \frac{1}{e}$ and divergent if $x \geq \frac{1}{e}$.

6. Gauss' Test

Let Σu_n be a series of positive terms and we can express $\frac{u_n}{u_{n+1}}$ in the form $1 + \frac{a}{n} + \frac{\beta_n}{n^p}\ (p > 1)$,

$$\text{i.e., } \frac{u_n}{u_{n+1}} = 1 + \frac{a}{n} + \frac{\beta_n}{n^p}\ (p > 1)$$

Where $\{\beta_n)$ is a bounded sequence. It will be sufficient if $\{\beta_n\}$ is a convergent sequence because it will be necessarily bounded. Then,

(*i*) Σu_n converges if $a > 1$.

(*ii*) Σu_n diverges if $a \leq 1$.

Note: When D'Alembert's ratio test fails, one may try Gauss' test, without going through other tests.

Example 2.11: Test the convergence of the series $\sum\left\{\frac{2\,.\,4\,.\,6\,.\,8\ldots 2n}{3\,.\,5\,.\,7\,.\,9\ldots(2n+1)}\right\}^2$.

Solution: Here, $u_n = \left\{\frac{2\,.\,4\,.\,6\,.\,8\ldots 2n}{3\,.\,5\,.\,7\,.\,9\ldots(2n+1)}\right\}^2$

And $$u_{n+1} = \left\{\frac{2\,.\,4\,.\,6\,.\,8\ldots 2n\,(2n+2)}{3\,.\,5\,.\,7\,.\,9\ldots(2n+1)\,(2n+3)}\right\}^2$$

$$\therefore \frac{u_n}{u_{n+1}} = \left(\frac{2n+3}{2n+2}\right)^2 = \left(1 + \frac{1}{2n+2}\right)^2$$

$$= 1 + \frac{1}{n} + \frac{\beta_n}{n^2} \text{ where } \beta_n = n^2\left[\frac{1}{(2n+2)^2} + \frac{2}{2n+2} - \frac{1}{n}\right]$$

$$\text{Now, } \beta_n = n^2\left[\frac{1}{4(n+1)^2} + \frac{1}{n+1} - \frac{1}{n}\right]$$

$$= n^2\left[\frac{1}{4(n+1)^2} - \frac{1}{n(n+1)}\right]$$

$$= n^2\left[\frac{n - 4(n+1)}{4(n+1)^2\, n}\right] = n^2\left[\frac{-3n-4}{4(n+1)^2\, n}\right]$$

$$= \frac{-3 - \frac{4}{n}}{4\left(1 + \frac{1}{n}\right)^2} \to -\frac{3}{4} \text{ as } n \to \infty$$

$\therefore$ Hence, $\{\beta_n\}$ converges and hence is bounded.

$\therefore$ By the Gauss' test, the given series is divergent because $a = 1$.

7. De Morgan's and Bertrand's Test

Let Σu_n be a series of positive terms and $\lim\limits_{n\to\infty} \beta_n = l,$ where $\beta_n = (R_n - 1)$ $\log n$ and $R_n = \dfrac{u_n}{u_{n+1}} - 1$.

If $l < 1$, then Σu_n is divergent.

If $l > 1$, then Σu_n is convergent.

If $l = 1$, then the test fails.

Example 2.12: Test the convergence of the series

$$\frac{1^2}{2^2} + \frac{1^2 . 3^2 x}{2^2 . 4^2} + \frac{1^2 . 3^2 . 5^2}{2^2 . 4^2 . 6^2} x^2 + ... + \frac{1^2 . 3^2 . 5^2 ... (2n-1)^2}{2^2 . 4^2 . 6^2 ... (2n)^2} x^{n-1} + ... \text{ to } \infty .$$

Solution: Here, $u_n = \dfrac{1^2 . 3^2 . 5^2 ... (2n-1)^2}{2^2 . 4^2 . 6^2 ... (2n)^2} x^{n-1}$

And $u_{n+1} = \dfrac{1^2 . 3^2 . 5^2 ... (2n-1)^2 \ (2n+1)^2}{2^2 . 4^2 . 6^2 ... (2n)^2 \ (2n+2)^2} x^n$

$$\therefore \quad \frac{u_{n+1}}{u_n} = \frac{(2n+1)^2}{(2n+2)^2} x = \frac{\left(1+\dfrac{1}{2n}\right)^2}{\left(1+\dfrac{2}{2n}\right)^2} x \to x \text{ as } n \to \infty$$

Hence, by the D'Alembert's ratio test, Σu_n converges if $0 < x < 1$ and diverges if $x > 1$. If $x = 1$, then the test fails.

When $x = 1$, we consider Raabe's sequence $\{R_n\}$ where $R_n = \left(\dfrac{u_n}{u_{n+1}} - 1\right) n$.

$$\therefore \quad R_n = n\left(\frac{(2n+2)^2}{(2n+1)^2} - 1\right) = n\left\{\frac{4n^2 + 8n + 4 - 4n^2 - 4n - 1}{(2n+1)^2}\right\}$$

$$= \frac{n(4n+3)}{(2n+1)^2} = \frac{4\left(1+\dfrac{3}{4n}\right)n^2}{4n^2\left(1+\dfrac{1}{2n}\right)^2} = \frac{\left(1+\dfrac{3}{4n}\right)}{\left(1+\dfrac{1}{2n}\right)^2} \to 1 \text{ as } n \to \infty$$

$\therefore$ Raabe's test also does not gives any conclusion if $x = 1$.

We next try, D'Morgan's and Bertrand's test.

Now, $B_n = \log n \, (R_n - 1)$

$$= \left(\frac{4n^2 + 3n}{(2n+1)^2} - 1\right) \log n$$

$$= \frac{4n^2 + 3n - 4n^2 - 4n - 1}{(2n+1)^2} \log n$$

$$= \frac{-n-1}{(2n+1)^2}\log n = -\frac{\log n}{n}\frac{\left(1+\frac{1}{n}\right)}{\left(2+\frac{1}{n}\right)^2}$$

NOTES

$$\to 0 \text{ as } n \to \infty \quad \left(\because \lim_{n\to\infty}\frac{\log n}{n} = 0\right)$$

Since $B_n \to 0$ as $n \to \infty$ for $x = 1$, hence Σu_n is divergent for $x = 1$

$(\because B_n = 0 < 1)$.

$\therefore$ The given series is convergent for $0 < x < 1$ and divergent for $x \geq 1$.

8. Alternative Bertrand's Test

Let Σu_n be a series of positive terms and $\lim_{n\to\infty}\left\{n \log \frac{u_n}{u_{n+1}} - 1\right\} \log n = l$, then the given series Σu_n is convergent if $l > 1$ and Σu_n is divergent if $l < 1$.

Example 2.13: Test the convergence or divergence of the series,

$$1^p + \left(\frac{1}{2}\right)^p + \left(\frac{1.3}{2.4}\right)^p + \left(\frac{1.3.5}{2.4.6}\right)^p + \ldots$$

Solution: Here, $u_n = \left\{\frac{1.3.5\ldots(2n-1)}{2.4.6\ldots 2n}\right\}^p$

$$u_{n+1} = \left\{\frac{1.3.5\ldots(2n-1)(2n+1)}{2.4.6\ldots(2n)(2n+2)}\right\}^p$$

$$\therefore \quad \frac{u_{n+1}}{u_n} = \left(\frac{2n+1}{2n+2}\right)^p = \left\{\frac{\left(1+\frac{1}{2n}\right)}{\left(1+\frac{1}{n}\right)}\right\}^p \to 1 \text{ as } n \to \infty$$

So D'Alembert's ratio test does not give any conclusion.

$$\text{Now } n \log \frac{u_n}{u_{n+1}} = n \log\left\{\left(\frac{2n+2}{2n+1}\right)^p\right\}$$

$$= n \log\left\{\frac{\left(1+\frac{1}{n}\right)}{\left(1+\frac{1}{2n}\right)}\right\}^p$$

$$= pn\left[\log\left(1+\frac{1}{n}\right) - \log\left(1+\frac{1}{2n}\right)\right]$$

$$= np\left[\left(\frac{1}{n} - \frac{1}{2}\cdot\frac{1}{n^2} + \frac{1}{3}\cdot\frac{1}{n^3} - \ldots\right)\right.$$
$$\left. - \left(\frac{1}{2n} - \frac{1}{2}\cdot\frac{1}{4n^2} + \frac{1}{3}\cdot\frac{1}{8n^3} - \ldots\right)\right]$$

$$= np\left[\frac{1}{2n} - \frac{3}{8n^2} + \frac{7}{24n^3} - \ldots\right]$$

$$= p\left[\frac{1}{2}-\frac{3}{8n}+\frac{7}{24n^2}-...\right]$$

$$\therefore \lim_{n\to\infty} n\log\frac{u_n}{u_{n+1}} = \frac{p}{2}.$$

Hence, by the logarithmic test, the series Σu_n converges if $\frac{p}{2} > 1$ and diverges if $\frac{p}{2} < 1$.

When $\frac{p}{2} = 1$, i.e., when $p = 2$, the logarithmic test does not give any conclusion.

Now we try the alternative Bertrand's test:

$$\left(n\log\frac{u_n}{u_{n+1}}-1\right)\log n$$

$$= \left[2\left(\frac{1}{2}-\frac{3}{8n}+\frac{7}{24n^2}-...\right)-1\right]\log n$$

$$= \left[\left(1-\frac{3}{4n}+\frac{7}{12n^2}-...\right)-1\right]\log n$$

$$= n\left[-\frac{3}{4n}+\frac{7}{12n^2}-...\right]\frac{\log n}{n}$$

$$= n\left[-\frac{3}{4}+\frac{7}{12n}-...\right]\frac{\log n}{n}$$

$$\to 0\ \infty \text{ as } n\to\infty \quad \left(\because \lim_{n\to\infty}\frac{\log n}{n}=0\right)$$

$$\therefore \lim_{n\to\infty}\left(n\log\frac{u_n}{u_{n+1}}-1\right)\log n = 0 < 1.$$

Hence, by the alternative form of Bertrand's test, the series Σu_n is divergent for $p = 2$.

$\therefore$ The given series is convergent if $p > 2$ and divergent if $p \leq 2$.

Alternating Series: A series in which the terms are alternately positive and negative or negative and positive is called an alternating series.

Thus, $\sum_{n=1}^{\infty}(-1)^{n-1}u_n = u_1 - u_2 + u_3 - u_4 + ...$ is an alternating series.

if $u_n > 0\ \forall\ n$.

Or if $u_n < 0\ \forall\ n$.

Leibnitz's Test

Let the alternating series be,

$$\sum_{n=1}^{\infty}(-1)^{n-1}u_n = u_1 - u_2 + u_3 - u_4 + ...$$

This series converges if

(*i*) The sequence $\{u_n\}$ is Monotone Decreasing (M.D.).

(*ii*) $u_n \to 0$ as $n \to \infty$, i.e., $\lim_{n\to\infty} u_n = 0$.

NOTES

Example 2.14: Test the convergence of the series,

(*i*) $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - ...$ (*ii*) $\frac{1}{1^p} - \frac{1}{2^p} + \frac{1}{3^p} - \frac{1}{4^p} + ...$, for $p > 0$

Solution: (*i*) The terms of the given series are alternately positive and negative, so the given series is alternating series.

$$\text{Here,} \quad u_n = \frac{1}{n} \text{ and } u_{n+1} = \frac{1}{n+1}$$

$$\text{Now,} \quad u_{n+1} - u_n = \frac{1}{n+1} - \frac{1}{n} = \frac{n-n-1}{n(n+1)} = -\frac{1}{n(n+1)} < 0$$

$$\therefore \ u_{n+1} < u_n$$

Hence, the sequence $\{u_n\}$ is M.D.

$$\text{And} \quad \lim_{n\to\infty} u_n = \lim_{n\to\infty} \frac{1}{n} = 0.$$

$\therefore$ By the Leibnitz's test the given alternating series is convergent.

(*ii*) The terms of the given series are alternately positive and negative. So the given series is an alternating series.

$$\text{Here, } u_n = \frac{1}{n^p} \text{ and } u_{n+1} = \frac{1}{(n+1)^p}$$

$$\therefore \ u_{n+1} - u_n = \frac{1}{(n+1)^p} - \frac{1}{n^p} = \frac{n^p - (1+n)^p}{\{n(n+1)\}^p}$$

$$= \frac{n^p - \left[1 + pn + \frac{p(p-1)}{\lfloor n} n^2 + ... + n^p\right]}{\{n(n+1)\}^p}$$

$$= \frac{-\left(1 + pn + \frac{p(p-1)}{\lfloor 2} n^2 + ... + pn^{p-1}\right)}{\{n(n+1)\}^p} \le 0 \text{ for } p > 0$$

$$\therefore \ u_{n+1} \le u_n.$$

So the sequence $\{u_n\}$ is monotone decreasing and $\lim_{n\to\infty} \frac{1}{n^p} = 0$ for $p > 0$.

Hence, by the Leibnitz's test, the given series is convergent.

Theorem 2.4: Every absolutely convergent series is convergent.

Proof: Let the series Σu_n be absolutely convergent. Then $\Sigma \mid u_n \mid$ is convergent.

$\therefore$ By the Cauchy's general principle of convergence, for any given $\varepsilon > 0$, there exists a positive integer n_0 such that

$$|\,|u_{n+1}| + |u_{n+2}| + ... + |u_{n+p}|\,| < \varepsilon \text{ for all } n \ge n_0 \text{ and for all } p = 1, 2, 3 ...$$

$$\text{i.e.,} \quad |u_{n+1}| + |u_{n+2}| + ... + |u_{n+p}| < \varepsilon \text{ for all } n \ge n_0 \text{ and } p \ge 1$$

Now, for $n \geq n_0$ and all $p = 1, 2, 3...$,

$|u_{n+1} + u_{n+2} + ... + u_{n+p}| \leq |u_{n+1}| + |u_{n+2}| + ... + |u_{n+p}| < \varepsilon$

Hence, by the Cauchy's general principle of convergence, the series Σu_n is convergent.

2.4 POINTWISE AND UNIFORM CONVERGENCE

Definition: Suppose $\{f_n\}$, $n = 1, 2, 3, ...$ is a sequence of functions defined on a set S and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in S$. We can then define a function f by $f(x) = \lim_{n\to\infty} f_n(x), x \in S$.

Under these circumstances we say that $\{f_n\}$ converges to f pointwise on S and that f is the limit or the limit function of $\{f_n\}$.

Note: A sequence $\{f_n\}$ of functions is said to converge pointwise on a set S to a limit function f, if for each $x \in S$ and for each $\varepsilon > 0$ there exists an **N** (depending on x and ε) such that,

$$|f_n(x) - f(x)| < \varepsilon, \text{ for all } n > N.$$

Uniform Convergence

A sequence of real valued functions $\langle f_n \rangle$ defined on a set S is said to **converge uniformly** to a real valued function f on S if for $\varepsilon > 0$ $\exists$ $m \in \mathbf{N}$ such that,

$$|f_n(x) - f(x)| < \varepsilon \ \forall \ n \geq m \text{ and } x \in S$$

If $\langle f_n \rangle$ is not uniformly convergent on S but it is convergent for each x $\in S$ then $\langle f_n \rangle$ is said to be **pointwise convergent** on S. Evidently, uniform convergence of $\langle f_n \rangle$ to f on a set S implies that $\langle f_n \rangle$ also converges pointwise to f on S. But the converse is not true.

Note that if $f_n(x)$ converges pointwise to $f(x)$ on a set S then, if $f_n(x)$ is convergent uniformly on S, it converges uniformly to $f(x)$ on S.

Also note that if on S, $|f_n(x) - f(x)| < M_n$, where M_n is independent of x and if $M_n \to 0$, then $f_n(x) \to f(x)$ uniformly on S.

In the sense of usual convergence the natural number m involved in the definition depends upon the number ε and on x. In uniform convergence m essentially depends on ε only. Keeping in view this fact, various developments regarding convergence provide theory for uniform convergence.

Note that the pointwise convergence is a local property whereas the uniform convergence is a global property.

Example 2.15: The sqeuence $\langle x^n \rangle$ converges uniformly to 0 on $[0, a]$, where $0 < a < 1$, but not on $[0, 1)$.

Solution: For $\varepsilon > 0$, let $m > \log \varepsilon / \log a$, where $0 < a < 1$. So that $x \in [0, a]$ and $n \geq m$ implies that,

$$|x^n - 0| \leq a^n \leq a^m < a^{\log \varepsilon / \log a} = \varepsilon$$

Hence $\langle x^n \rangle$ converges uniformly to 0 on $[0, a]$, where $0 < a < 1$.

On $[0, 1)$ $x^n \to 0$.

NOTES

If possible, suppose x^n converges uniformly to 0 on $[0, 1)$. Then for $\varepsilon = \frac{1}{8}$ $\exists\, m \in \mathbf{N}$, such that

$$|x^n - 0| < \frac{1}{8}\ \forall\ n \geq m \text{ and } \forall\ x \in [0, 1)$$

$$\Rightarrow \left(1-\frac{1}{n}\right)^n < \frac{1}{8} \forall\ \ n \geq m \text{ taking } x = 1-\frac{1}{n} \in [0, 1), \text{ as } n \to \infty$$

$$\Rightarrow e^{-1} \leq \frac{1}{8}, \text{ a Contradiction.}$$

Thus, $\langle x^n \rangle$ does not converge uniformly on $[0, 1)$.

Example 2.16: If $P_0(x) = 0$ and $P_n(x)$ be the sequence of polynomials defined by,

$$P_{n+1}(x) = P_n(x) + \frac{x^2 - P_n^r(x)}{2}$$

For $n = 0, 1, 2, ...$ then $P_n(x)$ converges uniformly to $|x|$ on $[-1, 1]$.

Solution: Here,

$$|x| - P_{n+1}(x) = \{|x| - P_n(x)\}\left(1 - \frac{|x| + P_n(x)}{2}\right),$$

This implies that $0 \leq P_n(x) \leq P_{n+1}(x) \leq 1$, for $n = 0, 1, 2, ...$ and for $\varepsilon > 0\ \exists\ m \in N$ such that $\forall\, |x| \leq 1$.

$$0 \leq |x| - P_n(x) \leq |x|\left(1 - \frac{|x|}{2}\right)^n, \text{ for } n = 0, 1, 2, ...$$

$$\leq \frac{2}{n+1}\left(\frac{n}{n+1}\right)^n \text{ by Maximum}$$

$$< \frac{2}{n+1} \varepsilon\ \forall\ n \geq m\left(> \frac{2}{\varepsilon} - 1\right)$$

Hence, the sequence of polynomials $P_n(x)$ defined as above, converges uniformly to $|x|$ on $[-1, 1]$.

2.5 CAUCHY CRITERION FOR UNIFORM CONVERGENCE

Cauchy's General Principle of Uniform Convergence

Theorem 2.5: A sequence of real valued functions $< f_n >$ defined on a set S converges uniformly on S iff to each given $\varepsilon > 0\ \exists\ m \in \mathbf{N}$ such that,

$$\left|f_{n+p}(x) - f_n(x)\right| < \varepsilon \ \forall \ n \geq m, p \geq 0 \ \text{ and } x \in S.$$

Proof: Let $\langle f_n \rangle$ converges uniformly to f on S. Then for $\varepsilon > 0 \ \exists \ m \in \mathbf{N}$ such that,

$\left|f_n(x) - f(x)\right| < \frac{1}{2}\varepsilon \ \forall \ n \geq m$ and $x \in S$. So that,

$$\left|f_{n+p}(x) - f_n(x)\right| = \left|f_{n+p}(x) - f(x) + f(x) - f_n(x)\right|$$

$$\leq \left|f_{n+p}(x) - f(x)\right| + \left|f_n(x) - f(x)\right|$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon, \ \forall \ n \geq m, \ p \geq 0 \text{ and } x \in S.$$

Hence, the condition in the theorem is necessary.

To prove sufficiency, in view of the Theorem 2.5, let $\langle f_n \rangle$ converge to f on S, then for $\varepsilon > 0 \ \exists \ m \in \mathbf{N}$ such that,

$\left|f_{n+p}(x) - f_n(x)\right| < \frac{1}{2}\varepsilon \ \forall \ n \geq m, p \geq 0$ and $x \in S$. Thus,

$f_m(x) - \frac{1}{2}\varepsilon < f_{m+p}(x) < f_m(x) + \frac{1}{2}\varepsilon, \ \forall \ p \geq 0$ and $x \in S$.

When $p \to \infty$, then we get

$$f_m(x) - \frac{1}{2}\varepsilon \leq \lim_{p \to \infty} f_{m+p}(x) = f(x) \leq f_m(x) + \frac{1}{2}\varepsilon, \ \forall \ x \in S$$

i.e., $\left|f_m(x) - f(x)\right| < \frac{1}{2}\varepsilon, \ \forall \ x \in S.$

Therefore,
$$\left|f_n(x) - f(x)\right| = \left|f_n(x) - f_m(x) + f_m(x) - f(x)\right|$$
$$\leq \left|f_n(x) - f_m(x)\right| + \left|f_m(x) - f(x)\right|$$
$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon, \ \forall \ n \geq m \ \text{ and } x \in S.$$

Hence, $\langle f_n \rangle$ converges uniformly to f on S, i.e., the condition in the theorem is sufficient.

Theorem 2.6: $f_n(x)$ converges uniformly to $f(x)$ as S iff

$\max_{x \in S} \left|f_n(x) - f(x)\right| \to 0$ as $n \to \infty$.

Proof: $\max_{x \in S} \ |f_n(x) - f(x)| \to 0$ as $n \to \infty$.

$\Rightarrow$ For $\varepsilon > 0 \ \exists \ m \in \mathbf{N}$ such that $\max_{x \in S} |f_n(x) - f(x)| < \varepsilon \ \forall \ n \geq m$

$$\Rightarrow \quad |f_n(x) - f(x)| < \varepsilon \ \forall\ n \geq m \text{ and } \forall\ x \in S$$

$$\Rightarrow \quad f_n(x) \text{ converges uniformly to } f(x) \text{ on } S.$$

NOTES

Conversely, if $f_n(x)$ converges uniformly to $f(x)$ on S, then for $\varepsilon > 0\ \exists\ m \in \mathbf{N}$ such that,

$$|f_n(x) - f(x)| < \varepsilon/2 \ \forall\ n \geq m \text{ and } x \in S$$

$$\Rightarrow \max_{x \in S} |f_n(x) - f(x)| \leq \varepsilon/2 < \varepsilon \ \forall\ n \geq m.$$

$$\Rightarrow \max_{x \in S} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty.$$

Example 2.17: The sequence $\langle x^n \rangle$ converges on [0, 1] to the function f defined by $f(x) = 0$ when $x \neq 1$ and $f(1) = 1$. The convergence is not uniform on [0, 1]. For $\varepsilon = \frac{1}{8}$ and any $m \in \mathbf{N}$, let $x = 2^{-1/m}$, then $x^m = \frac{1}{2}$ and $x^{2m} = \frac{1}{4}$ gives $|x^m - x^{2m}| = \left|\frac{1}{2} - \frac{1}{4}\right| > \frac{1}{8}$, i.e., the above theorem is not satisfied.

It can be easily seen that the concept of uniform convergence is compatible with addition and subtraction. But the situation in respect of the multiplication is different. The uniform convergence may or may not remain. For example, the sequences $\left\langle \frac{n+x+1}{(n+1)x} \right\rangle, \left\langle \frac{(n+1)x^2}{1+n^2x^2} \right\rangle$ are uniformly convergent on (0, 1).

For,

$$\left|\frac{n+x+1}{(n+1)x} - \frac{1}{x}\right| = \frac{1}{n+1} \to 0 \Rightarrow \text{Given } \varepsilon > 0\ \exists\ m \in \mathbf{N}:$$

$$\left|\frac{n+x+1}{(n+1)x} - \frac{1}{x}\right| < \varepsilon \ \forall\ n \geq m \text{ and } x \in (0, 1).$$

And,

$$\frac{(n+1)x^2}{1+n^2x^2} = \frac{n+1}{1/x^2 + n^2} < \frac{n+1}{n^2} \to 0 \Rightarrow \text{Given } \varepsilon > 0\ \exists\ m \in \mathbf{N}:$$

$$\left|\frac{(n+1)x^2}{1+n^2x^2} - 0\right| < \varepsilon \ \forall\ n \geq m \text{ and } x \in (0, 1).$$

Now note that though $\left\langle \frac{n+x+1}{(n+1)x} \right\rangle, \left\langle \frac{(n+1)x^2}{1+n^2x^2} \right\rangle$ converge uniformly on

NOTES

(0, 1) but their product sequence $\left\langle \frac{nx+x^2+x}{1+n^2x^2} \right\rangle$ is not uniformly convergent on (0, 1).

The uniform convergence of the product sequence is preserved under additional condition of uniform boundedness.

A sequence $f_n(x)$ defined on a set S is said to be *uniformly bounded* on S if there exists $k > 0$ such that,

$$|f_n(x)| < k \ \forall \ x \in S, \text{ and } n.$$

Under the above additional concept of uniform boundedness we can define that if $f_n(x)$, $g_n(x)$ are uniformly convergent and uniformly bounded on a set S then $f_n(x)g_n(x)$ converges uniformly and is uniformly bounded on S.

Cauchy's General Principle of Convergence for Series

The sequence of partial sums $\langle s_n \rangle$ of a series $\Sigma\, u_n$ converges iff for given $\varepsilon > 0 \ \exists \ m \in \mathbf{N}$ such that,

$$|s_{n+p} - s_n| < \varepsilon \ \forall n \geq m \text{ and } p \geq 0$$

$$\text{i.e., } |u_{n+1} + u_{n+2} + ... + u_{n+p}| < \varepsilon \ \forall \ n \geq m \text{ and } p \geq 0.$$

This condition does not involve a separate evaluation of s_n or R_n.

Therefore, as for sequences, Cauchy's general principle of convergence for series is given by the following theorems.

Theorem 2.7: A series $\Sigma\, u_n$ converges iff to each $\varepsilon > 0 \ \exists \ m \in \mathbf{N}$ such that,

$$|u_{n+1} + u_{n+2} + ... + u_{n+p}| < \varepsilon \ \forall \ n \geq m \text{ and } p \geq 0.$$

Corollary 1: The series $\Sigma\, u_n$ converges iff $R_n \to 0$.

Proof: Let $\Sigma u_n \to s$, then $s_n \to s$. So that $s = s_n + R_n$ implies that $R_n \to 0$. On the other hand, if $R_n \to 0$ then for $\varepsilon > 0 \ \exists \ m \in \mathbf{N}$ such that $|R_n| < \frac{1}{2}\varepsilon \ \forall \ n \geq m$. Thus,

$$\begin{aligned} |u_{n+1} + u_{n+2} + ... + u_{n+p}| &= |R_n - \mathrm{R}_{n+p}| \\ &\leq |R_n| + |R_{n+p}| \\ &\leq \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \ \forall \ n \geq m \text{ and } p \geq 0 \end{aligned}$$

Hence, $\Sigma\, u_n$ converges.

Example 2.18: The series $1 + \frac{1}{3} + \frac{1}{5} + ... + \frac{1}{2n-1} + ...$ diverges to $+\infty$.

Solution: For $\varepsilon = \frac{1}{8}$, if the series converges by the above Theorem 2.7 $\exists \ m \in \mathbf{N}$ such that,

$$\left|\frac{1}{2n+1}+\frac{1}{2n+3}+...+\frac{1}{2n+2p-1}\right|<\frac{1}{8}\forall n\geq m \text{ and } p\geq 1$$

NOTES

In particular, for $n = p = m$,

$$\frac{m}{4m-1}<\left|\frac{1}{2m+1}+\frac{1}{2m+3}+...+\frac{1}{2m+2m-1}\right|<\frac{1}{8} \text{ a Contradiction.}$$

$$\Rightarrow 4m<-1$$

Therefore, the given series is non-convergent, and $\langle s_n \rangle$ being monotonically increasing and non-convergent, diverges to +, ∞, i.e., the given series diverges to $+\infty$.

A Necessary Condition for Convergent Series

Theorem 2.8: If Σu_n is convergent then $\lim u_n = 0$.

Proof: Σu_n is convergent iff its sequence of partial sums $\langle s_n \rangle$ converges.

Let $\lim s_n = l$. So that $s_{n-1} = l$.

It gives $\lim (s_n - s_{n-1}) = l - l = 0$, i.e., $\lim u_n = 0$.

$\Sigma 1/n \rightarrow +\infty$ but $lim\, u_n = \lim 1/n = 0$. Therefore, the condition in the above Theorem 2.8 is only necessary.

A Sufficient Condition for Divergent Series

Theorem 2.9: If $u_n \rightarrow l \neq 0$ then Σu_n diverges to $+\infty$ or $-\infty$ according as $l > 0$, or < 0 and is finite or infinite.

Proof: Let l be finite and $l > 0$.

Then, for $\varepsilon = l/2 > 0 \ \exists\ m \in \mathbf{N}$ such that $|u_n - l| < l/2$, i.e., $l/2 < u_n < 3l/2 \ \forall\ n \geq m$.

This implies that $s_n > (n - m + 1)\, l/2 + u_1 + ... + u_{m-1} \ \forall\ n \geq m$, i.e., $s_n \rightarrow +\infty$.

Hence, Σu_n diverges to $+\infty$ if $l > 0$.

If $l < 0$, for $\varepsilon = -l/2 > 0 \ \exists\ m \in \mathbf{N}$ such that $|u_n - l| < -l/2$, i.e., $3l/2 < u_n < l/2 \ \forall\ n \geq m$.

Thus, $s_n < (n - m + 1)\, l/2 + u_1 + ... + u_{m-1} \ \forall\ n \geq m$, i.e., $s_n \rightarrow -\infty$.

Hence, Σu_n diverges to $-\infty$, if $l < 0$.

Let $u_n \rightarrow +\infty$, then for any $k > 0 \ \exists\ m \in \mathbf{N}$ such that $u_n > k \ \forall\ n \geq m$. So that $s_n > (n - m + 1)\, k + u_1 + ... + u_{m-1} \ \forall\ n \geq m$. Hence $s_n \rightarrow +\infty$, i.e., Σu_n diverges to $+\infty$.

Similarly, if $u_n \rightarrow -\infty$, then for any $k > 0 \ \exists\ m \in \mathbf{N}$ such that $u_n < k \ \forall\ n \geq m$. So that $s_n < (n - m + 1)\, k + u_1 + ... + u_{m-1} \ \forall\ n \geq m$. Hence, $s_n \rightarrow -\infty$, i.e., Σu_n diverges to $-\infty$.

Since $\Sigma\ 1/n \to +\infty$ and $\Sigma\ -1/n \to -\infty$ and $\lim \pm 1/n = 0$, therefore, the condition in the above Theorem 2.9 is not a necessary condition for divergence of a series.

NOTES

Example 2.19: For any $x \in \mathbf{R}$, the series $\Sigma \cos \dfrac{x}{n}$ diverges to $+\infty$.

Solution: Since $\lim\limits_{n\to\infty} \cos\dfrac{x}{n} = 1$ for any $x \in \mathbf{R}$, then the series $\Sigma \cos\dfrac{x}{n}$ diverges to $+\infty$.

2.6 WEIERSTRASS'S M-TEST

Theorem 2.10: A series $\Sigma f_n(x)$ converges uniformly (and absolutely) on a set S if there exists a convergent series $\Sigma\, M_n$ of non-negative terms M_n such that,

$$|f_n(x)| \leq M_n\ \forall\ x \in S \text{ and } n \in \mathbf{N}.$$

Proof: If $\Sigma\, M_n$ is convergent then for $\varepsilon > 0\ \exists\ m \in \mathbf{N}$ such that,

$$|M_{n+1} + M_{n+2} + ... + M_{n+p}| < \varepsilon\ \forall\ n \geq m \text{ and } p \geq 1.$$

Therefore, $\forall\ x \in S$, $|f_{n+1}(x) + f_{n+2}(x) + ... + f_{n+p}(x)|$

$$\leq |f_{n+1}(x)| + |f_{n+2}(x)| + ... + |f_{n+p}(x)|$$

$$\leq M_{n+1} + M_{n+2} + ... + M_{n+p}$$

$$< \varepsilon\ \forall\ n \geq m, p \geq 1 \text{ and } x \in S.$$

Hence, $\Sigma f_n(x)$ converges uniformly on S.

From above given notation $(|f_{n+1}(x)| + |f_{n+2}(x)| + ... + |f_{n+p}(x)|) < \varepsilon\ \forall\ n \geq m,\ p \geq 1$ and $x \in S$, implies that $\Sigma f_n(x)$ also converges absolutely on S.

Example 2.20: The series $\Sigma\, 3^n \sin \dfrac{1}{4^n x}$ converges absolutely and uniformly on (a, ∞) where $a > 0$.

Solution: For any $0 < x \in (a, \infty)\ \exists\ m \in \mathbf{N}$ such that $4^n x \geq 1\ \forall\ n \geq m$. Hence, the series after a finite number of terms consists of positive terms. Since,

$$\lim\left|\frac{u_n}{u_{n+1}}\right| = \lim \frac{1}{3}\,\frac{\sin\dfrac{1}{4^n x}}{\sin\dfrac{1}{4^{n+1} x}} = \frac{4}{3} > 1,$$

Therefore, the given series converges absolutely on (a, ∞), if $a > 0$.

Also, for $n \geq m$, $\sin\dfrac{1}{4^n x} < \dfrac{1}{4^n x} < \dfrac{1}{4^{n-m}}$.

Thus, $\left|3^n \sin\dfrac{1}{4^n x}\right| < 4^m\left(\dfrac{3}{4}\right)^n\ \forall n \geq m.$

Since, $\Sigma 4^m \left(\frac{3}{4}\right)^n$ converges, therefore, by Weierstrass's M-test,

$\Sigma 3^n \sin 1/4^n x$ converges uniformly on (a, ∞), where $a > 0$.

Example 2.21: The series $\sum_{n=1}^{\infty} \frac{x}{1+n^2x^2}$ converges uniformly on $[a, 1]$, where $1 > a > 0$, but not on $[0, 1]$.

Solution: On $[a, 1]$, where $1 > a > 0$,

$$|u_n(x)| = \frac{x}{1+n^2x^2} \le \frac{1}{1+n^2a^2}$$

And $\Sigma 1/(1+n^2a^2)$ converges, therefore, by Weierstrass's M-test, the given series converges uniformly on $[a, 1]$.

On the other hand, let the given series be uniformly convergent on $[0, 1]$, then for $\varepsilon = \frac{1}{8} > 0\ \exists\ m \in \mathbf{N}$ such that (by Cauchy's principle, for $n = p = m$)

$$\left|\frac{x}{1+m^2x^2} + \frac{x}{1+(m+1)^2x^2} + ... + \frac{x}{1+(2m)^2x^2}\right| < \frac{1}{8}, \therefore \left|\frac{mx}{1+(2m)^2x^2}\right| < \frac{1}{8}$$

On putting $x = 1/m$ it gives $1/5 < 1/8$, a contradiction. Therefore, the series given is not uniformly convergent on $[0, 1]$.

Check Your Progress

1. Define the term sequence.
2. When is the sequence bounded above and bounded below?
3. State about the bounded sequence.
4. What do you understand by the term 'Series'?
5. Define pointwise and uniform convergence.
6. State Cauchy's general principle of uniform convergence.
7. What is the Weierstrass's *M*-test?

2.7 ABEL'S TEST FOR UNIFORM CONVERGENCE

Theorem 2.11: If $\Sigma f_n(x)$ converges uniformly on a set S and $\langle g_n(x)\rangle$ be monotonic and uniformly bounded on S, then the series $\Sigma f_n(x)\, g_n(x)$ converges uniformly on S.

Proof: Let ${}_pR_n(x) = f_{n+1}(x) + f_{n+2}(x) + ... + f_{n+p}(x),\ \forall\ x \in S$, then

$$f_{n+1}(x)\, g_{n+1}(x) + f_{n+2}(x)\, g_{n+2}(x) + ... + f_{n+p}(x)\, g_{n+p}(x)$$
$$= {}_1R_n(x)\, g_{n+1}(x) + [{}_2R_n(x) - {}_1R_n(x)]\, g_{n+2}(x) + ... + [{}_pR_n(x) - {}_{p-1}R_n(x)]\, g_{n+p}(x)$$

$= {}_1R_n(x)\,[g_{n+1}(x) - g_{n+2}(x) + ... + {}_{p-1}R_n(x)\,[g_{n+p-1}(x) - g_{n+p}(x)] + {}_pR_n(x)$
$g_{n+p}(x)$ ...(2.1)

Since $\Sigma f_n(x)$ converges uniformly, for $\varepsilon > 0\ \exists\ m \in \mathbf{N}$ such that,

$|{}_pR_n(x)| < \varepsilon\ \forall\ n \geq m$ and $p \geq 1$ and $x \in S$,

Therefore, Equation (2.1) gives

$|f_{n+1}(x)\,g_{n+1}(x) + f_{n+2}(x)\,g_{n+2}(x) + ... + f_{n+p}(x)\,g_{n+p}(x)|$

$< \varepsilon|g_{n+1}(x) - g_{n+2}(x) + g_{n+2}(x) - ... - g_{n+p}(x)| + \varepsilon\,|g_{n+p}(x)|$

$= \varepsilon\,|g_{n+1}(x) - g_{n+p}(x)| + \varepsilon\,|g_{n+p}(x)| < \varepsilon\,.\,2k + \varepsilon\,.$

$k = 3\,k\,\varepsilon,\ \forall\ n \geq m$ and $p \geq 1$ and $x \in S$.

For, $\langle g_n(x)\rangle$ being uniformly bounded $\exists\ k > 0$ such that,

$|g_n(x)| < k\ \forall\ x \in S$ and $n \in \mathbf{N}$.

Hence, $\Sigma f_n(x)\,g_n(x)$ converges uniformly on S.

Uniform convergence of functional series plays important role in term-by-term integration and differentiation of the infinite series.

NOTES

2.8 DIRICHLET'S TEST FOR UNIFORM CONVERGENCE

Theorem 2.12 (**Dirichlet's Test**): Let X be a metric space.

If the functions $f_n: X \to C$, $g_n: X \to R$, $n \in N$ satisfy the following:

1. $F_n(x) = \sum_{m=1}^{n} f_m(x)$ is bounded uniformly in n and x.
2. $g_{n+1} \leq g_n(x)$ for all $x \in X$ and $n \in N$.
3. $\{g_n(x)\}_{n\in N}$ converges uniformly to zero on X.

Then, $\sum_{n=1}^{\infty} f_n(x) g_n(x)$ converges uniformly on X.

Proof: We will prove this by using summation by parts formula.

$$s_n(x) = \sum_{k=1}^{n} f_k(x) g_k(x)$$

$$= F_1(x) g_1(x) + \sum_{k=2}^{n} [F_k(x) - F_{k-1}(x)] g_k(x)$$

$$= F_1(x) g_1(x) + \sum_{k=2}^{n} F_k(x)(g_k)(x) - \sum_{k=2}^{n} F_{k-1}(x) g_k(x)$$

$$= \sum_{k=1}^{n} F_k(x) g_k(x) - \sum_{k=1}^{n-1} F_k(x) g_{k+1}(x)$$

$$= \sum_{k=1}^{n} F_k(x)[g_k(x) - g_{k+1}(x)] + g_{n+1}(x) F_n(x)$$

So if $m > n$, the difference between the mth and nth partial sums is,

$$s_m(x) - s_n(x) = \sum_{k=n+1}^{m} F_k(x)[g_k(x) - g_{k+1}(x)] + g_{m+1}(x)F_m(x) - g_{n+1}(x)F_n(x)$$

NOTES

If $M = \sup\{|F_n(x)|\ x \in X, n \in N\}$,

$$|s_m(x) - s_n(x)| \le m| \sum_{k=1+1}^{m}[g_k(x) - g_{k-1}(x)] + M g_{m+1}(x) + M g_{n+1}(x)$$

$$= M[g_{n+1}(x) - g_{m+1}(x)] + M\, g_{m+1}(s) + M\, g_{n+1}(x) \qquad \ldots (2.2)$$

$$= 2M\, g_{n+1}(x)$$

Since $g_{m+1}(x) \ge 0$, $g_{n+1}(x) \ge 0$ and every $g_k(x) - g_{k+1}(x) \ge 0$

For each fixed x, $\lim_{n\to\infty} g_{n+1}(x) = 0$. So Equation (2.2) guarantees that $\{s_n(x)\}$ is a Cauchy sequence and hence converges. Call the limit $s(x)$. Taking the limit of Equation (2.2) as $m \to \infty$ gives,

$$|s(x) - s_n(x)| \le 2Mg_{n+1}(x)$$

Since $g_{n+1}(x)$ converges uniformly to zero as $n \to \infty$, $s_n(x)$ converges uniformly to $s(x)$ as $n \to \infty$.

Example 2.22: We shall consider the following three different power series: $\sum_{n=0}^{\infty}\left(\frac{z}{R}\right)^n, \sum_{n=0}^{\infty}\frac{1}{n}\left(\frac{z}{R}\right)^n$ and $\sum_{n=0}^{\infty}\frac{1}{n^2}\left(\frac{z}{R}\right)^n$ for some fixed $R > 0$. For all three series, the radius of convergence is exactly R since, for $l \in \{0, 1, 2\}$

$$\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n^l}\frac{1}{R^n}} = \frac{1}{R}\limsup_{n\to\infty}\left(\sqrt[n]{\frac{1}{n}}\right)^l = \frac{1}{R}$$

So all three series converge for all complex numbers z with $|z| < R$ and diverge for all complex numbers with $|z| > R$.

For $|z| = R$, start with the series $\sum_{n=0}^{\infty}\left(\frac{z}{R}\right)^n$. Then we can compute exactly the partial sum,

$$F_n(z) = \sum_{m=0}^{n}\left(\frac{z}{R}\right)^m = \begin{cases} \dfrac{1-\left(\frac{z}{R}\right)^{n+1}}{1-\frac{z}{R}} & \text{if } z \ne R \\ n+1 & \text{if } z = R \end{cases} \qquad \ldots(2.3)$$

Now, if $|z| < R$, this converges to $\dfrac{1}{1-\frac{z}{R}}$ as $n \to \infty$. Also, this diverges for $|z| > R$, because $\left|\left(\frac{z}{R}\right)^{n+1}\right| = \left|\frac{z}{R}\right|^{n+1} \to \infty$.

Claim that this also diverges whenever $|z| = R$. For $z = R$, it is obvious because $n+1 \to \infty$.

Also for $|z| = R$ with $z \neq R$, $\left(\frac{z}{R}\right)^{n+1}$ cannot converge, because

$$\left|\left(\frac{z}{R}\right)^{n+2} - \left(\frac{z}{R}\right)^{n+1}\right| = \left|\frac{z}{R}\right|^{n+1}\left|\frac{z}{R} - 1\right| = \left|\frac{z}{R} - 1\right|$$

This is independent of n. So the geometric series $\sum_{n=0}^{\infty}\left(\frac{z}{R}\right)^n$, which has radius of convergence R, converges if and only if $|z| < R$.

The third series, $\sum_{n=0}^{\infty}\frac{1}{n^2}\left(\frac{z}{R}\right)^n$, converges for all $|z| \leq R$, by comparison with $\sum_{n=0}^{\infty}\frac{1}{n^2}$. As the series has radius of convergence R, it converges if and only if $|z| \leq R$.

The middle series $\sum_{n=0}^{\infty}\frac{1}{n}\left(\frac{z}{R}\right)^n$ has a more interesting domain of convergence. Of course the radius of convergence is exactly R, so the series converges for all complex numbers z with $|z| < R$ and diverges for all complex numbers with $|z| > R$.

For $z = R$, the series is $\sum_{n=0}^{\infty}\frac{1}{n}\left(\frac{z}{R}\right)^n = \sum_{n=0}^{\infty}\frac{1}{n}$, which diverges. So, that leaves $|z| = R$ but with $z \neq R$.

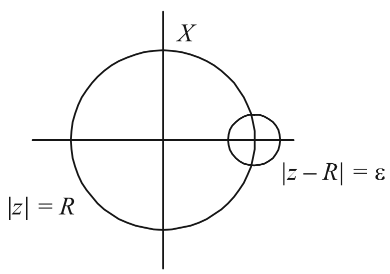
This is where the Dirichlet's test comes in handy. Fix any $\varepsilon > 0$ and set,

$$X = \{z \in \mathbb{C} \mid |z| = R, |z - R| \geq \varepsilon\}$$

$$f_n(z) = \left(\frac{z}{R}\right)^n$$

$$F_n(z) = \sum_{m=0}^{n}\left(\frac{z}{R}\right)^m \text{ as in Equation (2.3)}$$

$$g_n = \frac{1}{n}$$



For $z \in X$,

$$|F_n(z)| = \left|\frac{1-\left(\frac{z}{R}\right)^{n+1}}{1-\frac{z}{R}}\right| \le \frac{1+\left|\frac{z}{R}\right|^{n+1}}{\frac{1}{R}|R-z|} \le \frac{2R}{\varepsilon}$$

So that the hypotheses of the Dirichlet's test are satisfied and the series converges uniformly on X. We conclude that $\sum_{n=0}^{\infty}\frac{1}{n}\left(\frac{z}{R}\right)^n$ converges for $|z| < R$ and for $|z| = R$, $z \neq R$, and diverges for,

$|z| > R$ and for $z = R$.

2.9 UNIFORM CONVERGENCE AND CONTINUITY

Theorem 2.13: Assume that $f_n \to f$ uniformly on S. If each f_n is continuous at a point c of S, then the limit function f is also continuous at c.

Proof: Given that $f_n \to f$ uniformly on S.

$\Rightarrow$ For every $\varepsilon > 0$ there is an M such that,

$|f_n(x) - f(x)| < \frac{\varepsilon}{3}$ for all $n > m$ and for each $x \in S, f_m$ is continuous at c.

$\Rightarrow$ For above $\varepsilon > 0$, we can find a $\delta > 0$ such that,

$|f_m(x) - f_m(c)| < \frac{\varepsilon}{3}$ when $x \in (c - \delta, c + \delta) \cap S$

If $x \in (c - \delta, c + \delta) \cap S$ then,

$$\begin{aligned}
|f(x) - f(c) &= |f(x) - f_m(x) + f_m(x) - f_m(c) + f_m(c) - f(c)| \\
&= |f(x) - f_m(x)| + |f_m(x) - f_m(c)| + |f_m(c) - f(c)| \\
&= \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon
\end{aligned}$$

Or f is continuous at c.

Note: The converse of the theorem is not true, i.e., a sequence of continuous functions may converge to a continuous function, although the convergence is not uniform.

2.10 UNIFORM CONVERGENCE AND RIEMANN- STIELTJES INTEGRATION

Theorem 2.14: Let α be of bounded variation on $[a, b]$. Assume that each term of the sequence f_n is a real valued function such that $f_n \in R(\alpha)$ on $[a, b]$ for each n = 1, 2,... . Assume that $f_n \to f$ uniformly on $[a, b]$ and define, $g_n(x) = \int_a^x f_n(t)d\alpha(t)$ if $x \in [a, b]$, $n = 1, 2,\ldots$, then we have:

(*a*) $f \in R(\alpha)$ on $[a, b]$.

(*b*) $g_n \to g$ uniformly on $[a, b]$, where $g(x) = \int_\alpha^x f(t)d\alpha(t)$.

Proof: It is sufficient to prove the case when α is increasing.

(a) To Prove Part (a)

To prove $f \in R(\alpha)$ on $[a, b]$, it is enough to prove that f satisfies Riemann condition with respect to α.

$f_n \to f$ uniformly on $[a, b]$. This implies that given $\varepsilon > 0$, there exists N such that,

$$|f(x) - f_N(x)| < \frac{\varepsilon}{3[\alpha(b) - \alpha(a)]} \text{ for all } x \in [a, b]$$

Then for any partition P of $[a, b]$ we have,

$$\begin{aligned}
|U(P, f - f_N, \alpha)| &= \left|\sum_{k=1}^n M_k(f - f_N)\Delta\alpha_k\right| \\
&= \sum_{k=1}^n |M_k(f - f_N)|\Delta\alpha_k \\
&< \frac{\varepsilon}{3[\alpha(b) - \alpha(a)]}\sum_{k=1}^n \Delta\alpha_k \\
&= \frac{\varepsilon}{3[\alpha(b) - \alpha(a)]}[\alpha(b) - \alpha(a)] \\
&= \frac{\varepsilon}{3}
\end{aligned}$$

Similarly we get, $|u(P, f - f_N, \alpha)| < \frac{\varepsilon}{3}$ for each $f_n \in R(\alpha)$ on $[a, b]$. In particular $f_N \in R(\alpha)$ on $[a, b]$. So, for the above $\varepsilon > 0$ we can find a partition P_ε such that,

$$U(P, f_N, \alpha) - L(P, f_N, \alpha) < \frac{\varepsilon}{3} \text{ for all } P \supseteq P_\varepsilon$$

Then for such P we have,

$$U(P, f, \alpha) - L(P, f, \alpha) = U(P, h + f_N, \alpha) - L(P, h + f_N, \alpha), \text{ where } h = f - f_N$$

$$= \ U(P, h, \alpha) + U(P, f_N, \alpha) - L(P, h, \alpha) - L(P, f_N, \alpha)$$

NOTES

$$= U(P,h,\alpha) - L(P,h,\alpha) + U(P,f_N,\alpha) - L(P,f_N,\alpha)$$

$$= U(P,f-f_N,\alpha) - L(P,f-f_N,\alpha) + U(P,f_N,\alpha) - L(P,f_N,\alpha)$$

$$\leq U(P,f-f_N,\alpha) - L(P,f-f_N,\alpha) + U(P,f_N,\alpha) - L(P,f_N,\alpha)$$

$$\leq |U(P,f-f_N,\alpha)| + |L(P,f-f_N,\alpha)| + U(P,f_N,\alpha) - L(P,f_N,\alpha)$$

$$< \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}$$

$$= \varepsilon$$

$\Rightarrow f \in R(\alpha)$ on $[a, b]$

To Prove Part (b)

$f_n \to f$ uniformly on $[a, b]$

$\Rightarrow$ For any $\varepsilon > 0$ there exists an N such that,

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2[\alpha(b)-\alpha(a)]} \text{ for all } n > N \text{ and every } t \in [a,b]$$

If $x \in [a, b]$ then,

$$\begin{aligned} |g_n(x) - g(x)| &= \left|\int_a^x f_n(t)d\alpha(t) - \int_a^x f(t)d\alpha(t)\right| \\ &= \left|\int_a^x [f_n(t) - f(t)]d\alpha(t)\right| \\ &\leq \int_a^x |f_n(t) - f(t)|\, d\alpha(t) \\ &< \frac{\varepsilon}{2[\alpha(b)-\alpha(a)]}\int_a^x d\alpha(t)[\alpha(b)-\alpha(a)] \\ &= \frac{\varepsilon}{2[\alpha(b)-\alpha(a)]} \\ &\leq \frac{\varepsilon}{2} \\ &< \varepsilon \end{aligned}$$

$\Rightarrow g_n \to g$ uniformly on $[a, b]$

Theorem 2.15: Let α be of bounded variation on $[a, b]$ and assume that $\sum f_n(x) = f(x)$ (uniformly on $[a, b]$), where each f_n is a real valued function such that $f_n \in R(\alpha)$ on $[a, b]$. Then we have,

(a) $f \in R(\alpha)$ on $[a, b]$.

(b) $\int_a^x \sum_{n=1}^{\infty} f_n(t)d\alpha(t) = \sum_{n=1}^{\infty}\int_a^x f_n(t)d\alpha(t)$ (uniformly on $[a, b]$).

Proof: Define $s_n = \sum_{k=1}^{n} f_k(x)$. Then $s_n \in R(\alpha)$, since each $f_k \in R(\alpha)$ on $[a, b]$. Also, $\sum f_n(x)$ converges to f uniformly on $[a, b]$. So, s_n converges to f uniformly on $[a, b]$. Then, $f \in R(\alpha)$ on $[a, b]$. To Prove Part (b), $g_n(x) = \int_a^x s_n(t)d\alpha(t)$.

Then by Weierstrass M-test, $g_n \to g$ uniformly on $[a, b]$ where $g(x) = \int_a^x f(t)d\alpha(t)$, or $\lim_{n\to\infty} g_n(x) = g(x)$.

$$\Rightarrow \quad \lim_{n\to\infty} \int_a^x s_n(t)d\alpha(t) = \int_a^x f(t)d\alpha(t)$$

$$\Rightarrow \quad \lim_{n\to\infty} \int_a^x \sum_{k=1}^{n} f_k(t)d\alpha(t) = \int_a^x \sum_{n=1}^{\infty} f_n(t)d\alpha(t)$$

$$\Rightarrow \quad \lim_{n\to\infty} \sum_{k=1}^{n} \int_a^x f_k(t)d\alpha(t) = \int_a^x \sum_{n=1}^{\infty} f_n(t)d\alpha(t)$$

$$\Rightarrow \quad \sum_{k=1}^{\infty} \int_a^x f_k(t)d\alpha(t) = \int_a^x \sum_{n=1}^{\infty} f_n(t)d\alpha(t)$$

$$\Rightarrow \quad \sum_{n=1}^{\infty} \int_a^x f_n(t)d\alpha(t) = \int_a^x \sum_{n=1}^{\infty} f_n(t)d\alpha(t)$$

NOTES

2.11 UNIFORM CONVERGENCE AND DIFFERENTIATION

Theorem 2.16: Assume that each term of $\{f_n\}$ is a real valued function having a finite derivative at each point of an open interval (a, b). Assume that for at least one point x_0 in (a, b), the sequence $\{f_n(x_0)\}$ converges. Assume further that there exists a function g such that $f_n' \to g$ uniformly on (a, b). Then,

(a) There exists a function f such that $f_n \to f$ uniformly on (a, b).

(b) For each x in (a, b) the derivative $f'(x)$ exists and equals $g(x)$.

Proof:

(a) Assume that $c \in (a, b)$ and define $\{g_n\}$ as,

$$g_n(x) = \begin{cases} \dfrac{f_n(x) - f_n(c)}{x - c} & \text{if} \quad x \neq c \\ f'_n(c) & \text{if} \quad x = c \end{cases} \qquad \dots (2.4)$$

Now,

$$\lim_{x\to c} g_n(x) = f'_n(c)$$
$$= g_n(c)$$

Therefore, g_n is continuous for each n.

Also,

$$\lim_{n\to\infty} g_n(c) = \lim_{n\to\infty} f'_n(c)$$
$$= g(c)$$

That is, $\{g_n(c)\}$ is convergent and so is Cauchy. To prove that $\{g_n\}$ converges uniformly on (a, b), let $\varepsilon > 0$ be given. Since, $f_n' \to g$ uniformly on (a, b) and $\{g_n(c)\}$ is a Cauchy sequence,

NOTES

$\Rightarrow$ There exists a $k \in N$ such that,

$$|f'_n(x) - f'_m(x)| < \frac{\varepsilon}{2} \text{ for all } x \in (a,b), n, m \geq k \text{ and}$$

$$|g_n(c) - g_m(c)| < \frac{\varepsilon}{2} \text{ for all } n, m \geq k \quad \ldots(2.5)$$

Now, if $x \in (a, b)$, $x \neq c$ and n, $m \geq k$, then we have,

$$|g_n(x) - g_m(x)| = \left|\frac{f_n(x) - f_n(c)}{x - c} - \frac{f_m(x) - f_m(c)}{x - c}\right|$$

$$= \left|\frac{f_n(x) - f_m(x) - [f_n(c) - f_m(c)]}{x - c}\right|$$

$$\Rightarrow \quad |g_n(x) - g_m(x)|$$

$$= \left|\frac{h(x) - h(c)}{x - c}\right| \quad \ldots(2.6)$$

Clearly, h is differentiable on (a, b).

Therefore, by mean value theorem, there exists a point x_1 in between x and c such that,

$h(x) - h(c) = h'(x_1)(x - c)$

Now from Equation (2.6) we get,

$$|g_n(x) - g_m(x)| = \frac{|h'(x_1)||x - c|}{|x - c|}$$

$$= |h'(x_1)|$$

$$= |f'_n(x_1) - f'_m(x_1)|$$

$$< \frac{\varepsilon}{2}, \text{ for } n, m \geq k \text{ and } x \in (a,b) \text{ with } x \neq c$$

$$\Rightarrow \quad |g_n(x) - g_m(x)| < \varepsilon \quad \ldots(2.7)$$

From Equations (2.5) and (2.7) we get,

$$|g_n(x) - g_m(x)| < \varepsilon \text{ for } n, m \geq k \text{ and } x \in (a,b) \quad \ldots(2.8)$$

which implies that $\{g_n\}$ is uniformly convergent.

To prove that $f_n \to f$ uniformly on (a, b), form the particular sequence $\{g_n\}$ corresponding to the special point $c = x_0$ for which $\{f_n(x_0)\}$ is assumed to converge.

Now from Equation (2.4) we get,

$f_n(x) = f_n(x_0) + (x - x_0)g_n(x)$ for all $x \in (a, b)$

Hence we have,

$f_n(x) - f_m(x) = f_n(x_0) - f_m(x_0) + (x - x_0)\,[g_n(x) - g_m(x)]$

NOTES

Let $\varepsilon > 0$ be given.

g_n converges uniformly on (a, b) and $\{f_n(x_0)\}$ is a Cauchy sequence. This implies that there exists a $k \in N$ such that,

$$|g_n(x) - g_m(x)| < \frac{\varepsilon}{2(b-a)} \text{ and,}$$

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2} \text{ for all } n, m \geq k \text{ and } x \in (a, b)$$

Hence if $n, m \geq k$ and $x \in (a, b)$ we get,

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x_0) - f_m(x_0)| + |x - x_0||g_n(x) - g_m(x)| \\ &< \frac{\varepsilon}{2} + (b-a)\frac{\varepsilon}{2(b-a)} \\ &= \varepsilon \end{aligned}$$

That is, $\{f_n\}$ satisfies Cauchy's condition on (a, b) and so $\{f_n\}$ converges uniformly on (a, b) say to f.

(b) Assume $c \in (a, b)$ then,

$$\begin{aligned} \lim_{x\to c}\frac{f(x)-f(c)}{x-c} &= \lim_{x\to c}\frac{\lim_{n\to\infty} f_n(x) - \lim_{n\to\infty} f_n(c)}{x-c} \\ &= \lim_{x\to c}\lim_{n\to\infty}\frac{f_n(x)-f_n(c)}{x-c} \\ &= \lim_{x\to c}\lim_{n\to\infty} g_n(x) \\ &= \lim_{n\to\infty}\lim_{x\to c} g_n(x) \\ &= \lim_{n\to\infty} g_n(c) \\ &= \lim_{n\to\infty} f'_n(c) \\ &= g(c), \text{ which exists.} \end{aligned}$$

$\Rightarrow f$ is differentiable at c and $f'(c) = g(c) = \lim_{n\to\infty} f'_n(c)$

$\Rightarrow f_n'(c) \to f'(c)$

But, since c is arbitrary, $f_n'(x) \to f'(x)$.

Theorem 2.17: Assume that each $\{f_n\}$ is a real valued function defined on (a, b), such that the derivative exists for each $x \in (a, b)$. Assume that, for at least one point x_0 in (a, b), the series $\sum f_n(x_0)$ converges. Assume further that there exists a function g such that $\sum f'_n(x) = g(x)$ (uniformly on (a, b)). Then,

(a) There exists a function f such that $f_n(x) = f(x)$ (uniformly on (a, b)).

(b) If $x \in (a, b)$, the derivative $f'(x)$ exists and equals $\sum f'_n(x)$.

2.12 WEIERSTRASS APPROXIMATION THEOREM

NOTES

Theorem 2.18 (Weierstrass Approximation Theorem): Let I be a closed and bounded interval. Suppose $f: I \rightarrow R$ is a continuous function. Then for each $\varepsilon > 0$, there exists a polynomial function $p_\varepsilon: I \rightarrow R$ such that,

$|f(x) - p_\varepsilon(x)| < \varepsilon$ for all x in I, or equivalently,

$\sup\ \{|f(x) - p_\varepsilon(x)|\ : x \in I\} < \varepsilon$.

We will prove this as follows:

1. The polynomial functions form a subalgebra that separates points of I.
2. The closure of this subalgebra is a lattice in $C(I, R)$, the space of all continuous function on I with the sup norm.
3. Using the compactness of I, and (1) and (2), one can find a point on this lattice which is arbitrarily near f. This relies on compactness and argument involving finite subcover of an open cover.

We shall prove a special case of Theorem 2.18 when $I = [0, 1]$ first. We now describe the Bernstein polynomials.

Let f: $[0,1] \rightarrow R$ be a function. Then for each integer $n \geq 0$, we define the Bernstein polynomial of degree n associated with f to be,

$$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$$

Theorem 2.19: Suppose f: $[0, 1] \rightarrow R$ is a continuous function. Then for each $\varepsilon > 0$, there exists a polynomial function p: $I \rightarrow R$ such that,

$\sup\{|f(x) - p(x)| : x \in I\} < \varepsilon$

More specially, the sequence of Bernstein polynomial, $(B_n(f))$, as defined above converges uniformly to f.

Before we proceed with the proof, we shall derive a series of identities, which are needed for the proof. The binomial theorem states that for integer $n \geq 0$,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \quad \ldots (2.9)$$

Hence, from Equation (2.9), for integer $n \geq 1$,

$$(x+y)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} x^k y^{n-k-1} \quad \ldots (2.10)$$

Multiplying Equation (2.10) by nx, we obtain

$$nx(x+y)^{n-1} = \sum_{k=0}^{n-1} n\binom{n-1}{k} x^{k+1} y^{n-k-1} \quad \ldots (2.11)$$

NOTES

Now, $n\binom{n-1}{k} = n.\frac{(n-1)!}{k!(n-1-k)!}$

$$= (k+1)\frac{n!}{(k+1)!(n-1-k)!}$$

$$= (k+1)\binom{n}{k+1} \text{ and so}$$

From Equation (2.11), we have

$$nx(x+y)^{n-1} = \sum_{k=0}^{n-1} n\binom{n-1}{k} x^{k+1} y^{n-k-1}$$

$$= \sum_{k=0}^{n-1} (k+1)\binom{n}{k+1} x^{k+1} y^{n-k-1}$$

$$= \sum_{k=1}^{n} k\binom{n}{k} x^k y^{n-k}$$

$$= \sum_{k=0}^{n} k\binom{n}{k} x^k y^{n-k}$$

Evidently, the above equality is true when $n = 0$ and so we have that for any integer $n \geq 0$,

$$nx(x+y)^{n-1} = \sum_{k=0}^{n} k\binom{n}{k} x^k y^{n-k} \qquad \ldots (2.12)$$

From Equation (2.12) for integer $n \geq 1$,

$$(n-1)x(x+y)^{n-2} = \sum_{k=0}^{n-1} k\binom{n-1}{k} x^k y^{n-k-1} \qquad \ldots (2.13)$$

Multiplying Equation (2.13) by nx, similarly, we obtain

$$n(n-1)x^2(x+y)^{n-2} = \sum_{k=0}^{n-1} kn\binom{n-1}{k} x^{k+1} y^{n-k-1}$$

$$= \sum_{k=0}^{n-1} k(k+1)\binom{n}{k+1} x^{k+1} y^{n-k-1}$$

$$= \sum_{k=1}^{n} (k-1)k\binom{n}{k} x^k y^{n-k}$$

$$= \sum_{k=0}^{n} (k-1)k\binom{n}{k} x^k y^{n-k}$$

NOTES

The above equality also holds when $n = 0$ and so we have for any integer $n \geq 0$,

$$n(n-1)x^2(x+y)^{n-2} = \sum_{k=0}^{n}(k-1)k\binom{n}{k}x^k y^{n-k} \qquad \ldots (2.14)$$

Note that the identities in Equations (2.9), (2.12) and (2.14) contain the same factor $\binom{n}{k}x^k y^{n-k}$.

Let $r_k(x) = \binom{n}{k}x^k y^{n-k}$.

Now taking y to be $1-x$, so that $x+y=1$, we have

$$r_k(x) = \binom{n}{k}x^k y^{n-k}$$

$$= \binom{n}{k}x^k (1-x)^{n-k}$$

We obtain from Equation (2.9),

$$1 = (x+y)^n = \sum_{k=0}^{n} r_k(x)$$

Or, for any integer $n \geq 0$

$$1 = \sum_{k=0}^{n} r_k(x) \qquad \ldots (2.15)$$

Similarly, from Equation (2.12) we obtain for any integer $n \geq 0$,

$$nx = \sum_{k=0}^{n} kr_k(x) \qquad \ldots (2.16)$$

And from Equation (2.14) we obtain for any integer $n \geq 0$,

$$n(n-1)x^2 = \sum_{k=0}^{n}(k-1)kr_k(x) \qquad \ldots (2.17)$$

Then for any integer $n \geq 0$,

$$\sum_{k=0}^{n}(k-nx)^2 r_k(x) = \sum_{k=0}^{n} n^2x^2 r_k(x) - 2nx\sum_{k=0}^{n} kr_k(x) + \sum_{k=0}^{n} k^2 r_k(x)$$

$$= n^2x^2\sum_{k=0}^{n} r_k(x) - 2nx\sum_{k=0}^{n} kr_k(x) + \sum_{k=0}^{n}[k-1)k+k]r_k(x)$$

$$= n^2x^2 - 2nx \,.\, nx + nx + n(n-1)x^2$$

$$= nx(1-x)$$

By Equations (2.15), (2.16) and (2.17).

Therefore, for any integer $n \geq 0$,

$$\sum_{k=0}^{n} (k - nx)^2 r_k(x) = nx(1 - x) \qquad \text{... (2.18)}$$

NOTES

We now proceed to the proof of Theorem 2.19.

Proof of Theorem 2.19: Since f is continuous and $[0,1]$ is compact, the image $f([0,1])$ is compact and so by the Heine-Borel theorem, $f([0, 1])$ is closed and bounded. Therefore, there exists a real number $M > 0$ such that $|f(x)| \leq$ M for all x in $[0, 1]$.

Given $\varepsilon > 0$, since f is continuous on $[0, 1]$ there exists $\delta > 0$ such that for all x, y in $[0, 1]$,

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon/2 \qquad \text{... (2.19)}$$

We now estimate how close the Bernstein polynomial $B_n(f)$ is to f for integer $n \geq 1$.

$$|f(x) - B_n(f)(x)| = \left| f(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k} \right|$$

$$= \left| f(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) r_k(x) \right|$$

$$= \left| f(x)\sum_{k=0}^{n} r_k(x) - \sum_{k=0}^{n} f\left(\frac{k}{n}\right) r_k(x) \right| \text{(Using Equation (2.15))}$$

$$= \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right)\right) r_k(x) \right| \qquad \text{...(2.20)}$$

We next examine the sum on the right of Equation (2.20) according to whether $|x - \frac{k}{n}| < \delta$ or $|x - \frac{k}{n}| \geq \delta$, where δ is given in Equation (2.19).

If $|x - \frac{k}{n}| < \delta$, then by Equation (2.19),

$$\left| f(x) - f\left(\frac{k}{n}\right) \right| < \frac{\varepsilon}{2} \qquad \text{... (2.21)}$$

Suppose now $|x - \frac{k}{n}| \geq \delta$.

Then, $|nx - k| \geq n\delta$.

Hence,

$$\left| f(x) - f\left(\frac{k}{n}\right) \right| = |f(x)| + |f\left(\frac{k}{n}\right)| \leq 2M$$

$$\leq 2M \frac{(nx-k)^2}{n^2\delta^2}$$

Because $\quad \dfrac{|nx-k|}{n\delta} \geq 1 \qquad \ldots (2.22)$

Therefore, for any x in [0, 1] and for $0 \leq k \leq n$,

NOTES

$$\left| f(x) - f\left(\frac{k}{n}\right) \right| \leq \frac{\varepsilon}{2} + \frac{2M}{\delta^2}\left(x - \frac{k}{n}\right)^2 \qquad \ldots (2.23)$$

We add the sum ε/2 so that we can combine Equations (2.21) and (2.22) in one inequality for simplicity.

Using Equation (2.20) and the fact that $r_k(x) \geq 0$ for all x in [0, 1] and $0 \leq k \leq n$, we get

$$\begin{aligned}
|f(x) - B_n(f)(x)| &= \left| \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right)\right) r_k(x) \right| \\
&\leq \sum_{k=0}^{n} \left| \left(f(x) - f\left(\frac{k}{n}\right)\right) \right| r_k(x) \\
&\leq \sum_{k=0}^{n} \left(\frac{\varepsilon}{2} + \frac{2M}{\delta^2}\left(x - \frac{k}{n}\right)^2 \right) r_k(x) \quad \text{By Equation (2.23)} \\
&= \frac{\varepsilon}{2}\sum_{k=0}^{n} r_k(x) + \frac{2M}{\delta^2}\sum_{k=0}^{n}\left(x - \frac{k}{n}\right)^2 r_k(x) \\
&= \frac{\varepsilon}{2} + \frac{2M}{\delta^2 n^2}\sum_{k=0}^{n}(nx-k)^2 r_k(x) \qquad \text{By Identity (2.15)} \\
&= \frac{\varepsilon}{2} + \frac{2M}{\delta^2 n^2} nx(1-x) \qquad \text{By Identity (2.18)} \\
&= \frac{\varepsilon}{2} + \frac{2M}{\delta^2 n} x(1-x) \\
&< \frac{\varepsilon}{2} + \frac{2M}{\delta^2 n}
\end{aligned}$$

Because $x\,(1-x) < 1$ for x in [0,1].

Hence, for any x in [0,1] and any $n \geq 1$,

$$|f(x) - B_n(f)(x)| < \frac{\varepsilon}{2} + \frac{2M}{\delta^2 n} \qquad \ldots (2.24)$$

Since $\dfrac{2M}{\delta^2 n} \to 0$, there exists a positive integer N such that,

$$n \geq N \Rightarrow \frac{2M}{\delta^2 n} < \varepsilon/4.$$

It then follows from Equation (2.24) that,

$$n \geq N \Rightarrow |f(x) - B_n(f)(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4}$$

$$= \frac{3\varepsilon}{4} \text{ for all } x \text{ in } [0, 1].$$

Hence, $n \geq \mathrm{N} \Rightarrow \sup\{|f(x) - B_n(f)(x)| : x \in [0,1]\} \leq \frac{3\varepsilon}{4} < \varepsilon$ for all x in $[0,1]$. This shows that $B_n(f) \to f$ uniformly on $[0,1]$. We may take the polynomial function p to be $B_N(f)$ and $\sup\{|f(x) - p(x)| : x \in [0,1]\} < \varepsilon$. This completes the proof of Theorem 2.19.

Proof of Theorem 2.18: Suppose $I = [a, b]$ is a closed and bounded interval and $f: I \to R$ is a continuous function. Let $g: [0, 1] \to [a, b]$ be the bijective linear map defined by $g(t) = a + t(b - a)$ for t in $[0,1]$. Then g is continuous, $g(0) = a$ and $g(1) = b$. Since f is continuous, the composite $f \circ g : [0,1] \to R$ is also continuous. Hence, by Theorem 2.19, for any $\varepsilon > 0$, there exists a positive integer N such that for any integer $n \geq N$, the Bernstein polynomial $B_n(f \circ g)$ satisfies,

$$|f \circ g(x) - B_n(f \circ g)(x)| < \varepsilon \text{ for all } x \text{ in } [0, 1]. \qquad \dots (2.25)$$

Now g is continuous injective map and so g has a continuous inverse function. Indeed the inverse function $g^{-1}: [a, b] \to [0,1]$ is given by,

$g^{-1(t)} = t{-}a/b{-}a$ for t in $[a, b]$

Thus by Equation (2.25), for all t in $[a, b]$,

$$|f(t) - B_N(f \circ g)(g^{-1}(t))| < \varepsilon$$

Hence, $\left| f(t) - B_N(f \circ g)\left(\dfrac{t-a}{b-a}\right) \right| < \varepsilon$, for all t in $[a, b]$.

Since $B_n(f \circ g)$ is a polynomial function, $p_\varepsilon(t) = B_N(f o g)\left(\dfrac{t-a}{b-a}\right)$ is a polynomial in t and $|f(t) - p_\varepsilon(t)| < \varepsilon$ for all x in I.

If we let $q_n(t) = B_n(f o g)\left(\dfrac{t-a}{b-a}\right)$, then

$$q_n(t) = B_n(f \circ g)\left(\frac{t-a}{b-a}\right) = \sum_{k=0}^{n} f \circ g\left(\frac{k}{n}\right)\binom{n}{k}\left(\frac{t-a}{b-a}\right)^k \left(1 - \left(\frac{t-a}{b-a}\right)\right)^{n-k}$$

$$= \sum_{k=0}^{n} f \circ g\left(\frac{k}{n}\right)\binom{n}{k}\left(\frac{t-a}{b-a}\right)^k \left(\frac{b-t}{b-a}\right)^{n-k}$$

$$= \sum_{k=0}^{n} f\left(a + \frac{k}{n}(b-a)\right)\binom{n}{k}\left(\frac{t-a}{b-a}\right)^k \left(\frac{b-t}{b-a}\right)^{n-k}$$

It follows from Equation (2.25) that q_n converges uniformly on $[a, b]$.

NOTES

Check Your Progress

8. Define the Abel's test for uniform convergence.
9. Write Dirichlet's test for uniform convergence.
10. Can a sequence of continuous functions converge to a continuous function?
11. State a necessary condition for a sequence to converge uniformly in an open interval.
12. What does the Weierstrass approximation theorem state?
13. Define Bernstein polynomial.

2.13 POWER SERIES

An infinite series, in the ascending integral powers of a real variable x, of the form

$$a_0 + a_1 x + a_2 x^2 + ... + a_n x^n + ...,$$

where the coefficients $a_0, a_1, a_2,...$ are constant, independent of x, is called a real power series. We shall simply call it a *Power Series* (P.S.).

The primary characteristics regarding convergence of the power series are given below in Theorems 2.20, and 2.21.

Hadamard's Formula

Theorem 2.20: The power series $a_0 + a_1x + a_2x^2 + ...$ converges absolutely only at $x = 0$, on $(-r, r)$, or on every bounded interval according as,

$$\overline{\lim}|a_n|^{1/n} = \infty, \frac{1}{r}, \text{ or } 0.$$

According to these three cases the power series is said to have the radius of convergence zero, r, or infinity.

Proof: $\overline{\lim}|a_n|^{1/n} = \infty \Rightarrow \overline{\lim}|a_n|^{1/n}|x| = 0 \text{ iff } x = 0$

$\Rightarrow$ The power series converges absolutely only at $x = 0$

$$\overline{\lim}|a_n|^{1/n} = \frac{1}{r} \Rightarrow \overline{\lim}|a_n|^{1/n}|x| = \frac{|x|}{r} < 1 \text{ iff } |x| < r$$

$\Rightarrow$ The power series converges absolutely on $(-r, r)$

$$\overline{\lim}|a_n|^{1/n} = 0 \Rightarrow \overline{\lim}|a_n|^{1/n}|x| = 0 < 1 \text{ for every finite } x$$

$\Rightarrow$ The power series converges absolutely on every bounded interval

Corollary 1: The power series $a_0 + a_1x + a_2x^2 + ...$ and its derived power series $a_1 + 2a_2x + 3a_3x^2 + ...$ have alike radius of convergence. Thus, it follows by induction, that the successive derived power series have the same radius of convergence.

Proof: $\lim n^{1/n} = 1$ implies that $\overline{\lim}|na_n|^{1/n} = \overline{\lim}|a_n|^{1/n}$, therefore, both of the given power series have alike radius of convergence. The rest of the statement of the corollary readily follows by induction.

Similar to the above Corollary, we have the following Corollary.

Corollary 2: The power series $a_0 + a_1x + a_2x^2 + ...$ and its integrated power series $c + a_0x + \frac{a_1x^2}{2} + \frac{a_1x^3}{3} + ...$ (where c is a Constant)

have alike radius of convergence. And, it follows by induction that the successive integrated power series have the same radius of convergence.

As an example, note that the power series,

$$(i)\ 1 + 2x + 3x^2 + ... \quad (ii)\ 1 + \frac{x}{r} + \left(\frac{x}{r}\right)^2 + ... \quad (iii)\ 1 + x = \frac{x^2}{2!} + ...,$$ have

radius of convergence 0, r and ∞, respectively.

The power series,

$$(i)\ 1 - \frac{x}{r} + \frac{x^2}{2r^2} - \frac{x^3}{3r^3} + ... \qquad (ii)\ 1 + \frac{x}{r} + \frac{x^2}{2r^2} + \frac{x^3}{3r^3} + ...$$

$$(iii)\ 1 + \frac{x}{r} + \left(\frac{x}{r}\right)^2 + \left(\frac{x}{r}\right)^3 + ... \qquad (iv)\ 1 + \frac{x}{r} + \frac{x^2}{2!r^2} + \frac{x^3}{3!r^3} + ...$$

readily illustrate that if r (> 0) be the radius of converges of a power series, then the power series may or may not be convergent for $x = \pm r$ (four cases).

In case $\lim|a_n|^{1/n}$ exists and is evaluated, it outrightly provides the radius of convergence as per the preceding theorem.

For practical purpose, when $\lim\left|\frac{a_{n+1}}{a_n}\right|$ is more convenient to evaluate than $\lim|a_n|^{1/n}$, then $\lim|a_n|^{1/n}$ exists and,

$$\lim\left|\frac{a_{n+1}}{a_n}\right| = \lim|a_n|^{1/n}$$

Consequently, the radius of convergence of the power series $\sum a_n x^n$ is given by,

$$\lim\left|\frac{a_n}{a_{n+1}}\right|$$

Theorem 2.21: If r (> 0) be the radius of convergence of the power series $a_0 + a_1x + a_2x^2 + ...$, in $(-r, r)$. Then prove that it converges uniformly and absolutely on every close interval contained in $(-r, r)$.

NOTES

Proof: For $0 < \alpha < r$, we have

$$\left|a_n x^n\right| < \left|a_n\right| \alpha^n \ \forall x \in [-\alpha, \alpha]$$

And since $\sum \left|a_n\right| \alpha^n$ is convergent as $\left|a_n\right| \alpha^n < \left|a_n\right| r^n$ and $\sum \left|a_n\right| r^n$ is convergent,

Therefore, $\sum_{n=0}^{\infty} a_n x^n$, i.e., $a_0 + a_1 x + a_2 x^2 + ...$ is uniformly convergent (by *M*-test and absolutely by comparison test). And, if $[a, b] \subset (-r, r)$ and let $\alpha = \max \{|a|, |b|\}$, so that $[a, b] \subset [-\alpha, \alpha] \subset [-r, r]$.

Thus, the given power series converges uniformly and absolutely on every closed interval contained in $(-r, r)$.

Note that with slight modification the analysis done in this section can be conveniently applied to the power series of the form,

$a_0 + a_1(x - \alpha) + a_2 (x - \alpha)^2 + ...$

Regarding the domain of convergence it may also be observed that if a series is pointwise convergent on every $[a, b]$ contained in an interval I then it is pointwise convergent on I but this may or may not hold for uniform convergence.

Certain infinite power series are of basic importance and represent familiar functions such as,

(*i*) $\log (1 + x) \equiv x - \frac{x^2}{2} + \frac{x^3}{3} - ... \forall x \in (-1, 1).$

(*ii*) $e^x \equiv 1 + x + \frac{x^2}{2!} + ... \forall x.$

(*iii*) $\sin x \equiv x - \frac{x^3}{3!} + \frac{x^5}{5!} - ... \forall x.$

(*iv*) $\cos x \equiv 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - ... \forall x.$

Each of the above power series converges for the values of x as given. These enable to establish many useful results.

Example 2.23: Show that,

$$0 < |x| < \frac{1}{2} \Rightarrow \frac{1}{6} x^2 < x - \log (1 - x) < \frac{5}{6} x^2.$$

Solution: For $0 < x < 1$, we have

$$\left|\frac{x - \log(1 + x)}{x^2} - \frac{1}{2}\right| = x\left|\frac{1}{3} - \frac{x}{4} + \frac{x^2}{5} - ...\right| < \frac{1}{3}, \qquad ...(1)$$

$$\therefore \qquad 0 < x < 1 \Rightarrow \frac{1}{6} x^2 < x - \log (1 + x) < \frac{5}{6} x^2.$$

NOTES

And for $-\frac{1}{2} < x < 0$, we have

$$\left|\frac{x-\log(1+x)}{x^2}-\frac{1}{2}\right| = |x|\left(\frac{1}{3}+\frac{|x|}{4}+\frac{|x|^2}{5}+...\right)$$

$$< |x|\left(\frac{1}{3}+\frac{|x|}{4}(1+|x|+|x|^2+...)\right) < |x|\left(\frac{1}{3}+\frac{|x|}{4}.\frac{1}{1-|x|}\right)$$

$$< \frac{1}{2}\left(\frac{1}{3}+\frac{1}{4}\right) < \frac{1}{3},$$

$$\therefore \qquad -\frac{1}{2} < x < 0 \Rightarrow -\frac{1}{6}x^2 < x - \log(1+x) < \frac{5}{6}x^2. \qquad ...(2)$$

Combining Equations (1) and (2), we get that,

$$0 < |x| < \frac{1}{2} \Rightarrow \frac{1}{6}x^2 < x - \log(1+x) < \frac{5}{6}x^2.$$

2.14 UNIQUENESS THEOREM FOR POWER SERIES

Theorem 2.22 (Uniqueness): Let $\Omega \subset C$ be a region and consider two analytic functions,

$$f, g : \Omega \to C.$$

Let S be a subset of Ω that has a limit point $p \in \Omega$ which need not lie in S. Suppose that $f = g$ on S. Then $f = g$. For example, the unique analytic function on C that vanishes at 1, 1/2, 1/4, 1/8, 1/16,…., is the zero function. Also, the extensions by power series of e^x, sin x, cos x and log x from their domains in R to analytic functions on C or on C minus the negative real axis for log, are the unique possible such extensions.

Proof: We may assume that $g = 0$. That is, we may assume that $f = 0$ on S. Let $B = B(p, r)$ be the largest ball about p in Ω.

Possibly $r = +\infty$, but in any case $r > 0$. The power series representation of f at p is,

$$f(z) = a_0 + a_1(z-p) + a_2(z-p)^2 + ..., \; z \in B.$$

Because p is a limit point of S, some sequence $\{z_n\}$ in S satisfies the conditions,

$$\lim_{n\to\infty}\{z_n\} = p, \; z_n \neq p \text{ for all } n$$

NOTES

Thus, since f is continuous at p and since $f = 0$ on S,

$$a_0 = f(p) = \lim_{z \to p} f(z) = \lim_{n\to\infty}\{f(z_n)\} = \lim_{n\to\infty}\{0\} = 0$$

Define a new function,

$$f_1 : \Omega \to \mathrm{C},\ f_1(z) = \begin{cases} \dfrac{f(z)}{z-p} & \text{if} \quad z \neq p, \\ f'(p) & \text{if} \quad z \neq p. \end{cases}$$

Then, $f_1 = 0$ on $S - \{p\}$, so that in particular $f_1(z_n) = 0$ for all n. Also, f_1 is clearly analytic at all $z \neq p$ and f_1 is analytic at $z = p$ as well, because it has a power series representation at p,

$$f_1(z) = a_1 + a_2(z-p) + a_3(z-p)^2 + \ldots,\ z \in B$$

In fact, this power series agrees with $f_1(z)$ both for $z \neq p$ and for $z = p$. The previous argument that $a_0 = 0$ now applies to f_1 to show that $a_1 = 0$ as well, and the whole process repeats to show that $a_n = 0$ for all $n \geq 0$. Therefore, f is identically zero on the ball B. However, we want f to be identically zero on all of Ω. So let q be any point of Ω. Since Ω is connected and open in C, a little topology shows that it is path-connected and the connecting paths can be taken to be rectifiable. The general topological principle here is that connected and locally path-oriented implies path-connected, and in our context, the connecting paths can be taken to be rectifiable by metric properties of C.

Thus, some rectifiable path γ in the region Ω connects p to q. Since γ is compact, some ribbon about it lies in the region as well,

$$R = \bigcup_{z\in\gamma} B(z, \rho) \subset \Omega\,.$$

Form a chain of finitely many discs of radius ρ, with their centers spaced atmost distance ρ apart along γ, starting at p and ending at q. Each consecutive pair of discs overlaps on a set S having the center of the second disc as a limit point. Since f is identically zero on the first disc, the argument just given shows that it is identically zero on the second disc as well, and so on up to last disc. In particular $f(q) = 0$. Hence, the theorem is proved.

Corollary: An analytic function $f: \Omega \to C$ that is not identically zero has isolated zeros in any compact subset K of Ω and hence, only finitely many zeros in any such K. More generally, if f is not constant then on any compact subset K of Ω and for any value $a \in C$, f has only finite many a-points, i.e., points where f takes the value a.

The corollary holds because any infinite subset S of a compact subset K has a limit point in K by the Bolzano-Weierstrass theorem. So, if $f = a$ everywhere on S then $f = a$ identically on Ω.

Theorem 2.23: Let $f: \Omega \to C$ be an analytic function. Then either $|f|$ assumes no maximum on Ω or f is constant.

Proof: Let $|f|$ assumes a maximum at some point $c \in \Omega$, or

$|f(z)| \leq |f(c)|$ for all $z \in \Omega$

Let $B = B(c, r)$ where $r > 0$ lies in Ω. For any ρ satisfying $0 < \rho < r$, let γ_ρ be the circle about c of radius ρ.

Now,

$$|f(c)| = \left|\frac{1}{2\pi i}\int_{\gamma_\rho}\frac{f(z)dz}{z-c}\right| \leq \frac{1}{2\pi}\int_{\gamma_\rho}\frac{|f(z)||dz|}{\rho} \leq \frac{1}{2\pi\rho}\sup_{z\in\gamma_\rho}\{|f(z)|\}2\pi\rho = |f(c)|$$

This implies that all of the terms are equal. So, $|f| = |f(c)|$ on γ_ρ.

Since $\rho \in (0, r)$ is arbitrary, $|f| = |f(c)|$ on B. thus, by the uniqueness theorem, f is constant on Ω.

Corollary: If f: $\Omega \to C$ is analytic and K is a compact subset of Ω then $\max_{z\in K}\{|f(z)|\}$ is assumed on the boundary of K.

Theorem 2.24 (Liouville's): Let f: $C \to C$ be analytic and bounded. Then f is Constant.

Proof: The power series representation of f at 0 is valid for all of C,

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \text{ , } z \in C$$

Let M bound $|f|$. Cauchy's inequality says that for any $r > 0$ and any $n \in N$,

$$|a_n| < \frac{M}{r^n}$$

Since r can be arbitrarily large, this proves that $a_n = 0$ for $n \geq 1$.

i.e., $f(z) = a_0$ for all z.

Theorem 2.25 (Fundamental Theorem of Algebra): Let $p(z)$ be a non-constant polynomial with complete coefficients. Then p has a complete root.

Proof: Consider,

$$p(z) = z^n + \sum_{j=0}^{n-1} a_j z^j \text{ , } n \geq 1$$

Note that for all z such that $|z| \geq 1$,

$$\left|\sum_{j=0}^{n-1} a_j z^j\right| \leq C|z|^{n-1} \text{ where } C = \sum_{j=0}^{n-1}|a_j|$$

It follows that for all z such that $|z| > C+1$,

$$|p(z)| \geq |z|^n - C|z|^{n-1} > |z|^{n-1} \geq 1 .$$

Now suppose that $p(z)$ has no complex root. Then the function $f(z) = 1/p(z)$ is entire. The function $f(z)$ is bounded on the compact set $\overline{B(0, C+1)}$, and it satisfies $|f(z)| < 1$ for all z such that $|z| > C+1$.

Therefore, $f(z)$ is entire and bounded, making it constant by Liouville's theorem, and this makes the original polynomial $p(z)$ constant as well. The proof is complete by contraposition.

NOTES

2.15 ABEL'S THEOREM

Theorem 2.26 (Abel) Theorem: Suppose $g(x) = \sum_{n\geq 0} c_n x^n$ be a power series which converges for $|x|<1$. If $\sum_{n\geq 0} c_n$ converges then, $\lim_{x\to 1^-} g(x) = \sum_{n\geq 0} c_n$.

Or, if a power series converges at $x=1$ then its value at $x=1$ is the limit of its values at $x\to 1^-$, so a power series has built-in continuity in its behaviour.

Example 2.24: Let $g(x) = \sum_{n\geq 1} (-1)^{n-1} x^n / n$ for $|x|<1$.

Then for $|x|<1$, $g(x) = \log(1+x)$.

The series $g(1)$ converges since it is alternating. Hence, by Abel's theorem,

$$\begin{aligned} g(1) &= \lim_{x\to 1^-} g(x) \\ &= \lim_{x\to 1^-} \log(1+x) \\ &= \log 2 \end{aligned}$$

Since the logarithm is a continuous function.

Example 2.25: Let $g(x) = \sum_{n\geq 0} \frac{(-1)^{n-1}(2n)!}{2^{2n} n!^2 (2n-1)} x^n$ for $|x|<1$. Hence, $g(x) = \sqrt{1+x}$.

The series $g(1)$ is absolutely convergent. So, by the continuity of $\sqrt{1+x}$ and Abel's theorem,

$$g(1) = \lim_{x\to 1^-} g(x) = \lim_{x\to 1^-} \sqrt{1+x} = \sqrt{2}$$

We know that the series converges at $x = 1$. Abel's theorem states that if a power series converges on $(-1, 1)$ and also at $x = 1$, then its value at $x = 1$ is determined by continuity from the left of 1.

Proof of Abel's Theorem: We have that $\sum_{n\geq 0} c_n x^n$ converges for $|x|<1$ and at $x = 1$. We need to prove that,

$$\lim_{x\to 1^-} \sum_{n\geq 0} c_n x^n = \sum_{n\geq 0} c_n$$

For $-1 < x < 1$ we work with the truncated sums $\sum_{n=0}^{N} c_n x^n$ and $\sum_{n=0}^{N} c_n$.

Put, $s_n = c_0 + c_1 + \ldots + c_n$ for $n \geq 0$.

Note that, $s_n - s_{n-1} = c_n$ for $n \geq 1$.

Then,

NOTES

$$\sum_{n=0}^{N} c_n x^n = c_0 + \sum_{n=1}^{N} x^n (s_n - s_{n-1})$$

$$= c_0 + \sum_{n=1}^{N} u_n (s_n - s_{n-1}), \text{ where } u_n = x^n$$

$$= c_0 + u_N s_N - u_1 s_0 - \sum_{n=1}^{N-1} s_n (u_{n+1} - u_n) \text{ by Summation by Parts}$$

$$= c_0 + x^N c_N - x c_0 - \sum_{n=1}^{N-1} s_n (x^{n+1} - x^n)$$

$$= (1-x) c_0 + x^N c_N + \sum_{n=1}^{N-1} s_n (x^n - x^{n+1})$$

$$= (1-x) c_0 + x^N c_N + \sum_{n=1}^{N-1} s_n (1-x) x^n$$

Or $$\sum_{n=0}^{N} c_n x^n = x^N c_N + (1-x) \sum_{n=0}^{N-1} s_n x^n \qquad ...(2.26)$$

By hypothesis, the left hand side of Equation (2.26) converges as $N \to \infty$. Also, $x^N c_N \to 0$ as $N \to \infty$, since $x^N \to 0$ and $c_N \to 0$, because the series $\sum_{n \geq 0} c_n$ converges, so its general term must tend to 0.

Therefore, the other term in Equation (2.26) converges as $N \to \infty$ and we get,

$$\sum_{n \geq 0} c_n x^n = (1-x) \sum_{n \geq 0} s_n x^n \qquad ... (2.27)$$

Suppose $s = \sum_{n \geq 0} c_n$. We have to show that $\sum_{n \geq 0} c_n x^n \to s$ as $x \to 1^-$.

Substract s from both sides of Equation (2.27) and write,

$$\sum_{n \geq 0} c_n x^n - s = (1-x) \sum_{n \geq 0} (s_n - s) x^n \qquad ... (2.28)$$

using the formula $(1-x) \sum_{n \geq 0} x^n = 1$. Now, we have to show that the right hand side of Equation (2.28) tends to 0 as $x \to 1^-$.

By assumption, $s_n \to s$ as $n \to \infty$. Choose a positive number ε. For all large n, say $n \geq M$, $|s_n - s| \leq \varepsilon$. Now, break up the right hand side of Equation (2.28) into two sums,

$$\sum_{n \geq 0} c_n x^n - s = (1-x) \sum_{n=0}^{M-1} (s_n - s)\, x^n + (1-x) \sum_{n \geq M} (s_n - s) x^n$$

And estimate as below:

$$\left| \sum_{n \geq 0} c_n x^n - s \right| \leq |1-x| \sum_{n=0}^{M-1} |s_n - s| |x|^n + |1-x| \sum_{n \geq M} |s_n - s| \, |x|^n$$

NOTES

$$\leq |1-x| \sum_{n=0}^{M-1} |s_n - s| |x|^n + |1-x| \sum_{n \geq M} \varepsilon |x|^n$$

$$= |1-x| \sum_{n=0}^{M-1} |s_n - s| \, |x|^n + |1-x| \, \varepsilon \frac{|x|^M}{1-|x|}$$

$$< |1-x| \sum_{n=0}^{M-1} |s_n - s| |x|^n + |1-x| \varepsilon \frac{1}{1-|x|}$$

Taking $0 < x < 1$, $|1-x| = 1-x$, this upper bound becomes

$$\left| \sum_{n \geq 0} c_n x^n - s \right| < |1-x| \sum_{n=0}^{M-1} |s_n - s| + \varepsilon \qquad \ldots (2.29)$$

When $x \to 1^-$, the first term on the right hand side of Equation (2.29) tends to 0 on account of the $1-x$ there. When x is close enough to 1, we can make the first term on the right side at most ε, so

$$\left| \sum_{n \geq 0} c_n x^n - s \right| \leq \varepsilon + \varepsilon = 2\varepsilon \qquad \ldots (2.30)$$

as $x \to 1^-$. Since ε is an arbitrary positive number, the left side of Equation (2.30) must go to zero as $x \to 1^-$.

Corollary: Let a power series $\sum_{n \geq 0} c_n x^n$ converges for $|x| < r$. If the series converges at r or $-r$ then there is the limit of the values of the series as x tends to the endpoint from inside the interval or,

(a) If $\sum_{n \geq 0} c_n r^n$ converges then,

$$\lim_{x \to r-} \sum_{n \geq 0} c_n x^n = \sum_{n \geq 0} c_n r^n$$

(b) If $\sum_{n \geq 0} c_n (-r)^n$ converges then,

$$\lim_{x \to -r+} \sum_{n \geq 0} c_n x^n = \sum_{n \geq 0} c_n (-r)^n$$

Proof: (a) Let $a_n = c_n r^n$ and $g(x) = \sum_{n \geq 0} a_n x^n = \sum_{n \geq 0} c_n (rx)^n$ for $|x| < 1$. This series converges at $x = 1$. So, by Abel's theorem,

$$\sum_{n \geq 0} a_n = \lim_{x \to 1-} \sum_{n \geq o} a_n x^n = \lim_{x \to 1-} \sum_{n \geq 0} c_n r^n x^n = \lim_{x \to r-} \sum_{n \geq 0} c_n x^n$$

Where the limit changed from $x \to 1^-$ to $x \to r^-$ in the last Equation (by replacing x with x/r). Since $a_n = c_n r^n$, the left hand side is $\sum_{n \geq 0} a_n = \sum_{n \geq 0} c_n r^n$

(b) The argument is similar, using $a_n = c_n (-r)^n$.

2.16 TAUBER'S THEOREM

NOTES

Theorem 2.27: The converse of Abel's theorem is false in general. If f is given by,

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \ , -r < x < r$$

Then the limit $f(r^-)$ may exist but the series $\sum a_n \ r^n$ may fail to converge. For example, if $a_n = (-1)^n$, then

$$f(x) = \frac{1}{1+x}, -1 < x < 1$$

And $f(x) \to 1/2$ as $x \to 1^-$. However, $\sum (-1)^n$ is not convergent. Tauber showed that the converse of Abel's theorem can be obtained by imposing additional condition on the coefficients a_n.

Theorem 2.28 (Tauber): Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $-1 < x < 1$ and suppose that $\lim_{n\to\infty} n\, a_n = 0$. If $f(x) \to S$ as $x \to 1^-$, then $\sum_{n=0}^{\infty} a_n$ converges and has the sum S.

Proof: Let $\sigma_n = \sum_{h=0}^{n} k\,|a_k|$. Then $\sigma_n \to 0$ as $n \to \infty$.

Also, $\lim_{n\to\infty} f(x_n) = S$ if $x_n = 1 - 1/n$.

Therefore to each $\varepsilon > 0$, we can choose an integer N such that $n \geq N$ implies,

$$|(fx_n)| - S| < \frac{\varepsilon}{3}$$

$$\sigma_n < \frac{\varepsilon}{3}$$

$$n|a_n| < \frac{\varepsilon}{3}$$

Let $S_n = \sum_{h=0}^{n} a_k$. Then for $-1 < x < 1$, we have

$$S_n - S = f(x) - S + \sum_{k=0}^{n} a_k (1 - x^k) - \sum_{h=n+1}^{\infty} a_k \ x^k$$

Let $x \in (0, 1)$. Then,

$(1 - x^k) = (1 - x)\,(1 + x + - + x^{k-1}) \leq k(r - x)$, for each k.

Therefore, if $n \geq N$ and $0 < x < 1$, we have

$$|S_n - S| \leq |f(x) - S| + (1 - x)\sum_{h=0}^{n} k|a_k| + \frac{\in}{3n(1-x)}$$

Putting $x = x_n = 1-1/n$, we find that

$$|S_n - S| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

NOTES

Which completes the proof.

Check Your Progress

14. What is power series?
15. State the uniqueness theorem of power series.
16. Define the Abel's theorem.
17. Give the statement of Tauber's theorem.

2.17 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A sequence is a function whose domain is the set of natural numbers.
2. A sequence $\{x_n\}$ is said to be bounded above if all its terms are less than or equal to a real number, i.e., there exists $K \in \mathbb{R}$ such that $x_n \leq K$ for all $n \in \mathbb{N}$.

   A sequence $\{x_n\}$ is said to be bounded below if all its terms are greater than or equal to a real number, i.e., there exists $K \in \mathbb{R}$ such that $x_n \geq k$ for all $n \in \mathbb{N}$.
3. A sequence is said to be bounded if it is bounded both above and below, i.e., if there exist $K, k \in \mathbb{R}$ such that $k \leq x_n \leq K$ for all $n \in N$.
4. An expression of the form,

   $u_1 + u_2 + u_3 + ... + u_n + ...$

   in which every term is followed by another according to some definite rule is called a series.
5. A sequence $\{f_n\}$ of functions is said to converge pointwise on a set S to a limit function f, if for each $x \in S$ and for each $\varepsilon > 0$ there exists an **N** (depending on x and ε) such that,

   $|f_n(x) - f(x)| < \varepsilon$, for all $n > N$.

   A sequence of real valued functions $\langle f_n \rangle$ defined on a set S is said to converge uniformly to a real valued function f on S if for $\varepsilon > 0 \; \exists \; m \in \mathbf{N}$ such that,

   $$|f_n(x) - f(x)| < \varepsilon \;\forall\; n \geq m \text{ and } x \in S$$
6. A sequence of real valued functions $< f_n >$ defined on a set S converges uniformly on S iff to each given $\varepsilon > 0 \; \exists \; m \in \mathbf{N}$ such that,

   $$|f_{n+p}(x) - f_n(x)| < \varepsilon \;\forall\; n \geq m, p \geq 0 \text{ and } x \in S.$$
7. A series $\Sigma f_n(x)$ converges uniformly (and absolutely) on a set S if there exists a convergent series $\Sigma\, M_n$ of non-negative terms M_n such that, $|f_n(x)| \leq M_n \;\forall\; x \in S$ and $n \in \mathbf{N}$.

8. If $\Sigma f_n(x)$ converges uniformly on a set S and $\langle g_n(x)\rangle$ be monotonic and uniformly bounded on S, then the series $\Sigma f_n(x)\, g_n(x)$ converges uniformly on S.
9. Let X be a metric space. If the functions $f_n: X \to C, g_n: X \to R, n \in N$ satisfy the following:

   (a) $F_n(x) = \sum_{m=1}^{n} f_m(x)$ is bounded uniformly in n and x.

   (b) $g_{n+1} \le g_n(x)$ for all $x \in X$ and $n \in N$.

   (c) $\{g_n(x)\}_{n\in N}$ converges uniformly to zero on X.

   Then $\sum_{n=1}^{\infty} f_n(x) g_n(x)$ converges uniformly on X.
10. A sequence of continuous functions may converge to a continuous function, although the convergence is not uniform.
11. Assume that each term of $\{f_n\}$ is a real valued function having a finite derivative at each point of an open interval (a, b). Assume that for at least one point x_0 in (a, b), the sequence $\{f_n(x_0)\}$ converges. Assume further that there exists a function g such that $f_n' \to g$ uniformly on (a, b).
12. Let I be a closed and bounded interval. Suppose $f: I \to R$ is a continuous function. Then for each $\varepsilon > 0$, there exists a polynomial function $p_\varepsilon: I \to R$ such that,

    $| f(x) - p_\varepsilon(x) | < \varepsilon$ for all x in I, or equivalently,

    sup $\{| f(x) - p_\varepsilon(x) | \ : x \in I\} < \varepsilon$.
13. Let $f: [0,1] \to R$ be a function. Then for each integer $n \ge 0$, we define the Bernstein polynomial of degree n associated with f to be,

    $$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$$
14. An infinite series, in the ascending integral powers of a real variable x, of the form,

    $a_0 + a_1 x + a_2 x^2 + ... + a_n x^n + ...,$

    Where the coefficients $a_0, a_1, a_2,...$ are constant, independent of x, is called a real power series.

15 Let $\Omega \subset C$ be a region and consider two analytic functions $f, g : \Omega \to C$.

Let S be a subset of Ω that has a limit point $p \in \Omega$ which need not lie in S. Suppose that $f = g$ on S. Then $f = g$.

16. Suppose $g(x) = \sum_{n\ge 0} c_n x^n$ be a power series which converges for $|x|<1$.

    If $\sum_{n\ge 0} c_n$ converges then, $\lim_{x\to 1} g(x) = \sum_{n\ge 0} c_n$.

Or, if a power series converges at $x=1$ then its value at $x=1$ is the limit of its values at $x \to 1^-$, so a power series has built-in continuity in its behaviour.

NOTES

17. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $-1 < x < 1$ and suppose that $\lim_{n\to\infty} n\, a_n = 0$.

If $f(x) \to S$ as $x \to 1^-$, then $\sum_{n=0}^{\infty} a_n$ converges and has the sum S.

2.18 SUMMARY

- A sequence is a function whose domain is the set of natural numbers.
- If the codomain is the set $\mathbb{R}$ of real numbers, it is called a real sequence.
- If the codomain is the set $\mathbb{C}$ of complex numbers then it is called a complex sequence.
- If it is a set of polynomials then it is a sequence of polynomials.
- A sequence $\{x_n\}$ is called convergent if it converges to a limit l.
- A sequence, which converges to zero, is called a null sequence.
- A sequence may or may not converge.
- If a sequence is convergent, it converges to a unique limit.
- Every convergent sequence is always bounded.
- A monotone increasing sequence bounded above is always convergent and converges to its Least Upper Bound (LUB).
- A monotone decreasing sequence bounded below is always convergent and converges to its Greatest Lower Bound (GLB).
- Every constant sequence is convergent.
- A sequence $\{x_n\}$ is said to diverge to $+\infty$ if for every large $G > 0$, there exists $n_0 \in \mathbb{N}$ such that $x_n \geq G$ for all $n \geq n_0$.
- An expression of the form $u_1 + u_2 + u_3 + \ldots + u_n + \ldots$ in which every term is followed by another according to some definite rule is called a series.
- If the number of terms is not finite, it is called an infinite series.
- The nature of a series is determined by the nature of the sequence of its nth partial sum.
- A sequence $\{f_n\}$ of functions is said to converge pointwise on a set S to a limit function f, if for each $x \in S$ and for each $\varepsilon > 0$ there exists an N (depending on x and ε) such that, $|f_n(x) - f(x)| < \varepsilon$, for all $n > N$.
- A sequence of real valued functions $<f_n>$ defined on a set S is said to converge uniformly to a real valued function f on S if for $\varepsilon > 0$ $\exists m \in N$ such that, $|f_n(x) - f(x)| < \varepsilon$ $\forall n \geq m$ and $x \in S$.

- A sequence of real valued functions $<f_n>$ defined on a set S converges uniformly on S iff to each given $\varepsilon > 0\ \exists m \in N$ such that,

  $|f_{n+p}(x) - f_n(x)| < \varepsilon\ \forall\, n \geq m,\ p \geq 0$ and $x \in S$.

  This is Cauchy's general principle of uniform convergence.
- A series $\sum f_n(x)$ converges uniformly on a set S if there exists a convergent series $\sum M_n$ of non-negative terms M_n such that $|f_n(x)| \leq M_n\ \forall\, x \in S$ and $n \in N$.
- If $\sum f_n(x)$ converges uniformly on a set Σ and $<g_n(x)>$ be monotonic and uniformly bounded on S, then the series $\sum f_n(x)\, g_n(x)$ converges uniformly on S. This is Abel's test for uniform convergence.
- Let X be a metric space. If the functions $f_n: X \to C$, $g_n: X \to R$, $n \in N$ satisfy the following:

  (a) $F_n(x) = \sum_{m=1}^{n} f_m(x)$ is bounded uniformly in n and x.

  (b) $g_{n+1} \leq g_n(x)$ for all $x \in X$ and $n \in N$.

  (c) $\{g_n(x)\}_{n \in N}$ converges uniformly to zero on X.

  Then $\sum_{n=1}^{\infty} f_n(x) g_n(x)$ converges uniformly on X. This is Dirichlet's test for uniform convergence.
- A sequence of continuous functions may converge to a continuous function, although the convergence is not uniform.
- $g_n(x) = \int_a^x f_n(t)\, d\alpha(t)$ if $x \in [a, b]$, $n = 1, 2, \ldots$, then we have

  (a) $f \in R(\alpha)$ on $[a, b]$.

  (b) $g_n \to g$ uniformly on $[a, b]$, where $g(x) = \int_a^x f(t)\, d\alpha(t)$.
- Let $f: [0, 1] \to \mathbb{R}$ be a function. Then for each integer $n \geq 0$, we define the Bernstein polynomial of degree n associated with f to be,

  $B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$.
- An infinite series, in the ascending integral powers of a real variable x, of the form $a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots$, where the coefficients $a_0, a_1, a_2, \ldots$ are constant and independent of x, is called a real power series.
- Any infinite subset S of a compact subset K has a limit point in K by the Bolzano-Weierstrass theorem.
- Abel's theorem states that if a power series converges on $(-1, 1)$ and also at $x = 1$, then its value at $x = 1$ is determined by continuity from the left of 1.

NOTES

- Let $f(x)=\sum_{n=0}^{\infty} a_n x^n$ for $-1< x < 1$ and suppose that $\lim_{n\to\infty} n\, a_n = 0$. If $f(x) \to S$ as $x \to 1^-$, then $\sum_{n=0}^{\infty} a_n x^n$ converges and has the sum S.

2.19 KEY TERMS

- **Sequence:** A function whose domain is the set of natural numbers.
- **Series:** An expression of the form $u_1 + u_2 + u_3 + \ldots + u_n + \ldots$ in which every term is followed by another according to some definite rule.
- **Pointwise convergence:** A sequence $\{f_n\}$ of functions is said to converge pointwise on a set S to a limit function f, if for each $x \in S$ and for each $\varepsilon > 0$ there exists an N (depending on x and ε) such that, $|f_n(x) - f(x)| < \varepsilon$, for all $n > N$.
- **Uniform convergence:** A sequence of real valued functions $<f_n>$ defined on a set S is said to converge uniformly to a real valued function f on S if for $\varepsilon > 0\ \exists m \in N$ such that $|f_n(x) - f(x)| < \varepsilon\ \forall n \geq m$ and $x \in S$.

2.20 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is sequence?
2. Define the term monotone sequence.
3. What is convergent sequence?
4. Name the two important series.
5. Test the convergence of the series $\frac{1}{3}+\left(\frac{2}{5}\right)^2+\left(\frac{3}{7}\right)^2+\ldots+\left(\frac{n}{2n+1}\right)^n+\ldots$ using Cauchy's root test.
6. Which convergence is a local property and which one is a global property?
7. Define uniformly bounded set.
8. What is the use of Weierstrass's M-test?
9. Write an application of Abel's test for uniform convergence.
10. What are the drawbacks of Dirichlet's test for uniform convergence?
11. What is the relation between uniform convergence and continuity?
12. State the significance uniform convergence.
13. What is the use of uniform convergence and differentiation?
14. Write the three cases of Weierstrass approximation theorem.

15. Give an example of power series.
16. What is the significance of uniqueness theorem for power series?
17. State the difference between Abel's theorem and Tauber's theorem.

Long-Answer Questions

1. Explain the concept of sequence with reference to bounded and unbounded sequences, monotone sequence and convergent sequence giving theorems, proofs and relevant examples.
2. Discuss about the Cauchy's criterion of convergence, divergent sequences and oscillatory sequences giving theorems, proofs and significant examples.
3. Show that the sequence $\{x_n\}$ is convergent when,

$$x_n = 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}$$

4. Briefly discuss about the series in detail with the help of appropriate examples.
5. What are the necessary and sufficient conditions for the convergence of an infinite series? Explain giving relevant examples.
6. Explain in detail the tests used for convergence series and divergence series giving definitions and appropriate examples of each test.
7. Prove that the following given series,

$$1 + \frac{1}{2} \cdot \frac{1}{3} + \frac{1.3}{2.4} \cdot \frac{1}{5} + \frac{1.3.5}{2.4.6} \cdot \frac{1}{7} + \dots \quad \text{to } \infty \text{ converges.}$$

8. Test the convergence of the following given series,

$$\sum \left\{ \frac{2.4.6.8 \dots 2n}{3.5.7.9 \dots (2n+1)} \right\}^2$$

9. Test the convergence or divergence of the following given series,

$$1^p + \left(\frac{1}{2}\right)^p + \left(\frac{1.3}{2.4}\right)^p + \left(\frac{1.3.5}{2.4.6}\right)^p + \dots$$

10. Briefly explain the pointwise and uniform convergence giving definitions and significant examples.
11. Prove that the pointwise convergence is a local property whereas the uniform convergence is a global property.
12. State and prove the Cauchy's criterion for uniform convergence giving definitions and related significant examples.
13. Prove that the sequence $\langle x^n \rangle$ converges on [0, 1] to the function f defined by $f(x) = 0$ when $x \neq 1$ and $f(1) = 1$. The convergence is not uniform on [0, 1].

NOTES

14. Discuss the Cauchy's general principle of convergence for series giving relevant examples.
15. Briefly explain the necessary conditions for convergent series and the sufficient conditions for divergent series giving appropriate examples.
16. Explain the Weierstrass's M-test giving the theorem, proof and appropriate examples.
17. Discuss in detail the Abel's test and the Dirichlet's test for uniform convergence with the help of theorems, proofs and relevant examples.
18. Describe the significance of uniform convergence and continuity giving the theorem and examples.
19. Briefly explain the uniform convergence and Riemann-Stieltjes integration with the help of theorems, proofs and appropriate examples.
20. Brief a note on uniform convergence and differentiation giving theorems and relevant examples.
21. Explain in detail the Weierstrass approximation theorems giving their proofs.
22. Discuss the concept of the power series and prove the uniqueness theorem for the power series.
23. If r (> 0) be the radius of convergence of the power series $a_0 + a_1x + a_2x^2 + \ldots$, in $(-r, r)$. Then prove that it converges uniformly and absolutely on every close interval contained in $(-r, r)$.
24. Consider that $f: \Omega \to C$ be an analytic function. Prove that either $|f|$ assumes no maximum on Ω or f is a constant.
25. State and prove Abel's theorem and Tauber's theorem giving appropriate examples.

2.21 FURTHER READING

Rudin, Walter. 2017. *Real and Complex Analysis*, Third Edition. Noida: McGraw-Hill Education.

Gupta, S. L. and Nisha Rani. 2004. *Fundamental Real Analysis*, Fourth Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, First Edition. Cambridge (U.K.): Cambridge University Press.

Bartle, Robert G. and Donald R. Sherbert. 2014. *Introduction to Real Analysis*, Fourth Edition. New York: Wiley.

Trench, William F. 2002. *Introduction to Real Analysis*, London: Pearson.

Loeb, Peter A. 2016. *Real Analysis*, Basel (Switzerland): Birkhäuser.

Royden, Halsey. 2015. *Real Analysis*, Fourth Edition. Noida: Pearson Education India.

NOTES

UNIT 3 FUNCTIONS OF SEVERAL VARIABLES AND HIGHER ORDER DIFFERENTIALS

Structure

3.0 INTRODUCTION

In real analysis, a function of several variables or multivariate function is a function with more than one argument, with all arguments being real variables. This concept extends the idea of a function of a real variable to several variables. The "Input" variables take real values, while the "Output", also called the "Value of the Function", may be real or complex. However, the complex valued functions may be easily reduced to the simple real valued functions on further analysis, by considering the real and imaginary parts of the complex function.

The domain of a function of n variables is the subset of $\mathbf{R}^n$ for which the function is defined. As usual, the domain of a function of several real variables is supposed to contain an open subset of $\mathbf{R}^n$. Some functions are defined for all real values of the variables such that they are everywhere defined, but some other functions are defined only if the value of the variable are taken in a subset X of $\mathbf{R}^n$, the domain of the function, which is always supposed to contain an open subset of $\mathbf{R}^n$.

The function $f(x, y)$ is a function of a single variable x when y is constant. Then the derivative of $f(x, y)$ (when exists) is called the partial derivative of $f(x, y)$ with respect to x. It is denoted by $f_x(x, y)$ or $\frac{\partial f}{\partial x}$.

A homogeneous function is one with multiplicative scaling behaviour. If all its arguments are multiplied by a factor, then its value is multiplied by some power of this factor. For example, a homogeneous real valued function of two variables x and y is a real valued function that satisfies the condition $f(rx, ry) = r^k f(x, y)$ for some constant and all real numbers r he constant k is called the degree of homogeneity.

The Jacobian determinant is used when making a change of variables when evaluating a multiple integral of a function over a region within its domain. To accommodate for the change of coordinates the magnitude of the Jacobian determinant arises as a multiplicative factor within the integral. This is because the n-dimensional dV element is in general a parallelepiped in the new coordinate system, and the n-volume of a parallelepiped is the determinant of its edge vectors.

In this unit, you will study about the functions of several variables, linear transformations, derivatives in an open subset of R^n, partial derivatives, higher order differentials, Taylor's theorem, explicit and implicit functions, implicit function theorem and inverse function theorem, change of variables, extreme values of explicit and stationary values of implicit functions, Lagrange's multipliers method, differential forms, Stoke's theorem and Jacobian and its properties.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the functions of several variables
- Define what linear transformations are
- Find the derivatives in open subset of R^n
- Evaluate the partial derivatives
- Explain the higher order differentials
- State about the Taylor's theorem
- Distinguish between explicit and implicit functions
- Define the concept of change of variables
- Describe Lagrange's multipliers method
- Discuss about the differential forms and Stoke's theorem
- Explain Jacobian and its properties

3.2 FUNCTIONS OF SEVERAL VARIABLES

A variable z is said to be a function of two variables x and y if for each pair (x, y) corresponds a value of z. This is expressed by $z = f(x, y)$. For example, if $z = x^2 + y^2$, then z is a function of x and y.

If $z = f(x, y)$, then z is a dependent variable and x, y are independent variables. The function $z = f(x, y)$ is called a *single-valued function* if only one value of z is corresponded by each pair (x, y) for which the function is defined. If there is more than one value of z, the function is called a *multi-valued function*.

The set of values (points) (x, y) for which a function is defined, is called the domain of definition or simply domain of the function.

For example, if $z = \sqrt{4-(x^2+y^2)}$, the domain for which z is real consists of the set of points (x, y) such that $x^2 + y^2 \geq 4$, i.e., the set of points inside and on a circle in the xy-plane having its centre at (0, 0) and radius 2.

Note: If z is a function of n independent variables $x_1, x_2, ..., x_n$, then we write $z = f(x_1, x_2, ..., x_n)$. For example, if $u = x^2 + y^2 + z^2$, then u is a function of three variables x, y and z, i.e., $u = f(x, y, z)$.

Limit and Continuity of a Function of Two Variables

Definition: Let $f(x, y)$ be a function of two variables defined in the region R and (a, b) be a point in R. The function $f(x, y)$ is said to have a limit l as (x, y) tends to (a, b) if for every small positive number ε, there exists a positive number δ such that,

$$| f(x, y) - l | < \varepsilon \text{ for } 0 < | x - a | < \delta \text{ and } 0 < | y - b | < \delta$$

Or $$| f(x, y) - l | < \varepsilon \text{ for } 0 < (x - a)^2 + (y - b)^2 < \delta^2$$

In this case, we write $\lim\limits_{(x,y)\to(a,b)} f(x, y) = l$ or $\lim\limits_{\substack{x\to a\\ y\to b}} f(x, y) = l$.

This is also called the *double limit* or the simultaneous limit of $f(x, y)$ as (x, y) tends to (a, b).

Repeated Limit: Let $f(x, y)$ be a function of two variables defined in the region R and (a, b) be a point in R. Let $\lim\limits_{x\to a} f(x, y)$ exist and it is a function of y, say $g(y)$. If $\lim\limits_{y\to b} g(y)$ exists is equal to l, then l is called the repeated or iterated limit of $f(x, y)$ as $x \to a$ and then $y \to b$ and we express this by,

$$\lim_{y\to b} \lim_{x\to a} f(x, y) = l$$

By changing the order of limits, we similarly define

$$\lim_{x \to a} \lim_{y \to b} f(x, y) = l$$

NOTES

Note: If the two repeated limits exist and are equal, then the double limit may or may not exist. Conversely the double limit may exist but the repeated limits may not exist but however if the repeated limits exist, they must be equal.

Continuity: Let $f(x, y)$ be a function of two variables defined in the region R and (a, b) be a point in R. The function $f(x, y)$ is said to be *continuous* at (a, b) if $\lim_{(x, y) \to (a, b)} f(x, y) = f(a, b)$.

If $f(x, y)$ is continuous for every point (a, b) in R, then we say $f(x, y)$ is continuous in R.

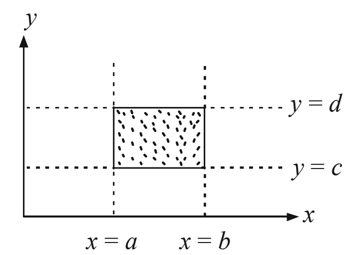
Analytical Definition: Let $f(x, y)$ be a function of two variables defined in the region R and (a, b) be a point in R. The function $f(x, y)$ is said to be continuous at (a, b) if for every small positive number ε, there exists a positive number δ depending on ε such that,

$$|f(x, y) - f(a, b)| < \varepsilon \text{ for } |x - a| < \delta, |y - b| < \delta \text{ (for Square Region)}$$

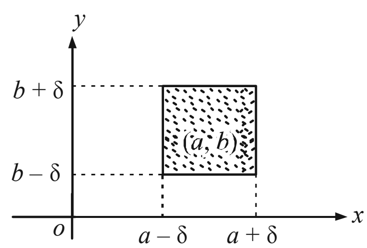
or $|f(x, y) - f(a, b)| < \varepsilon$ for $(x - a)^2 + (y - b)^2 < \delta^2$ (for Circular Region)

Region: If any two points of a set S can be joined by a path consisting of a finite number of broken line segments all of whose points belong to S, then S is called a connected set. A *region* is a connected open set. The following regions are generally used.

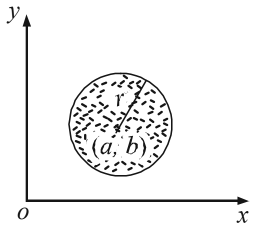
(*i*) **Rectangular Region:** A rectangular region R is a set of points (x, y) which satisfy the inequalities of the form $a \le x \le b$, $c \le y \le d$



(*ii*) **Square Region:** A square region R is a set of points (x, y) which satisfy the inequalities of the form $a - h \le x \le a + h$, $b - h \le x < b + h$



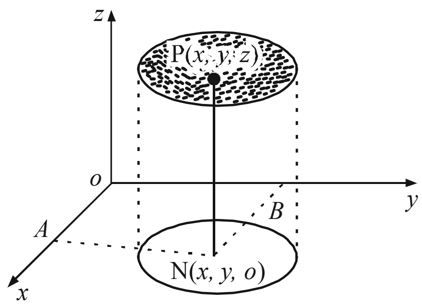
(*iii*) **Circular Region:** A circular region R is a set of points (x, y) which satisfy the inequalities of the form $(x - a)^2 + (y - b)^2 \leq r^2$



Note: A region is said to be closed or open according as the boundary points do or do not belong to the region. For example, the region $\{(x, y) \in \mathbb{R}^2;\ (x - a)^2 + (y - b)^2 < r^2\}$ is an open region, but $\{(x, y) \in \mathbb{R}^2,\ (x - a^2)^2 + (y - b)^2 \leq r^2\}$ is closed.

Geometrical Representation of Functions of Two Variables

The function of one variable represents a curve in the two dimensional plane. A function of two variables $z = f(x, y)$ represents a surface in the three dimensional space.



Let (x, y, z) be the coordinate of the point P. So, each point (x, y) in R corresponds to another point (x, y, z) in space which describes a surface. Hence, $P(x, y, z)$ is a point on the surface $z = f(x, y)$. This surface is the geometrical representation of the function.

Theorems on Limit and Continuity

Theorem 3.1: Let $f(x, y)$ and $g(x, y)$ be two functions defined in the same region R such that $\lim_{(x,y)\to(a,b)} f(x, y) = l$ and $\lim_{(x,y)\to(a,b)} g(x, y) = m$ then,

(*i*) $\lim_{(x,y)\to(a,b)} \{A\, f(x, y) \pm B\, g(x, y)\} = Al \pm Bm$ where A and B are constants.

(*ii*) $\lim_{(x,y)\to(a,b)} \{f(x, y) \cdot g(x, y)\} = l \cdot m$

(*iii*) $\lim_{(x,y)\to(a,b)} \dfrac{f(x,y)}{g(x,y)} = \dfrac{l}{m}$ provided $m \neq 0$.

NOTES

Theorem 3.2: If $f(x, y)$ and $g(x, y)$ are continuous at (a, b), then

(i) $f(x, y) \pm g(x, y)$ is continuous at (a, b).

(ii) $f(x, y) \cdot g(x, y)$ is continuous at (a, b).

(iii) $\dfrac{f(x,y)}{g(x,y)}$ is also continuous at (a, b) provided $g(a, b) \neq 0$.

Example 3.1: Prove that $\lim\limits_{(x,y)\to(0,0)} xy \dfrac{x^2 - y^2}{x^2 + y^2} = 0.$

Solution: Let $f(x, y) = xy \dfrac{x^2 - y^2}{x^2 + y^2}$.

We shall show that for any given small positive number ε, then we can find $\delta > 0$ such that $| f(x, y) - 0| < \varepsilon$ for $0 < x^2 + y^2 < \delta^2$.

Now, $| f(x, y) - 0 | = \left| xy\dfrac{x^2 - y^2}{x^2 + y^2} - 0 \right| = | x | | y | \left|\dfrac{x^2 - y^2}{x^2 + y^2}\right| < | x | | y |$

$(\because | x^2 - y^2 | < | x^2 + y^2|)$

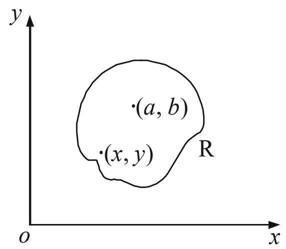
$\therefore$ $|f(x, y) - 0| < \varepsilon$ whenever $| x | | y | < \varepsilon$, i.e., whenever,
$x^2 + y^2 < \varepsilon$,

Since, $| x | < \sqrt{x^2 + y^2}$ and $| y | < \sqrt{x^2 + y^2}$.

Thus, if we take $\delta = \sqrt{\varepsilon}$, then $|f(x, y) - 0| < \varepsilon$ whenever $0 < x^2 + y^2 < \delta^2$.

Hence, the given double limit exists, i.e., $\lim\limits_{(x,y)\to(0,0)} xy \dfrac{x^2 - y^2}{x^2 + y^2} = 0.$

Note: Let (a, b) be an interior point of the region R and (x, y) be any point of R. The point (x, y) varies over the region R and approaches the point (a, b) along any specified curve in R but for the existence of the double limit, the limiting value must be unique along whatever path (x, y) approaches (a, b). If the limiting values are different for different approaches to the point (a, b) along different curves in R, then the limit does not exist.



Example 3.2: Show that the repeated limits exist and

$$\lim_{y\to 0}\lim_{x\to 0} \frac{xy}{x^2 + y^2} = \lim_{x\to 0}\lim_{y\to 0} \frac{xy}{x^2 + y^2}$$

But the double limit $\lim\limits_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ does not exist.

NOTES

Solution: Here, $\lim\limits_{y\to 0}\lim\limits_{x\to 0} \frac{xy}{x^2+y^2} = \lim\limits_{y\to 0} \frac{0\cdot y}{0+y^2} = 0$

And $\lim\limits_{x\to 0}\lim\limits_{y\to 0} \frac{xy}{x^2+y^2} = \lim\limits_{x\to 0} \frac{0\cdot x}{x^2+0} = 0$

$\therefore \lim\limits_{y\to 0}\lim\limits_{x\to 0} \frac{xy}{x^2+y^2} = \lim\limits_{x\to 0}\lim\limits_{y\to 0} \frac{xy}{x^2+y^2}$

2nd Part: Let $y = mx$, then $x \to 0$ implies $(x, y) \to (0, 0)$.

$$\therefore \lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{(x\to 0)} \frac{x\, mx}{x^2+m^2x^2} = \lim_{x\to 0} \frac{m}{1+m^2} = \frac{m}{1+m^2}$$

Which depends on m and varies as m varies.

Thus for different values of m, we get different limits. Hence the double limit does not exist.

3.3 LINEAR TRANSFORMATIONS

Let V and U be two vector spaces over the same field F, then a mapping $T : V \to U$ is called a **homomorphism** or a linear transformation if,

$$T(x + y) = T(x) + T(y) \quad \text{for all } x, y \in V$$

$$T(\alpha x) = \alpha T(x), \quad \alpha \in F$$

One can combine the two conditions to get a single condition,

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \text{ where } x, y \in V;\ \alpha, \beta \in F$$

It is easy to see that both are equivalent. If a homomorphism happens to be one-one onto also, we call it an *isomorphism*, and say that the two spaces are isomorphic. (Notation: $V \cong U$).

Example 3.3: Identity map $I : V \to V$, such that,

$$I(v) = v$$

And, the zero map $O : V \to V$, such that,

$$O(v) = 0$$

are clearly linear transformations.

Example 3.4: For a field F, consider the vector spaces F^2 and F^3. Define a map $T : F^3 \to F^2$, by

$$T(\alpha, \beta, \gamma) = (\alpha, \beta)$$

Then T is a linear transformation as,

for any $x, y \in F^3$, if $\quad x = (\alpha_1, \beta_1, \gamma_1)$

$$y = (\alpha_2, \beta_2, \gamma_2)$$

NOTES

Then, $T(x+y) = T(\alpha_1+\alpha_2,\ \beta_1+\beta_2,\ \gamma_1+\gamma_2) = (\alpha_1+\alpha_2,\ \beta_1+\beta_2)$

$= (\alpha_1, \beta_1) + (\alpha_2, \beta_2) = T(x) + T(y)$

And $T(\alpha x) = T(\alpha\ (\alpha_1, \beta_1, \gamma_1)) = T((\alpha\alpha_1, \alpha\beta_1, \alpha\gamma_1)$

$= (\alpha\alpha_1, \alpha\beta_1) = \alpha(\alpha_1, \beta_1) = \alpha T(x)$

Example 3.5: Let V be the vector space of all polynomials in x over a field F. Define,

$T : V \to V$, such that,

$$T(f(x)) = \frac{d}{dx} f(x)$$

Then $$T(f+g) = \frac{d}{dx}(f+g) = \frac{d}{dx} f + \frac{d}{dx} g = T(f) + T(g)$$

$$T(\alpha f) = \frac{d}{dx}(\alpha f) = \alpha \frac{d}{dx} f = \alpha T(f)$$

Shows that T is a linear transformation.

In fact, if $\theta : V \to V$ be defined such that

$$\theta(f) = \int_0^x f(t)\ dt$$

Then θ will also be a linear transformation.

Example 3.6: Consider the mapping,

$T : \mathbf{R}^3 \to \mathbf{R}$, such that,

$$T(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$$

Then T is not a linear transformation.

Consider, for instance,

$T((1, 0, 0) + (1, 0, 0)) = T(2, 0, 0) = 4$

$T(1, 0, 0) + T(1, 0, 0) = 1 + 1 = 2.$

3.4 DERIVATIVES IN AN OPEN SUBSET OF R^n

For arriving at a definition of the derivative of a function whose domain is R^n or an open subset of R^n, let us consider the case n=1. Let f be a real function with domain $(a, b)\subset R$ and $c\in(a, b)$ then $f'(c)$ is defined to be the real number,

$$\lim_{h\to 0} \frac{f(c+h)-f(c)}{h}$$, provided that the limit exists.

Let $$E_c(h) = \frac{f(c+h)-f(c)}{h} - f'(c),\ \text{if}\ h \neq 0 \qquad \ldots(3.1)$$

And $E_c(h) = 0$, if $h = 0$. Then we have,

$$hE_c(h) = f(c+h) - f(c) - hf'(c)$$

Or the equation which holds also for $h = 0$,

$$f(c+h) = f(c) + hf'(c) + hE_c(h) \qquad \ldots (3.2)$$

The Equation (3.2) is called first order Taylor formula for approximating $f(c+h)-f(c)$ by $f'(c)h$. The error committed is $hE_c(h)$. Also from Equation (3.1) we get $E_c(h)\to 0$ as $h\to 0$. Following are the two properties of Equation (3.2):

1. The quantity $f'(c)h$ is a linear function, i.e., if we write $T_c(h) = f'(c)h$, then,

   $T_c(ah_1 + bh_2) = aT_c(h_1) + bT_c(h_2)$
2. The error term $hE_c(h)$ is of smaller order than h as $h\to 0$.

   Now the total derivative of a function f from R^n to R^m will be defined in such a way that it preserves the above two properties.

Definition: Let f:$S \to R^m$ be a function defined on a set S in R^n with values in R^m. Let c be an interior point of S and let $B(c;r)$ be an n-ball lying in S. Let V be a point in R^n with $\|v\| < r$, so that $c+v \in B(c;r)$. Then the function f is said to be differentiable at c if there exists a linear function $T_c: R^n \to R^m$ such that,

$$f(c+v) = f(c) + T_c(v) + \|v\| E_c(v), \quad \text{where } E_c(v) \to 0 \text{ as } v \to 0 \qquad \ldots (3.3)$$

Note: Equation (3.3) is called a first order Taylor formula. The linear function T_c is called the total derivative of f at c. We also write Equation (3.3) as,

$$f(c+v) = f(c) + T_c(v) + 0(\|v\|) \text{ as } v \to 0$$

Theorem 3.3: Let f is differentiable at c with total derivative T_c. Then the directional derivative $f'(c;v)$ exists for every u in R^n and we have,

$$T_c(u) = f'(c; v)$$

Proof: Let f is differentiable at c. Then we have,

$$f(c+v) = f(c) + T_c(v) + \|v\| E_c(v), \text{ where } T_c \text{ is linear and } E_c(v) \to 0 \text{ as } v \to 0 \qquad \ldots (3.4)$$

If $v = 0$, then $f'(c;0) = 0 = T_c(u)$. Now assume that $v \neq 0$.

Then taking $v = hu$ in Equation (3.4) we get,

$$f(c+hu) = f(c) + T_c(hu) + \|hu\| E_c(hu)$$

$$\Rightarrow f(c+hu) - f(c) = hT_c(u) + |h| \|u\| E_c(hu)$$

$$\Rightarrow \frac{f(c+hu)}{h} - f(c) = hT_c(u) + \frac{|h|}{h} \|u\| E_c(hu)$$

Taking $\lim\limits_{h\to 0}$ on both sides we get,

$$\Rightarrow \frac{f(c+hu)}{h} - f(c) = hT_c(u) + \frac{|h|}{h} \|u\| E_c(hu)$$

That is, $f'(c;0) = T_c(u)$

Theorem 3.4: If f is differentiable at c, then f is continuous at c.

Proof: f is differentiable at c then we have,

NOTES

$$f(c+v) = f(c) + T_c(v) + \|v\| E_c(v) \qquad \ldots (3.5)$$

Where T_c is linear and $E_c(v) \to 0$ as $v \to 0$.

As $v \to 0$ in Equation (3.5) we get,

$$\lim_{v\to 0} f(c+v) = f(c) + \lim_{v\to 0} T_c(v) + \lim_{v\to 0} \|v\| E_c(v)$$

$$= f(c) + T_c(0) + 0$$

Or $\lim_{v\to 0} f(c+v) = f(c)$, since T_c is linear, $T_c(v) = 0$

That is, f is continuous at c.

For example, total derivative of linear function is the function itself. Let f be a linear function differentiable at c. Then $f(c+v) = f(c) + T_c(v) + \|v\| E_c(v)$ where $E_c(v) \to 0$ as $v \to 0$.

$$\Rightarrow f(c) + f(v) = f(c) + T_c(v) + \|v\| E_c(v)$$

$$\Rightarrow f(v) = T_c(v) + \|v\| E_c(v)$$

Theorem 3.5: Let $f: S \to R^m$ be differentiable at an interior point c of S, where $S \subseteq R^n$. If $v = v_1 u_1 + v_2 u_2 + \ldots + v_n u_n$, where $U_1, \ldots, U_n$ are the unit coordinate vectors in R^n, then $f(c)(v) = \sum_{k=1}^{n} v_k D_k f(c)$. If f is real valued (i.e., $m=1$) we have $f'(c)(v) = \nabla f(c) \,.\, (v)$ which is the dot product of v with the vector $\nabla f(c) = (D_1 f(c), \ldots, (D_n f(c))$.

Proof: Given $v = v_1 u_1 + v_2 u_2 + \ldots + v_n u_n$, where $u_1, \ldots, u_n$ are the unit coordinate vectors and v_i, $i = 1, \ldots, n$ are reals. Since $f'(c)$ is linear, so

$$\begin{aligned} f'(c)(v) &= f'(c)(v_1 u_1 + v_2 u_2 + \ldots + v_n u_n) \\ &= v_1 f'(c)(u_1) + v_2 f'(c)(u_2) + \ldots + v_n f'(c)(u_n) \\ &= v_1 f'(c;u_1) + v_2 f'(c;u_2) + \ldots + v_n f'(c:u_n) \\ &= v_1 D_1 f(c) + v_2 D_2 f(c) + \ldots + v_n D_n f(c) \\ &= \sum_{k=1}^{n} v_k D_k f(c) \end{aligned}$$

In particular for $m = 1$, we get

$$f'(c)(v) = \sum_{k=1}^{n} v_k D_k f(c)$$

$$= (D_1 f(c), D_2 f(c), ..., D_n f(c)).(v_1, v_2, ..., v_n)$$

$$= \nabla f(c).(v), \quad \text{where } \nabla f(c) = (D_1 f(c), ..., D_n f(c))$$

Note: The vector $\nabla f(c)$ is known as the gradient vector of f at c.

NOTES

Check Your Progress

1. Distinguish between single-valued and multi-valued functions.
2. Define linear transformation.
3. Give a necessary condition for a function to be continuous at a point.

3.5 PARTIAL DERIVATIVES

Let $f(x, y)$ be a function of two independent variables x and y, defined in the region R.

The function $f(x, y)$ is a function of a single variable x when y is constant. Then the derivative of $f(x, y)$ (when exists) is called the partial derivative of $f(x, y)$ with respect to x. It is denoted by $f_x(x, y)$ or $\frac{\partial f}{\partial x}$.

$$\therefore \quad \frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h} \qquad ... (3.6)$$

The function $f(x, y)$ is a function of a single variable y when x is constant. Then the derivative of $f(x, y)$ (when exists) is called the partial derivative of $f(x, y)$ with respect to y. It is denoted by $f_y(x, y)$ or $\frac{\partial f}{\partial y}$

$$\therefore \quad \frac{\partial f}{\partial y} = \lim_{k \to 0} \frac{f(x, y+k) - f(x, y)}{k} \qquad ... (3.7)$$

Notes:

1. When Equations (3.6) and (3.7) exist at (a, b) in R, then they are denoted by $f_x(a, b)$ and $f_y(a, b)$.
2. $f_x(x, y)$ and $f_y(x, y)$ are also functions of x and y.
3. The function $f(x, y)$ is derivable means that both the partial derivatives $f_x(x, y)$ and $f_y(x, y)$ exist.
4. Let $f(x, y, z)$ be a function of three independent variables x, y and z defined in R. Then $f(x, y, z)$ has three partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$.

$$\therefore f_x(x, y, z) = \lim_{h \to 0} \frac{f(x+h, y, z) - f(x, y, z)}{h}$$

$$f_y(x, y, z) = \lim_{k \to 0} \frac{f(x, y+k, z) - f(x, y, z)}{k}$$

And $f_z(x, y, z) = \lim\limits_{w \to 0} \dfrac{f(x, y, z+w) - f(x, y, z)}{w}$.

Example 3.7: Find from the definition of partial derivative $f_x(1, 1)$ and $f_y(1, 1)$ where $f(x, y) = \dfrac{x+y}{x+y+1}$.

NOTES

Solution: Here,

$$f_x(1, 1) = \lim_{h \to 0} \frac{f(1+h, 1) - f(1, 1)}{h} = \lim_{h \to 0} \left[\frac{1+h+1}{1+h+1+1} - \frac{2}{3} \right] \frac{1}{h}$$

$$= \lim_{h \to 0} \left[\frac{h+2}{h+3} - \frac{2}{3} \right] \frac{1}{h} = \lim_{h \to 0} \frac{3h+6-2h-6}{3(h+3)} \cdot \frac{1}{h}$$

$$= \lim_{h \to 0} \frac{h}{3(h+3)} \cdot \frac{1}{h} = \lim_{h \to 0} \frac{1}{3(h+3)} = \frac{1}{9}$$

$$\text{And } f_y(1, 1) = \lim_{k \to 0} \frac{f(1, 1+k) - f(1, 1)}{k} = \lim_{k \to 0} \left[\frac{1+1+k}{1+1+k+1} - \frac{2}{3} \right] \frac{1}{k}$$

$$= \lim_{k \to 0} \left[\frac{k+2}{k+3} - \frac{2}{3} \right] \frac{1}{k} = \lim_{k \to 0} \left[\frac{3k+6-2k-6}{3(k+3)} \right] \frac{1}{k}$$

$$= \lim_{k \to 0} \frac{k}{3(k+3)} \cdot \frac{1}{k} = \lim_{k \to 0} \frac{1}{3(k+3)} = \frac{1}{9}.$$

Example 3.8: Let $f(x, y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$.

Find $f_x(0, 0)$ and $f_y(0, 0)$.

Solution: Here $f_x(0, 0) = \lim\limits_{h \to 0} \dfrac{f(0+h, 0) - f(0, 0)}{h} = \lim\limits_{h \to 0} \dfrac{0-0}{h} = 0$

$$\text{And } f_y(0, 0) = \lim_{k \to 0} \frac{f(0, 0+k) - f(0, 0)}{k} = \lim_{k \to 0} \frac{\frac{0}{k^2} - 0}{k} = \lim_{k \to 0} \frac{0-0}{k} = 0.$$

Example 3.9: Find from definition given that $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ at (x, y) where $f(x, y) = x^2 + 2xy + y^2$.

Solution: Here,

$$f_x(x, y) = \frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}$$

$$= \lim_{h \to 0} \frac{(x+h)^2 + 2(x+h)y + y^2 - (x^2 + 2xy + y^2)}{h}$$

$$= \lim_{h \to 0} \frac{x^2 + 2hx + h^2 + 2xy + 2hy + y^2 - x^2 - 2xy - y^2}{h}$$

$$= \lim_{h \to 0} \frac{h^2 + 2hx + 2hy}{h} = \lim_{h \to 0} h + 2x + 2y = 2x + 2y$$

Similarly, $f_y(x, y) = \dfrac{\partial f}{\partial y} = 2x + 2y$.

Example 3.10: If $f(x, y) = \dfrac{x-y}{x+y}$, then find $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ at $(2, -1)$ from the given definition.

Solution: Now,

$$f_x(2,-1) = \left[\frac{\partial f}{\partial x}\right]_{(2,-1)} = \lim_{h\to 0}\frac{f(2+h,-1)-f(2,-1)}{h}$$

$$= \lim_{h\to 0}\frac{\dfrac{2+h+1}{2+h-1}-\dfrac{3}{2-1}}{h} = \lim_{h\to 0}\left[\frac{3+h}{h+1}-3\right]\frac{1}{h}$$

$$= \lim_{h\to 0}\frac{3+h-3h-3}{h(h+1)} = \lim_{h\to 0}\frac{-2h}{h(h+1)} = \lim_{h\to 0}\frac{-2}{h+1} = -2$$

And $$f_y(2,-1) = \left[\frac{\partial f}{\partial y}\right]_{(2,-1)} = \lim_{k\to 0}\frac{f(2,-1+k)-f(2,-1)}{k}$$

$$= \lim_{k\to 0}\left[\frac{2+1-k}{2-1+k}-3\right]\frac{1}{k} = \lim_{k\to 0}\left[\frac{3-k}{k+1}\right]\frac{1}{k}$$

$$= \lim_{k\to 0}\frac{3-k-3k-3}{k(k+1)}$$

$$= \lim_{k\to 0}\frac{-4k}{k(k+1)} = \lim_{k\to 0}\frac{-4}{k+1} = -4.$$

Partial Derivative of Higher Order

Let $z = f(x, y)$ be a function of two independent variables x and y. Then the partial derivatives $f_x(x, y)$ and $f_y(x, y)$ are also functions of x and y. The partial derivatives of $f_x(x, y)$ and $f_y(x, y)$ are called second order partial derivatives of $f(x, y)$. The partial derivatives of $f_x(x, y)$ with respect to x and y are given by,

$$\frac{\partial}{\partial x}(f_x(x,y)) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

$$= \frac{\partial^2 f}{\partial x^2} = f_{xx} = \lim_{h\to 0}\frac{f_x(x+h,y)-f_x(x,y)}{h}$$

And $$\frac{\partial}{\partial y}[f_x(x,y)] = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\partial x} = f_{yx} = \lim_{k\to 0}\frac{f_x(x,y+k)-f_x(x,y)}{k}$$

The partial derivatives of $f_y(x, y)$ with respect to x and y are given by,

$$\frac{\partial}{\partial x}[f_y(x,y)] = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\partial y} = f_{xy} = \lim_{h\to 0}\frac{f_y(x+h,y)-f_y(x,y)}{h}$$

And $$\frac{\partial}{\partial y}[f_y(x,y)] = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy} = \lim_{k\to 0}\frac{f_y(x,y+k)-f_y(x,y)}{k}$$

The four second order partial derivatives of $f(x, y)$ are f_{xx}, f_{yy}, f_{xy} and f_{yx}.

NOTES

The third order partial derivatives of $f(x, y)$ are given by,

f_{xxx}, f_{yxx}, f_{xyx}, f_{yyx}, f_{xxy}, f_{yxy}, f_{xyy} and f_{yyy} where,

$$f_{yyx} = \frac{\partial}{\partial y}\left[\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right] \text{ first with respect to } x,$$

Then with respect to y and then again with respect to y.

In general, $$\frac{\partial^n f}{\partial x^n} = \frac{\partial}{\partial x}\left(\frac{\partial^{n-1} f}{\partial x^{n-1}}\right) = \frac{\partial^2}{\partial x^2}\left(\frac{\partial^{n-2} f}{\partial x^{n-2}}\right)$$

And $$\frac{\partial}{\partial x}\left(\frac{\partial^{n-1} f}{\partial y^{n-1}}\right) = \frac{\partial^n f}{\partial x\, \partial y^{n-1}}.$$

In general $f_{xy} \neq f_{yx}$ but if f_x and f_y exist in some neighbourhood of (a, b) and if they are differentiable at (a, b), then $f_{xy} = f_{yx}$ at (a, b) which is known as *Young's theorem*. Another set of sufficient conditions for the above equality has been given by Schwarz as follows:

Theorem 3.6 (Schwarz's): Let $f(x, y)$ be a function defined in the region R of the xy-plane and (a, b) be any point in R such that:

(i) $\dfrac{\partial f}{\partial x}$ exists in some neighbourhood of (a, b)

(ii) $\dfrac{\partial^2 f}{\partial x \partial y}$ is continuous at (a, b),

Then $\dfrac{\partial^2 f}{\partial y\, \partial y}$ exists at (a, b) and $f_{xy}(a, b) = f_{yx}(a, b)$.

Example 3.11: The function $f(x, y)$ is defined by

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$$

Prove that $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

Solution: Here,

$$f_{xy}(0, 0) = \lim_{h \to 0} \frac{f_y(0 + h, 0) - f_y(0, 0)}{h}$$

$$= \lim_{h \to 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} \qquad \dots (1)$$

And, $$f_{yx}(0, 0) = \lim_{k \to 0} \frac{f_x(0, 0 + k) - f_x(0, 0)}{k}$$

$$= \lim_{k \to 0} \frac{f_x(0, k) - f_x(0, 0)}{k} \qquad \dots (2)$$

Also, $$f_y(h, 0) = \lim_{k \to 0} \frac{f(h, 0 + k) - f(h, 0)}{k} = \lim_{k \to 0} \frac{f(h, k) - f(h, 0)}{k}$$

NOTES

$$= \lim_{k\to 0}\left[hk\frac{h^2-k^2}{h^2+k^2} - h\cdot 0\frac{h^2-0^2}{h^2+0}\right]\frac{1}{k}$$

$$= \lim_{k\to 0}\frac{h(h^2-k^2)}{h^2+k^2} = \frac{h^3}{h^2} = h \qquad ...(3)$$

And, $$f_y(0, 0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{0-0}{k} = 0 \qquad ... (4)$$

Hence, from Equations (1), (3) and (4), we get

$$f_{xy}(0, 0) = \lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h} = \lim_{h\to 0}\frac{h-0}{h} = 1$$

Again, $$f_x(0, k) = \lim_{h\to 0}\frac{f(0+h, k)-f(0,k)}{h} = \lim_{h\to 0}\frac{f(h,k)-f(0,k)}{h}$$

$$= \lim_{h\to 0}\left[hk\frac{h^2-k^2}{h^2+k^2} - 0\,.\,k\,.\frac{0^2-k^2}{0^2+k^2}\right]\frac{1}{h}$$

$$= \lim_{h\to 0}\frac{k(h^2-k^2)}{h^2+k^2} = \frac{k(-k^2)}{k^2} = -k \qquad ... (5)$$

And, $$f_x(0, 0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}$$

$$= \lim_{h\to 0}\frac{0-0}{h} = 0 \qquad ... (6)$$

Hence, from Equations (2), (5) and (6), we get

$$f_{yx}(0, 0) = \lim_{k\to 0}\frac{f_x(0,k)-f_x(0,0)}{k} = \lim_{k\to 0}\frac{-k-0}{k} = -1$$

Since, $f_{xy}(0, 0) = 1$ and $f_{yx}(0, 0) = -1$, hence $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

Example 3.12: Show that for the function,

$$f(x, y) = \begin{cases}\dfrac{x^2y^2}{x^2+y^2} & \text{for } (x,y)\neq(0,0)\\ 0 & \text{for } (x,y)=(0,0)\end{cases}$$

the equality $f_{xy}(0, 0) = f_{yx}(0, 0)$ holds.

Solution: Here, $f_{xy}(0, 0)$

$$= \lim_{h\to 0}\frac{f_y(0+h,0)-f_y(0,0)}{h} = \lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h}$$

Also, $$f_y(h, 0) = \lim_{k\to 0}\frac{f(h,0+k)-f(h,0)}{k} = \lim_{k\to 0}\frac{f(h,k)-f(h,0)}{k}$$

$$= \lim_{k\to 0}\left[\frac{h^2k^2}{h^2+k^2} - \frac{h^2\,.\,0}{h^2+0^2}\right]\frac{1}{k} = \lim_{k\to 0}\frac{h^2k}{h^2+k^2} = \frac{0}{h^2} = 0$$

And, $$f_y(0, 0) = \lim_{k\to 0}\frac{f(0,0+k)-f(0,0)}{k} = \lim_{h\to 0}\left[\frac{0\,.\,k^2}{0+k^2} - 0\right]\frac{1}{k}$$

NOTES

$$= \lim_{k \to 0} \frac{0-0}{k} = 0$$

Hence, $f_{xy}(0, 0) = \lim_{h \to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = \lim_{h \to 0} \frac{0-0}{h} = 0$

Again, $f_{yx}(0, 0) = \lim_{k \to 0} \frac{f_x(0,0+k) - f_x(0,0)}{k} = \lim_{k \to 0} \frac{f_x(0,k) - f_x(0,0)}{k}$

Also, $f_x(0, k) = \lim_{h \to 0} \frac{f(h,k) - f(0,k)}{h} = \lim_{h \to 0} \left[\frac{h^2 k^2}{h^2 + k^2} - \frac{0 . k^2}{0^2 + k^2} \right] \frac{1}{h}$

$$= \lim_{h \to 0} \frac{hk^2}{h^2 + k^2} = \frac{0 . k^2}{0^2 + k^2} = 0$$

And, $f_x(0, 0) = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h \to 0} \left[\frac{h^2 . 0}{h^2 + 0^2} - 0 \right]$

$$= \lim_{h \to 0} \frac{0-0}{h} = 0$$

Hence, $f_{yx}(0, 0) = \lim_{k \to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = \lim_{k \to 0} \frac{0-0}{k} = 0$

$\therefore \quad f_{xy}(0, 0) = 0 = f_{yx}(0, 0)$

Homogeneous Functions

A function $f(x, y)$ is said to be homogeneous of degree n in the variables x and y if it can be expressed in the form $x^n \phi(y/x)$ or in the form $y^n \phi(x/y)$.

Alternatively, a function $f(x, y)$ is said to be homogeneous of degree n in the variables x and y if $f(tx, ty) = t^n f(x, y)$ for all values of t independent of x and y.

Generalized Definition: A function $f(x, y, z, ...)$ is said to be homogeneous function of degree n in the variables x, y, z ..., if $f(tx, ty, tz, ...) = t^n f(x, y, z, ...)$ for all values of t independent of $x, y, z, ...$.

Illustrations: (*i*) The function $f(x, y) = x^2 + y^2$ is homogeneous of degree two because $f(x, y)$ can be written in the form $f(x, y) = x^2 \left(1 + \frac{y^2}{x^2}\right) = x^2 \phi(y/x)$ where $\phi(y/x) = 1 + y^2/x^2$.

(*ii*) The function $f(x, y) = \frac{xy}{x^4 + y^4}$ is a homogeneous function of degree (– 2) because $f(tx, ty) = \frac{t^2 xy}{t^4 (x^2 + y^2)} = t^{-2} \frac{xy}{x^2 + y^4} = t^{-2} f(x, y)$.

(*iii*) The function $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ is a homogeneous function of degree 1 because $f(tx, ty, tz) = \sqrt{t^2 x^2 + t^2 y^2 + t^2 z^2} = t\sqrt{x^2 + y^2 + z^2} = t^1 f(x, y, z)$.

(*iv*) The function $f(x, y, z) = \dfrac{x+y+z}{\sqrt{x}+\sqrt{y}+\sqrt{z}}$ is a homogeneous function of degree 1/2 because $f(tx, ty, tz) = \dfrac{tx+ty+tz}{\sqrt{tx}+\sqrt{ty}+\sqrt{tz}} = \dfrac{t}{t^{1/2}} \dfrac{x+y+z}{\sqrt{x}+\sqrt{y}+\sqrt{z}} = t^{1/2} f(x, y, z)$.

Euler's Theorem on Homogeneous Functions

Theorem 3.7 (For Two Variables): Let $f(x, y)$ be a homogeneous function of two variables x and y of degree n having continuous partial derivative, then

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = n\, f(x, y).$$

Proof: Since $f(x, y)$ is a homogeneous function of degree n, then we write $f(x, y) = x^n \phi(y/x)$.

$$\therefore \quad \frac{\partial f}{\partial x} = n\, x^{n-1}\, \phi(y/x) + x^n\, \phi'(y/x)\, (-\, y/x^2)$$

$$= nx^{n-1}\, \phi(y/x) - x^{n-2}\, y\, \phi'(y/x)$$

$$\text{And} \quad \frac{\partial f}{\partial y} = x^n\, \phi'(y/x) \left(\frac{1}{x}\right) = x^{n-1}\, \phi'(y/x)$$

$$\therefore\ x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = n\, x^n\, \phi(y/x) - x^{n-1}\, \phi'(y/x) + x^{n-1}\, \phi'(y/x)$$

$$= n\, x^n \phi(y/x) = n\, f(x, y)$$

Euler's Theorem on Homogeneous Functions of More than Two Independent Variables

Let $u = f(x_1, x_2, ..., x_n)$ be a homogeneous function of n independent variables $x_1, x_2, ..., x_n$ of degree k having continuous partial derivatives, then

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \ldots + x_n\frac{\partial f}{\partial x_n} = k\, f(x_1, x_2, ..., x_n).$$

Example 3.13: Verify Euler's theorem for $f(x, y) = x^3 + y^3$.

Solution: Here, the function $f(x, y)$ is a homogeneous function of degree three.

$\therefore \quad f_x = 3x^2$ and $f_y = 3y^2$

$$\therefore \quad x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = x \cdot 3x^2 + y \cdot 3y^2 = 3(x^3 + y^3) = 3\, f(x, y)$$

Hence, Euler's theorem is verified.

Example 3.14: Verify the Euler's theorem for,

$$f(x, y, z) = 3x^2yz + 5xy^2z + 5z^4.$$

Solution: Here, $f(tx, ty, tz) = 3t^2x^2\; ty\; tz + 5\; tx\; t^2y^2\; tz + 5\; t^4z^4$

$$= t^4\, (3x^2yz + 5xy^2z + 5z^4)$$

$$= t^4 f(x, y, z)$$

NOTES

$\therefore\ f(x, y, z)$ is a homogeneous function of Degree 4.

Now $\frac{\partial f}{\partial x} = 6xyz + 5y^2z \quad \frac{\partial f}{\partial y} = 3x^2z + 10xyz$

$\frac{\partial f}{\partial z} = 3x^2y + 5xy^2 + 20z^3$

$$\begin{aligned}\therefore\ x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} \\ &= x(6xyz + 5y^2z) + y(3x^2z + 10xyz) + z(3x^2y + 5xy^2 + 20z^3) \\ &= 6x^2yz + 5xy^2z + 3x^2yz + 10xy^2z + 3x^2yz + 5xy^2z + 20z^4 \\ &= 12x^2yz + 20xy^2z + 20z^4 \\ &= 4(3x^2yz + 5xy^2z + 5z^4) \\ &= 4f(x, y, z)\end{aligned}$$

Hence, Euler's theorem is verified.

Harmonic Function

A function $f(x, y)$ is said to be a harmonic function if $\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$ (which is known as Laplace Equation).

Similarly, a function $f(x, y, z)$ is harmonic if,

$\nabla^2 f = 0$, i.e., if $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0.$

Example 3.15: Show that the function $f(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$ is a harmonic function.

Solution: Now

$$\frac{\partial f}{\partial x} = -\frac{x}{(x^2 + y^2 + z^2)^{3/2}}$$

And
$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= -\frac{(x^2 + y^2 + z^2)^{3/2} - x\frac{3}{2}(x^2 + y^2 + z^2)^{1/2}2x}{(x^2 + y^2 + z^2)^3} \\ &= -\frac{(x^2 + y^2 + z^2)^{1/2}[x^2 + y^2 + z^2 - 3x^2]}{(x^2 + y^2 + z^2)^3} \\ &= \frac{2x^2 - y^2 - z^2}{(x^2 + y^2 + z^2)^{5/2}}\end{aligned}$$

Similarly, $\frac{\partial^2 f}{\partial y^2} = \frac{2y^2 - x^2 - z^2}{(x^2 + y^2 + z^2)^{5/2}}$ and $\frac{\partial^2 f}{\partial z^2} = \frac{2z^2 - x^2 - y^2}{(x^2 + y^2 + z^2)^{5/2}}$

$\therefore\ \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$

$$= \frac{2x^2 - y^2 - z^2}{(x^2+y^2+z^2)^{5/2}} + \frac{2y^2 - x^2 - z^2}{(x^2+y^2+z^2)^{5/2}} + \frac{2z^2 - x^2 - y^2}{(x^2+y^2+z^2)^{5/2}}$$

$$= \frac{2x^2 - y^2 - z^2 + 2y^2 - x^2 - z^2 + 2z^2 - x^2 - y^2}{(x^2+y^2+z^2)^{5/2}} = 0.$$

Hence, $f(x, y, z)$ is a harmonic function in x, y, z.

Total Differential

Let $z = f(x, y)$ be a function of two independent variables x and y and f_x, f_y exist at (x, y), then $dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ is called the total differential of z.

Example 3.16: Find the total differential of $z = x^2 + xy + y^2$.

Solution: Now, $\frac{\partial z}{\partial x} = 2x + y$ and $\frac{\partial z}{\partial y} = x + 2y$.

Then the total differential of z is given by,

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy$$

$$= (2x + y)dx + (x + 2y)dy.$$

Chain Rules for Functions of Two or More Variables

(*i*) Let $z = f(x, y)$ be a function of two variables x and y where $x = \phi(t)$, $y = \psi(t)$ (assume that f, ϕ and ψ are differentiable functions) then z is a function of t only and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

(*ii*) Let $z = f(x, y)$ be a function of two variables x and y where $x = \phi(u, v)$ and $y = \psi(u, v)$ (assume that f, ϕ and ψ are differentiable functions), then z is a function of u and v and

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial u} \qquad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial v}$$

(*iii*) Let $z = f(u_1, u_2, ..., u_n)$ be a function of n variables $u_1, u_2, ..., u_n$ where $u_1 = \phi_1(x_1, x_2, ..., x_p)$, $u_2 = \phi_2(x_1, x_2, ..., x_p)$, ..., $u_n = \phi_n(x_1, x_2, ..., x_p)$ (assume that $f_\phi, \phi_1, \phi_2, ..., \phi_n$ are differentiable functions), then z is a function of $x_1, x_2, ..., x_p$ and

$$\frac{\partial z}{\partial x_i} = \frac{\partial z}{\partial u_1}\frac{\partial u_1}{\partial x_i} + \frac{\partial z}{\partial u_2}\frac{\partial u_2}{\partial x_i} + ... + \frac{\partial z}{\partial u_n}\frac{\partial u_n}{\partial x_i} \quad (i = 1, 2, ..., p)$$

$$= \sum_{r=1}^{n} \frac{\partial z}{\partial u_r}\frac{\partial u_r}{\partial x_i} \text{ where } i = 1, 2, ..., p.$$

The above results are known as chain rules.

NOTES

3.6 HIGHER ORDER DIFFERENTIALS

Let $f(x)$ be a derivable function of x in a given interval. Then its derivative $f'(x)$ is also a function of x. This function $f'(x)$ may have a derivative in a certain interval. This derivative is called the second order derivative of $f(x)$ and it is denoted by $f''(x)$. Similarly, the derivative of the 2nd order derivative is called the third derivative and so on.

The nth order derivative of $f(x)$ with respect to x is denoted by y_n or $f^{(n)}(x)$ or $\frac{d^n y}{dx^n}$ or $y^{(n)}$ or $\frac{d^n}{dx^n}\{f(x)\}$ or $D^n f(x)$ whre $D \equiv \frac{d}{dx}$ and $y = f(x)$.

The n^{th} Derivative of Some Functions

1. $y = e^{ax}$ where a is a constant.

Now $y_1 = ae^{ax}$

$$y_2 = aae^{ax} = a^2e^{ax}$$

$$y_3 = a^2ae^{ax} = a^3e^{ax}$$

....

....

$$y_n = a^n e^{ax}$$

Note: If $y = e^{ax+b}$, $y_n = a^n e^{ax+b}$

2. $y = \frac{1}{x}$

Here $y_1 = -\frac{1}{x^2} = (-1)^1 \frac{1}{x^{1+1}}$

$$y_2 = (-1)\left(\frac{-2}{x^3}\right) = (-1)^2 \frac{\lfloor 2}{x^{2+1}}$$

$$y_3 = (-1)^2 \cdot \frac{2.(-3)}{x^4} = \frac{(-1)^3 \lfloor 3}{x^{3+1}}$$

....

....

$$y_n = \frac{(-1)^n \lfloor n}{(x)^{n+1}}$$

Corollary: For $y = \frac{1}{x \pm a}$, $y_n = \frac{(-1)^n \lfloor n}{(x \pm a)^{n+1}}$.

3. $y = \log x$

Here, $y_1 = \frac{1}{x}$, $y_2 = -\frac{1}{x^2} = \frac{(-1)1}{x^{1+1}}$

$$y_3 = (-1)\frac{(-2)}{x^3} = (-1)^2 \frac{\lfloor 2}{x^{2+1}}$$

$$y_4 = (-1)^2 \cdot \frac{2 \cdot (-3)}{x^4} = (-1)^3 \frac{\lfloor 3}{x^{3+1}}$$

....

....

$$y_n = \frac{(-1)^{n-1} \lfloor n-1}{x^n}$$

Corollary: For $y = \log (x \pm a), y_n = \dfrac{(-1)^{n-1} \lfloor n-1}{(x \pm a)^n}$.

4. $y = \dfrac{1}{x^2 - a^2}$

Now $y = \dfrac{1}{x^2 - a^2} = \dfrac{1}{(x+a)(x-a)} = \left[\dfrac{1}{x-a} - \dfrac{1}{x+a}\right] \dfrac{1}{2a}$

$\therefore \quad y_n = \dfrac{1}{2a}\left[\dfrac{(-1)^n \lfloor n}{(x-a)^{n+1}} - \dfrac{(-1)^n \lfloor n}{(x+a)^{n+1}}\right]$ [by corollary of 2]

$$= \frac{1}{2a} (-1)^n \lfloor n \left[\frac{1}{(x-a)^{n+1}} - \frac{1}{(x+a)^{n+1}}\right]$$

Let $x = r \cosh \theta$ and $a = r \sinh \theta$, then $r^2 = x^2 - a^2$

And $\theta = \sinh^{-1} (a/r)$

$\therefore \quad x - a = r (\cosh \theta - \sinh \theta)$

$$= r\left[\left(\frac{e^{\theta} + e^{-\theta}}{2}\right) - \left(\frac{e^{\theta} - e^{-\theta}}{2}\right)\right] = re^{-\theta}$$

And $x + a = r (\cosh \theta + \sinh \theta)$

$$= r\left[\left(\frac{e^{\theta} + e^{-\theta}}{2}\right) - \left(\frac{e^{\theta} - e^{-\theta}}{2}\right)\right] = re^{\theta}$$

$\therefore \quad y_n = \dfrac{1}{2a} (-1)^n \lfloor n \left[\dfrac{1}{r^{n+1} e^{(n+1)\theta}} - \dfrac{1}{r^{n+1} e^{(n+1)\theta}}\right]$

$$= \frac{(-1)^n \lfloor n}{2a\, r^{n+1}} \left[e^{(n+1)\theta} - e^{-(n+1)\theta}\right] = \frac{(-1)^n \lfloor n}{a\, r^{n+1}} \sinh (n+1)\theta$$

Where $\theta = \sinh^{-1}\left(\dfrac{a}{r}\right)$

5. $y = \dfrac{1}{x^2 + a^2}$

Now, $y = \dfrac{1}{x^2 + a^2} = \dfrac{1}{x^2 - i^2 a^2} = \dfrac{1}{(x - ia)(x + ia)}$

$$= \frac{1}{2ai}\left[\frac{1}{x - ia} - \frac{1}{x + ia}\right]$$

$\therefore \quad y_n = \dfrac{1}{2ai} (-1)^n \lfloor n \left[\dfrac{1}{(x-ia)^{n+1}} - \dfrac{1}{(x+ia)^{n+1}}\right]$

Let $x = r\cos \theta$, $a = r \sin \theta$, then $r^2 = x^2 + a^2$ and $\theta = \tan^{-1}\left(\dfrac{a}{x}\right)$

$x - ia = r(\cos \theta - i \sin \theta) = re^{-i\theta}$ and $x + ia = r(\cos \theta + i \sin \theta) = re^{i\theta}$

NOTES

$$\therefore \quad y_n = \frac{1}{2ai}(-1)^n \lfloor n \left[\frac{1}{r^{n+1}\, \overline{e}^{(n+1)i\theta}} - \frac{1}{r^{n+1}\, e^{(n+1)i\theta}}\right]$$

$$= \frac{1}{2ai}\, \frac{(-1)^n \lfloor n}{r^{n+1}} \left[e^{(n+1)\theta i} - \overline{e}^{(n+1)\theta i}\right]$$

$$= \frac{(-1)^n \lfloor n}{a\, r^{n+1}} \sin(n+1)\,\theta \text{ where } \theta = \tan^{-1}\left(\frac{a}{x}\right)$$

6. $y = (ax+b)^m$ where m is any positive integer.

Now, $\quad y_1 = (ax+b)^{m-1}(m)a = ma\,(ax+b)^{m-1}$

$$y_2 = (ma)(m-1)a\,(ax+b)^{m-2}$$

$$y_3 = (ma)(m-1)a\,(m-2)a\,(ax+b)^{m-3}$$

$$= m(m-1)(m-2)\,a^3\,(ax+b)^{m-3}$$

....

....

$$y_n = m\,(m-1)(m-2)\,...\,(m-n+1)\,a^n\,(ax+b)^{m-n}$$

If m be a positive integer greater than n, then

$$y_n = m\,(m-1)(m-2)\,...\,(m-n+1)\,a^n\,(ax+b)^{m-n}$$

If m be a positive integer less than n, then

$$y_n = 0$$

If $m = n$, then $y_n = n(n-1)(n-2)\,...\,3 \cdot 2 \cdot 1\, a^n\,(ax+b)0$

$$= \lfloor n\; a^n.$$

7. $\quad y = \sin(ax+b)$

$$\therefore \quad y_1 = a\cos(ax+b) = a\sin\left(\frac{\pi}{2}+ax+b\right)$$

$$y_2 = -a^2\sin(ax+b) = a^2\sin\left(2\,.\frac{\pi}{2}+ax+b\right)$$

$$y_3 = -a^3\cos(ax+b) = a^3\sin\left(3\cdot\frac{\pi}{2}+ax+b\right)$$

....

....

$$y_n = a^n\sin\left(n\cdot\frac{\pi}{2}+ax+b\right)$$

Note: If $b = 0$, then $y_n = D^n\sin(ax) = a^n \sin\left(\frac{n\pi}{2}+ax\right)$

8. $\quad y = \cos(ax+b)$

$$\therefore \quad y_1 = -a\sin(ax+b) = a\cos\left(\frac{\pi}{2}+ax+b\right)$$

$$y_2 = -a^2\cos(ax+b) = a^2\cos\left(2\cdot\frac{\pi}{2}+ax+b\right)$$

$$y_3 = a^3 \sin(ax+b) = a^3 \cos\left(3 \cdot \frac{\pi}{2} + ax + b\right)$$

....

....

$$y_n = a^n \cos\left(\frac{n\pi}{2} + ax + b\right)$$

NOTES

Note: If $b = 0$, then $y_n = D^n \cos ax = a^n \cos\left(\frac{n\pi}{2} + ax\right)$

9. $y = \dfrac{1}{ax+b}$ where a and b are constants.

Now, $y_1 = -\dfrac{a}{(ax+b)^2} = (-1)^1 \dfrac{a}{(ax+b)^{1+1}}$

$$y_2 = (-a)\frac{(-2a)}{(ax+b)^3} = (-1)^2 \frac{a^2 \lfloor 2}{(ax+b)^{2+1}}$$

$$y_3 = (-1)^2 a^2 \frac{\lfloor 2\,(-3a)}{(ax+b)^4} = \frac{(-1)^3 a^3 \lfloor 3}{(ax+b)^{3+1}}$$

....

....

$$y_n = \frac{(-1)^n a^n \lfloor n}{(ax+b)^{n+1}}$$

10. $y = e^{ax} \sin bx$ where a and b are constants.

Now, $y_1 = ae^{ax} \sin bx + e^{ax} b \cos bx = e^{ax}(a \sin bx + b \cos bx)$

Let $a = r\cos\theta$ and $b = r\sin\theta$, then $a^2 + b^2 = r^2$ and $\tan\theta = \dfrac{b}{a}$ or

$$\theta = \tan^{-1}\left(\frac{b}{a}\right)$$

$$\therefore \quad y_1 = e^{ax}[r\cos\theta \sin bx + r\sin\theta \cos bx] = re^{ax}\sin(bx+\theta)$$

....

....

$$y_n = r^n e^{ax} \sin(bx + n\theta)$$

$$= (a^2+b^2)^{\frac{n}{2}} e^{ax} \sin\left[bx + n\tan^{-1}\frac{b}{a}\right]$$

Note: Similarly for $y = e^{ax}\cos bx$, $y_n = (a^2+b^2)^{\frac{n}{2}} e^{ax} \cos\left(bx + n\tan^{-1}\frac{b}{a}\right)$.

We now show some applications of the above.

Example 3.17: Find y_n in the following cases:

(*i*) $y = \log\left(\dfrac{a-x}{a+x}\right)$ (*ii*) $y = \dfrac{a-x}{a+x}$

(*iii*) $y = \dfrac{x^n}{x-1}$ (*iv*) $y = \dfrac{x^2}{(x-1)(x-2)(x-3)}$

NOTES

Solution: (*i*) Now $y = \log\left(\frac{a-x}{a+x}\right) = \log(a-x) - \log(a+x)$

$$\therefore \quad y_1 = \frac{-1}{a-x} - \frac{1}{a+x}$$

$$y_2 = (-)\frac{1}{(a-x)^2} - \frac{(-1)}{(a+x)^2}$$

$$y_3 = (-)\frac{\lfloor 2}{(a-x)^3} - \frac{(-1)^2 \lfloor 2}{(a+x)^3}$$

$$y_4 = (-)\frac{\lfloor 3}{(a-x)^4} - \frac{(-1)^3 \lfloor 3}{(a+x)^4}$$

....

....

$$y_n = (-)\frac{\lfloor n-1}{(a-x)^n} - \frac{(-1)^{n-1} \lfloor n-1}{(a+x)^n}$$

$$= \lfloor n-1 \left[-\frac{1}{(a-x)^n} + \frac{(-1)^n}{(a+x)^n} \right]$$

(*ii*) $y = \frac{a-x}{a+x} = \frac{2a-(x+a)}{a+x} = \frac{2a}{a+x} - 1$

$$\therefore \quad y_1 = 2a \ \frac{(-1)^1}{(a+x)^2}$$

$$y_2 = 2a \ \frac{(-1)^2 2}{(a+x)^3}$$

$$y_3 = 2a \ \frac{(-1)^3 \lfloor 3}{(a+x)^4}$$

....

....

$$y_n = 2a \ \frac{(-1)^n \lfloor n}{(a+x)^{n+1}}$$

(*iii*) $\quad y = \frac{x^n}{x-1} = \frac{x^n - 1 + 1}{x-1} = \frac{x^n - 1}{x-1} + \frac{1}{x-1}$

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

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

Since, the nth order derivative of $x^{n-1}, x^{n-2}, ..., x$ are zero and

$$D^n\left(\frac{1}{x-1}\right) = \frac{(-1)^n \lfloor n}{(x-1)^{n+1}},$$

$$y_n = \frac{(-1)^n \lfloor n}{(x-1)^{n+1}}$$

(*iv*) $\quad y = \frac{x^2}{(x-1)(x-2)(x-3)}$

$$\text{Let } y = \frac{x^2}{(x-1)(x-2)(x-3)} = \frac{A}{(x-1)} + \frac{B}{(x-2)} + \frac{C}{(x-3)}$$

$$\text{or } x^2 = A(x-2)(x-3) + B(x-1)(x-3) + C(x-1)(x-2)$$

Now putting $x = 1, 2$ and 3, we get

$$1 = A(-1)(-2) \Rightarrow A = \frac{1}{2}$$

$$4 = B(-1)(-1) \Rightarrow B = -4$$

And $9 = C(-2)(1) \Rightarrow C = \frac{9}{2}$

$$\therefore \quad y = \frac{1}{2}\frac{1}{x-1} + \frac{(-4)}{x-2} + \frac{9}{2}\cdot\frac{1}{x-3}$$

We know that $D^n\left(\frac{1}{x-a}\right) = \frac{(-1)^n \lfloor n}{(x-a)^{n+1}}$

$$\therefore \quad y_n = \frac{1}{2}\frac{(-1)^n\lfloor n}{(x-1)^{n+1}} - (4)\frac{(-1)^n\lfloor n}{(x-2)^{n+1}} + \frac{9}{2}\frac{(-1)^n\lfloor n}{(x-3)^{n+1}}$$

Leibnitz's Theorem for the *n*th Order Derivative of the Product of Two Functions

Theorem 3.8: Let u and v be two functions of x, both derivable at least upto n times, then the nth derivative of their product is given by,

$$(uv)_n = \sum_{r=0}^{n} {}^nc_r\, u_{n-r}\, v_r$$

$$= {}^nc_o\, u_n v + {}^nc_1\, u_{n-1}\, v_1 + \ldots + {}^nc_n\, u\, v_n$$

$$= u_n\, v + n\, u_{n-1}\, v_1 + \frac{n(n-1)}{\lfloor 2}\, u_{n-2}\, v_2 + \ldots + u\, v_n$$

Where the suffixes of u and v denote the orders of differentiation of u and v with respect to x.

Example 3.18: Find y_n when $y = x^3 \log x$.

Solution: Let $u = \log x$ and $v = x^3$, then $u_k = \frac{(-1)^{k-1}\lfloor k-1}{x^k}$ and $v_k = 0$ for $k \geq 4$.

Then by the Leibnitz's theorem, we get

$$y_n = (x^3 \log x)_n = (uv)_n$$

$$= u_n\, v + n\, u_{n-1}\, v_1 + \frac{n(n-1)}{\lfloor 2}\, u_{n-2}\, v_2 + \frac{n(n-1)(n-2)}{\lfloor 3}\, u_{n-3}\, v_3$$

(Other terms are Zero, $\because v_k = 0$ for $k \geq 4$)

$$= \frac{(-1)^{n-1}\lfloor n-1}{x^n}x^3 + n\frac{(-1)^{n-2}\lfloor n-2}{x^{n-1}}3x^2 + \frac{n(n-1)(-1)^{n-3}\lfloor n-3}{x^{n-2}\lfloor 2}6x$$

$$+ \frac{n(n-1)(n-2)}{\lfloor 3}\frac{(-1)^{n-4}\lfloor x-4}{x^{n-3}}6$$

$$= \frac{(-1)^n}{x^{n-3}}\lfloor n-2\,(-n+1+3n) + \frac{n(n-1)\lfloor n-4}{x^{n-3}}(-1)^n[(-n+3)3 + n - 2]$$

$$= \frac{(-1)^n\lfloor n-2}{x^{n-3}}(2n+1) + \frac{n(n-1)\lfloor n-4}{x^{n-3}}(-1)^n(-2n+7)$$

NOTES

$$= \frac{(-1)^n \underline{|n-4}}{x^{n-3}} [(n-2)(n-3)(2n+1) + n(n-1)(-2n+7)]$$

$$= \frac{(-1)^n \underline{|n-4}}{x^{n-3}} [(n^2-5n+6)(2n+1) + (n^2-n)(-2n+7)]$$

$$= \frac{(-1)^n \underline{|n-4}}{x^{n-3}} [2n^3 - 10n^2 + 12n + n^2 - 5n + 6 - 2n^3 + 2n^2 + 7n^2 - 7n]$$

$$= \frac{(-1)^n \, 6\underline{|n-4}}{x^{n-3}}.$$

Example 3.19: If $x + y = 1$, then prove that $\frac{d^n}{dx^n}(x^n y^n) = \underline{|n} \{y^n - ({}^nC_1)^2 y^{n-1} x + ({}^nC_2)^2 y^{n-2} x^2 - ({}^nC_3)^2 y^{n-3} x^3 + ... + (-1)^n x^n\}$.

Solution: Since $x + y = 1$, then $y = 1 - x$.

Let $u = x^n$ and $v = (1-x)^n$, then $u_r = n(n-1)(n-2) ... (n-r+1)x^{n-r}$

$$= \frac{n(n-1)(n-2)...(n-r+1)\underline{|n-r}}{\underline{|n-r}} x^{n-r} = \frac{\underline{|n}}{\underline{|n-r}} x^{n-r}$$

Differentiating n times by Leibnitz's theorem, we get

$$\frac{d^n}{dx^n}(x^n y^n) = \frac{d^n}{dx^n}(x^n (1-x)^n) = (uv)_n$$

$$= \sum_{r=0}^{n} {}^nC_r \; u_{n-r} v_r = {}^nC_0 \, u_n v + {}^nC_1 \, u_{n-1} v_1 + {}^nC_2 \, u_{n-2} v_2 + ... + {}^nC_n \, uv_n$$

$$= \underline{|n} (1-x)^n + {}^nC_1 \frac{\underline{|n}}{\underline{|n-n+1}} x (1-x)^{n-1} (-1)n + {}^nC_2 \frac{\underline{|n}}{\underline{|n-n+2}} x^2 \, n(n-1)(-1)^2 (1-x)^{n-2} + ... + {}^nC_n x^n \underline{|n} (-1)^n \quad [\because v_n = (-1)^n \underline{|n}]$$

$$= \underline{|n} \left[(1-x)^n - \left({}^nC_1\right) x(1-x)^{n-1} \left({}^nC_1\right) + {}^nC_2 \frac{n(n-1)}{\underline{|2}} x^2 (1-x)^{n-2} + ... + ({}^nC_n)(-1)^n x^n \right]$$

$$= \underline{|n} [y^n - ({}^nC_1)^2 xy^{n-1} + ({}^nC_2)^2 x^2 y^{n-2} + ... + (-1)^n x^n]$$

Check Your Progress

4. Define partial derivative of a function of two variables.
5. What is a homogeneous function?
6. Define a harmonic function.
7. State Leibnitz's theorem for the nth order derivative of the product of two functions.

NOTES

3.7 TAYLOR'S THEOREM

Theorem 3.9: Let $f(x+h)$ be expandable into a power series in the variable h. Again the flat assumption is that this series can be differentiated term-by-term.

Taylor's theorem states that,

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots \infty$$

Proof: Let $f(a+h) = \alpha_0 + \alpha_1 h + \alpha_2 h^2 + \alpha_3 h^3 + \ldots \infty$

Put $h = 0$ to get $\alpha_0 = f(a)$.

Differentiating each side with respect to h, we obtain

$$\frac{d}{dh}[f(a+h)] = \alpha_1 + 2\alpha_2 h + 3\alpha_3 h^2 + \ldots \infty$$

But $\frac{d}{dh}[f(a+h)] = \frac{d[f(a+h)]}{d(a+h)} \cdot \frac{d(a+h)}{d(h)}$

$$= f'(a+h) \,.\, 1 = f'(a+h)$$

This in turn yields that,

$$\alpha_1 = f'(a)$$

[Observe that $f'(a+h)$ is first derivative of $f(a+h)$ with respect to $a+h$.]

Again, differentiate both sides with respect to h.

Thus, $f''(a+h) = 2\alpha_2 + (3.2)\,\alpha_3 h + \ldots \infty$

$$\Rightarrow \qquad 2\alpha_2 = f''(a) \quad \text{or} \quad \alpha_2 = \frac{f''(a)}{2!}$$

Proceeding in this manner, we get

$$\alpha_r = \frac{1}{r!} f^{(r)}(a)$$

Hence, $f(a+h) = \sum_{r=0}^{\infty} \frac{h^r}{r!} f^{(r)}(a)$.

Example 3.20: Show that,

$$\log(n+1) = \log n + \left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \frac{1}{4n^4} + \ldots\right).$$

Solution: We expand $\log(n+h)$ in terms of h and then put $h = 1$.

$$f(n+h) = \log(n+h) \Rightarrow f(n) = \log n$$

Also, $f'(n+h) = \frac{d}{d(n+h)}[\log(n+h)] = \frac{1}{n+h}$

$$\Rightarrow \qquad f'(n) = \frac{1}{n}.$$

NOTES

Further, $f''(n+h) = -\dfrac{1}{(n+h)^2} \Rightarrow f''(n) = -\dfrac{1}{n^2}$ ans so on.

In general, $f^{(r)}(n+h) = \dfrac{(-1)^{r-1}(r-1)^r}{(n+h)^r}$

$$\Rightarrow \qquad f^{(r)}(n) = \frac{(-1)^{r-1}(r-1)}{n^r}.$$

Consequently by Taylor's expansion,

$$\log(n+h) = f(n) + h f'(n) + \frac{h^2}{2!} f''(n) + \ldots + \frac{h^r}{r!} f^{(r)}(n) + \ldots$$

$$= \log n + \frac{h}{n} - \frac{h^2}{2n^2} + \frac{2!}{3!}\frac{h^3}{n^3} + \ldots$$

$$+ \frac{(-1)^{r-1} hr\,(r-1)!}{(r!)\,nr} + \ldots$$

Put $h = 1$, to obtain

$$\log(n+1) = \log n + \left[\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \frac{1}{4n^4} + \ldots\right].$$

Example 3.21: Prove that,

$$\tan^{-1}(x+2) = \tan^{-1} x + 2\sin z\ \frac{\sin z}{1} - \frac{(2\sin z)^2}{2}\sin^2 z + \frac{(2\sin z)^3 \sin 3z}{3} - \ldots$$

Where, $\cot z = x$.

Solution: We expand $\tan^{-1}(x+h)$ by Taylor's expansion and then put $h = 2$.

Here, $f(x+h) = \tan^{-1}(x+h) \Rightarrow f(x) = \tan^{-1} x.$

In this case, $f^{(r)}(x+h) = (-1)^{r-1}(r-1)!\sin^n\theta \sin n\theta,$

Where, $\quad \cos\theta = x + h$

Thus, $\quad f^{(r)}_{(x)} = (-1)^{r-1}(r-1)!\sin^n z \sin nz$

Since by definition the value of $\cot\theta$ at $h = 0$ is $\cot z$, i.e., $\theta = z$

$$\text{Hence, } \tan^{-1}(x+h) = \tan^{-1} x + h\sin z \sin z + \frac{h^2}{2!}(-\sin^2 z \sin 2z)$$

$$+ \frac{h^3}{3!}(2!\sin^3 z \sin 3z)\ldots$$

$$= \tan^{-1} x + h\sin z\ \frac{\sin z}{1} - \frac{h^2\sin^2 z}{2}\sin 2z$$

$$+ \frac{h^3}{3}\sin^3 z \sin 3z \ldots$$

Put $h = 2$, to get the required result.

Example 3.22: Show that if x is numerically less than 1,

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 +$$

NOTES

Solution. Here, $f(a+h) = \frac{1}{(a+h)^2}$

$$\Rightarrow \qquad f(a) = \frac{1}{a^2}$$

$$f'(a+h) = -\frac{2}{(a+h)^3}$$

$$\Rightarrow \qquad f'(a) = -\frac{2}{a^3}$$

Again, $$f''(a+h) = \frac{6}{(a+h)^4}$$

$$\Rightarrow \qquad f''(a) = \frac{6}{a^4}$$

and so on.

By Taylor's, expansion

$$\frac{1}{(a+h)^2} = \frac{1}{a^2} - \frac{2}{a^3}h + \frac{6}{a^4}\frac{1}{2!}h^2 ...$$

Put $a = -1$, $h = x$ to get

$$\frac{1}{(-1+x)^2} = 1 + 2x + 3x^2 + ...$$

Or $$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + ...$$

Notes: 1. Series on RHS is convergent only when x is numerically less than 1.

2. One might think that in the proof of $\frac{d}{dx}\left(\frac{1}{x^2}\right)$ (by definition), we use the series $(1-x)^{-2} = 1 + 2x + 3x^2 + ...$. We could avoid this circular argument by finding out derivative of x^{-n}, n is any integer by following technique

$$\frac{d}{dx}\left(\frac{1}{x^n}\right) = \frac{\frac{d}{dx}(1)x^n - 1.\,nx^{n-1}}{x^{2n}}$$

$$= \frac{0.x^n - nx^{n-1}}{x^{2n}}$$

$$= \frac{-nx^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}} = -nx^{-n-1}$$

Now, put $n = 1, 2, 3, ...$, to get derivative of reciprocal of any integral power of x.

3.8 EXPLICIT AND IMPLICIT FUNCTIONS

If the dependent variable y is expressed in terms of the independent variable x, we call y an explicit function of x and denote such a function by $y = f(x)$. One can similarly define an explicit function $x = f(y)$ where y is the independent variable and x depends on y. Thus $y = x \sin x + 5 \log x - 2^x$ is an explicit function. But often it may not be possible to relate a dependent variable to the independent variable in such an explicit form, yet it may be possible to get y as a function of x or x as a function of y under some stringent conditions. These conditions are given by the well-known implicit function theorem. It is to be clearly understood that the conditions of this theorem assert the existence of a function but do not provide the function itself.

Derivatives of Implicit Functions

Theorem 3.10: If $F(x, y) = 0$ defines y as an implicit function of x, then

$$\frac{dy}{dx} = -\frac{F_x}{F_y} \text{ provided } F_y \neq 0.$$

Proof: Since $F(x, y) = 0$, then $\dfrac{dF}{dx} = 0$. ... (3.8)

By the chain rule, we get

$$\frac{d}{dx}\{F(x, y)\} = \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx}$$

or $F_y \dfrac{dy}{dx} + F_x = 0$ (By Equation (3.8)) or $\dfrac{dy}{dx} = -\dfrac{F_x}{F_y}$.

Notes:

1. If $F(x, y, z) = 0$ defines z as an implicit function of x and y

   then $\dfrac{\partial z}{\partial x} = -\dfrac{F_x}{F_z}$ and $\dfrac{\partial z}{\partial y} = -\dfrac{F_y}{F_z}$ provided $F_z \neq 0$.

2. If $F(x, y) = 0$ defines y as an implicit function of x, then

   $$\frac{d^2y}{dx^2} = -\frac{(F_y)^2 F_{xx} - 2F_x F_y F_{xy} + (F_x)^2 F_{yy}}{F_y^{\,3}} \text{ provided } F_y \neq 0.$$

Example 3.23: Find $\dfrac{dy}{dx}$ if $a^2x^3 + b^2y^3 - 3abxy = 0$.

Solution: Let $F(x, y) = a^2x^3 + b^2y^3 - 3abxy$;

$\therefore F_x = 3a^2x^2 - 3aby$ and $F_y = 3b^2y^2 - 3abx$.

$$\therefore \frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{3a(ax^2 - by)}{3b(by^2 - ax)} = -\frac{a}{b}\frac{(ax^2 - by)}{(by^2 - ax)}.$$

3.9 INVERSE FUNCTION THEOREM AND IMPLICIT FUNCTION THEOREM

NOTES

The Inverse Function Theorem

Theorem 3.11: Let $f = (f_1, f_2, \ldots, f_n) \in C'$ on an open set S in R^n, and let $T = f(s)$. If the Jacobian determinant $J_f(a) \neq 0$ for some point a in S, then there are two open sets $X \subseteq S$ and $Y \subseteq T$ and a uniquely determined function g such that,

(*a*) $a \in X$ and $f(a) \in Y$.

(*b*) $Y = f(X)$.

(*c*) F is one to one on X.

(*d*) G is defined on Y, $g(Y) = X$, and $g[f(X)]$ for every $x \in X$.

(*e*) $g \in C'$.

Proof: $f = (f_1, f_2, \ldots, f_n) \in C'$ on S. So f is continuous on S.

$\Rightarrow J_f$ is continuous on S.

$\Rightarrow J_f(a) \neq 0$.

$\Rightarrow$ There is an n-ball $B_1(a)$ such that $J_f(X) \neq 0$ for all X in $B_1(a)$.

$\Rightarrow$ There is an n-ball $B(a) \subseteq B_1(a)$ on which f is one to one.

Let B be an n-ball with center at a and radius smaller than that of $B(a)$.

Then $f(B)$ contains an n-ball with center at $f(a)$. Denote this by Y and suppose $X = f^{-1}(Y) \cap B$.

$f \in C'$ on S.

$\Rightarrow$ f is differentiable on S.

$\Rightarrow$ f is continuous on S.

Therefore, $f^{-1}(Y)$ is open (being an n-ball Y is open).

$\Rightarrow$ $f^{-1}(Y) \cap B$ is open.

$\Rightarrow$ X is open.

(a) We have,

$a \in B$ and $f(a) \in Y$

$\Rightarrow$ $a \in B$ and $a \in f^{-1}(Y)$

$\Rightarrow$ $a \in B \cap f^{-1}(Y)$

$\Rightarrow$ $a \in X$

(b) $X = f^{-1}(Y) \cap B$

$f(X) = f[f^{-1}(Y) \cap B]$

$\subseteq f(f^{-1}(Y)) \cap f(B)$

$= Y \cap f(B)$

NOTES

$$f(X) \subseteq Y \qquad \dots(3.9)$$

Let $y \in Y \quad \Rightarrow \quad y \in f(B)$

$\Rightarrow \quad y = f(b)$ for some $b \in B$

That is,

$$f(b) = y \in Y$$

So,

$$b \in f^{-1}(Y)$$

Thus,

$y \in Y \Rightarrow \quad b \in B$ and $b \in f^{-1}(Y)$

$\Rightarrow \quad b \in B \cap f^{-1}(Y)$

$\Rightarrow \quad b \in X$

That is,

$y \in Y \Rightarrow \quad y = f(b)$ for some $b \in X$

That is,

$$Y \subseteq f(X) \qquad \dots(3.10)$$

From Equations (3.9) and (3.10),

$$Y = f(X)$$

(c) We have f is one to one on B and $X \subseteq B$. Therefore f is one to one on X.

(d) The set B is compact and f is one to one and continuous on B.

Let $f: S \to T$ be a function from one metric space to another metric space. Assume that f is one to one on S, so that the inverse function f^{-1} exists. If S is compact and if f is continuous on S, then f^{-1} is continuous on $f(S)$. There exists a function g defined on $\overline{B}$ such that,

$g(f(X)) = X$ for all $x \in B$

Also,

g is continuous on $f(\overline{B})$

Since,

$Y \subseteq f(\overline{B})$, g is defined on Y.

We have,

$$f(X) = Y$$

So from the definition of g, we get $g(Y) = X$ and $g(f(X)) = X$ for all $x \in X$.

To prove g is unique, let there exists h on $f(\overline{B})$ satisfying (d).

Then,

$h(f(X)) = X$ for all $x \in X$

Also, we have

$$g(f(X))=X \text{ for all } x \in X$$

Let $y \in Y$.

Then there exists a unique $x \in X$ such that,

$$y = f(X) \text{ [Because } f \text{ is one-to-one]}$$

Therefore,

$$\begin{aligned} h(Y) &= h(f(X)) \\ &= x \\ &= g(f(X)) \\ &= g(Y) \end{aligned}$$

That is,

$$h(Y) = g(Y) \text{ for all } y \in Y$$

$$\Rightarrow \quad h = g$$

(e) Define a real valued function h by,

$$h(Z) = \det[D_j f_i(z_i)] \quad \text{where } z_1, z_2, ..., z_n \in S$$

And $(z_1, z_2, \ldots, z_n)$ is the corresponding element in R^{n^2}. h is continuous at the point Z in R^{n^2} where $h(Z)$ is defined. Let Z be the special point in R^{n^2} obtained by putting $z_1 = z_2 = \ldots = z_n = a$.

Then,

$$\begin{aligned} h(Z) &= \det [D_j f_i(z_i)] \\ &= J_r(a) \\ &\neq 0 \end{aligned}$$

Hence, by continuity of h, there is some n-ball $B_2(a)$ such that $h(Z) \neq 0$ for all $z_1, z_2, \ldots, z_n \in B_2(a)$.

We can now assume that, the n-ball $B(a)$ was chosen so that $B(a) \subseteq B_2(a)$.

Then, $\overline{B} \subseteq B_2(a)$ and hence $h(Z) \neq 0$ for each $z_i \in \overline{B}$.

In order to prove that $g = (g_1, g_2, \ldots, g_n) \in C'$ on Y, it is enough to show that $g_k \in C'$ on Y.

For proving that $D_r g_k$ exists on Y, let $y \in Y$. Since, Y is open, $y + tu_r \in Y$ for sufficiently small t, where u_r is the rth unit coordinate vector in R^n.

Consider $\dfrac{gk(y + tu_r) - gk(y)}{t}$

And let $x = g(y)$ and $x' = g(y + tu_r)$. Then both x and x' are in X and $f(x') - f(x) = y + tu_r - y = tu_r$.

NOTES

Hence,

$$f_i(x') - f_i(x) = t, \text{ if } i = r$$
$$= 0, \text{ if } i \neq r$$

By the mean value theorem we have,

$$\frac{f_i(x') - f_i(x)}{t} = \nabla f_i(\overline{z_i}) \cdot \left(\frac{x' - x}{t}\right) \text{ for } i = 1, 2, ..., n,$$

Where $\overline{z_i}$ lies in the line segment joining x and x' and hence $\overline{z_i} \in B$.

Therefore,

$$\nabla f_i(z_i).\left(\frac{x' - x}{t}\right) = 0, \quad \text{if} \quad i \neq r$$
$$= 1, \quad \textit{if} \quad i = r \text{ for } i = 1, 2, ..., n$$

That is,

$$\sum_{j=1}^{n} D_j f_i(z_i) \cdot \left(\frac{x' - x}{t}\right) = 0, \quad \text{if} \quad i \neq r$$
$$= 1, \quad \text{if} \quad i = r \text{ for } i = 1, 2, ..., n$$

This is a system of n linear equations in n unknowns $\frac{x_j' - x_j}{t}$ and has a unique solution, since det $[D_j f_i(z_i)] = h(Z) \neq 0$. On solving the kth unknown by Cramer's rule we get,

$$\frac{x'_k - x_k}{t} = \frac{g_k(y + tu_r) - g_k(y)}{t}$$
$$= \frac{D_k}{D},$$

Where, D_k is the determinant of the matrix obtained by replacing the kth column by u_r and $D = \det[D_j f_i(z_i)]$.

Now,

$t \to 0 \Rightarrow z_i \to X$, since z_i is on the line segment joining x and x'.

Also, $\lim_{t \to 0} D = \lim_{t \to 0} \det[D_j f_i(z_i)]$

$$= \det[D_j f_i(x)]$$
$$= J_f(x)$$
$$\neq 0, \text{ since } x \in X$$

Therefore,

$$\lim_{t \to 0} \frac{g_k(y + tu_r) - g_k(y)}{t} \text{ exists.}$$

Or,

$D_r g_k(y)$ exists for $r = 1, 2, \ldots, n$.

Furthermore, $D_r g_k(y)$ is a quotient of two determinants involving the derivatives $D_j f_i(X)$. But each $D_j f_i(X)$ is continuous. Therefore, $D_r g_k(y)$ is continuous. That is, $g_k \in C^1$ on Y. Hence, $g \in C^1$ on Y.

Implicit Function Theorem

Notation: Points in the $(n+k)$ dimensional space R^{n+k} will be written in the form $(x;t)$ where $x = (x_1, \ldots, x_n) \in R^n$ and $t = (t_1, \ldots, t_k) \in R^k$.

Theorem 3.12: Let $f = (f_1, \ldots, f_n)$ be a vector valued function defined on an open set S in R^{n+k} with values in R^n. Suppose $f \in C^1$ on S. Let $(x_0;t_0)$ be a point in S for which the $n \times n$ determinant $\det[D_j f_i(x_0;t_0)] \neq 0$. Then there exists a k dimensional open set T_0 containing t_0 and, one and only one, vector valued function g, defined on T_0 and having values in R^n, such that

(a) $g \in C^1$ on T_0

(b) $g(t_0) = x_0$

(c) $f(g(t;t)) = 0$ for every t in T_0

Proof: Define a vector valued function $F = (F_1, F_2, \ldots, F_n; F_{n+1})$ on S having values in R^{n+k} and apply inverse function theorem to F. The function F is defined as follows:

Let $F_m(x;t) = f_m(x;t)$ for $1 \leq m \leq n$ and let $F_{n+m}(x;t) = t_m$ for $1 \leq m \leq k$.

Thus, $F = (f;I)$, where $f = (f_1, f_2 \ldots f_n)$ and I is the identity function defined by $I(t) = t$ for each $t \in R^k$.

Now,

$F \in C^1$ on S, since $f \in C^1$ and $I \in C^1$ on S.

Also,

$$J_F(x;t) = \det \begin{pmatrix} D_1F_1(x;t) & \cdots & D_nF_1(x;t) & D_{n+1}F_1(x;t) & \cdots & D_{n+k}F_1(x;t) \\ D_1F_2(x;t) & \cdots & D_nF_2(x;t) & D_{n+1}F_2(x;t) & \cdots & D_{n+k}F_2(x;t) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ D_1F_n(x;t) & \cdots & D_nF_n(x;t) & D_{n+1}F_n(x;t) & \cdots & D_{n+k}F_n(x;t) \\ D_1F_{n+1}(x;t) & \cdots & D_nF_{n+1}(x;t) & D_{n+1}F_{n+1}(x;t) & \cdots & D_{n+k}F_{n+1}(x;t) \\ D_1F_{n+2}(x;t) & \cdots & D_nF_{n+2}(x;t) & D_{n+1}F_{n+2}(x;t) & \cdots & D_{n+k}F_{n+2}(x;t) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ D_1F_{n+k}(x;t) & \cdots & D_nF_{n+k}(x;t) & D_{n+1}F_{n+k}(x;t) & \cdots & D_{n+k}F_{n+k}(x;t) \end{pmatrix}$$

NOTES

$$= \det\begin{pmatrix} D_1f_1(x;t) & \cdots & D_nf_1(x;t) & D_{n+1}f_1(x;t) & \cdots & D_{n+k}f_1(x;t) \\ D_1f_2(x;t) & \cdots & D_nf_2(x;t) & D_{n+1}f_2(x;t) & \cdots & D_{n+k}f_2(x;t) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ D_1f_n(x;t) & \cdots & D_nf_n(x;t) & D_{n+1}f_n(x;t) & \cdots & D_{n+k}f_n(x;t) \\ D_1f_{n+1}(x;t) & \cdots & D_nf_{n+1}(x;t) & D_{n+1}f_{n+1}(x;t) & \cdots & D_{n+k}f_{n+1}(x;t) \\ D_1f_{n+2}(x;t) & \cdots & D_nf_{n+2}(x;t) & D_{n+1}f_{n+2}(x;t) & \cdots & D_{n+k}f_{n+2}(x;t) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ D_1f_{n+k}(x;t) & \cdots & D_nf_{n+k}(x;t) & D_{n+1}f_{n+k}(x;t) & \cdots & D_{n+k}f_{n+k}(x;t) \end{pmatrix}$$

$$= \det\begin{pmatrix} M & N \\ 0 & I \end{pmatrix}, \text{ where } M = \det\begin{pmatrix} D_1f_1(x;t) & \cdots & D_nf_1(x;t) \\ D_1f_2(x;t) & \cdots & D_nf_2(x;t) \\ \vdots & & \vdots \\ D_1f_n(x;t) & \cdots & D_nf_n(x;t) \end{pmatrix}$$

And $\quad N = \det\begin{pmatrix} D_{n+1}f_1(x;t) & \cdots & D_{n+k}f_1(x;t) \\ D_{n+1}f_2(x;t) & \cdots & D_{n+k}f_2(x;t) \\ \vdots & & \vdots \\ D_{n+1}f_n(x;t) & \cdots & D_{n+k}f_n(x;t) \end{pmatrix}$

$= \det M$

$= \det [D_j\, f_i\, (\text{x; t})]$

So $J_F\,(x_0;\, t_0) = \det\,[D_jf_i\,(x_0;\, t_0)]$

$\neq 0$

Also,

$F\,(x_0;t_0) = (f(x_0;t_0);t)$

$\qquad = (0,t_0)$

Now, by inverse function theorem, there exist open sets X and Y and a unique function,

G: $Y \rightarrow X$ which satisfy the following properties:

1. $(x_0;t_0) \in X$ and $F(x_0;t_0) = (0,t_0) \in Y$.
2. $Y= F(X)$.
3. F is one to one on X.
4. $G(Y) = X$ and $G(\,F(x;t))$ for every $(x;t) \in X$.
5. $G \in C^1$ on Y.

Now, G can be reduced to components as follows:

$G = (v;w)$ where $v = (v_1,\ldots, v_n)$ is a vector valued function defined on Y with values in R^n and

$w = (w_1,\ldots, w_k)$ is also defined on Y but has values in R^k.

To determine v and w explicitly, F is one to one on X and $F^{-1}(Y))$ contains X. Hence, for every $(x;t)$ in Y can be written uniquely as $(x;t) = F(x';t')$ for some $(x';t')$ in X. From the way in which F was defined, we must have $t'=t$.

Hence,

$G(F(x';t')) = G(x;t)$

$\Rightarrow (x';t) = (v(x;t);w(x;t))$

$\Rightarrow v(x;t) = x'$ and $w(x;t) = t$

Additionally,

$F(x';t') = F(v(x;t);t)$ for every $(x;t) \in Y$.

Now define the set T_0 and the function g as,

$T_0 = \{t/t \in R^k, (0;t) \in Y\}$

and for every t in T_0 define $g(t) = v(0;t)$.

Then, T_0 is open, since Y is open.

$G \in C^1$ on Y.

Therefore, $g \in C^1$ on Y, since the components of g are taken from the components of G.

Also, $g(t_0) = v(0;t_0)$

$= x_0$

$f(g(t;t) = f(v(0;t);t)$

$= f(x_0;t)$

$= 0$, for every $t \in T_0$

Now, we will prove the uniqueness of g. If there were another function h which satisfies (c) then,

$f(g(t);t) = f(h(t);t)$ for all $t \in T_0$

$\Rightarrow (g(t);t) = (h(t);t)$, since f is one-to-one

$\Rightarrow g(t) = h(t)$

$\Rightarrow g = h$

Check Your Progress

8. Give the statement of Taylor's theorem.
9. Define an explicit function.
10. State the inverse function theorem.

3.10 CHANGE OF VARIABLES

Theorem 3.13: If (i) $f \in \Re\,[a, b]$, (ii) ϕ is derivable, strictly monotonic on $[\alpha, \beta]$ and maps it onto $[a, b]$, and (iii) $\phi' \in \Re\,[\alpha, \beta]$, then

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(t)\phi'(t)dt.$$

Proof: Let ϕ be strictly monotonically increasing on $[\alpha, \beta]$, and $Q = \{\alpha = t_0, t_1, t_2, ..., t_n = \beta\}$ be a partition of $[\alpha, \beta]$, then $P = \{a = x_0, x_1, x_2, ..., x_n = b\}$ be the corresponding partition of $[a, b]$, for $x_i = \phi(t_i)$, $i = 0, 1, 2, ..., n$.

By Lagrange's mean value theorem,

$\delta x_r = x_r - x_{r-1} = \phi(t_r)\ \phi(t_{r-1}) = \phi'(\eta_r)\ \delta t_r$, where $\eta_r \in \delta t_r$ and $\xi_r = \phi(\eta_r)$, $r = 1, 2, ..., n$.

so that

$$\Sigma f(\xi_r)\ \delta\xi_r = \Sigma f(\phi(\eta_r))\ \phi'(\eta_r)\ \delta t_r. \qquad ...(3.11)$$

and ϕ being derivable is uniformly continuous on $[\alpha, \beta]$, and consequently $\|Q\| \to 0$ as $\| P \| \to 0$. Thus, letting $\| P \| \to 0$ as

$$\Sigma f(\xi_r)\delta x_r \to \int_a^b f(x)dx, \text{ and } \Sigma f(\phi(\eta_r))\phi'(\eta_r)\delta t_r \to \int_\alpha^\beta f(\phi(t))\ \phi'(t)dt,$$

Equation (3.11) gives that

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(t)\ \phi'(t)dt.$$

With some adjustments in the above proof, the theorem also holds for strictly monotonically decreasing ϕ.

Notes: **1.** If $\phi' \neq 0$ on $[\alpha, \beta]$ then ϕ is strictly monotonic on $[\alpha, \beta]$. Hence in the theorem the condition of strictly monotonic of ϕ can be replaced by $\phi' \neq 0$ on $[\alpha, \beta]$.

2. The theorem still holds even if $\phi' = 0$ for a finite number of times on $[\alpha, \beta]$. In that case $[\alpha, \beta]$ can be divided into a finite number of subintervals in each of which ϕ is strictly monotonic and the change of variable being valid in each of the subintervals, the result follows.

Conclusively, to evaluate $\int_\alpha^\beta f(x)dx$, if we put $x = g(t)$, where $g(a) = \alpha$, $g(b) = \beta$ and g' is continuous on $[a, b]$ vanishing at the most a finite number of times, then

$$\int_\alpha^\beta f(x)dx = \int_a^b f(g(t))\ g'(t)dt$$

The auxiliary function g mapping $[a, b]$ onto $[\alpha, \beta]$ is chosen in such a way so that the last integral is easily known.

Hence, $\int_\alpha^\beta f$ can be evaluated in many cases.

Example 3.24: To evaluate $\int\limits_0^1 (1 + x^{1/3})dx$.

Let $(1 + x^{1/3}) = t$, then $x = (t^3 - 1)^3 = t^9 - 3t^6 + 3t^3 - 1$ gives

$$\int\limits_0^1 (1 + x^{1/3})\ dx = \int\limits_1^{2^{1/3}} t(9t^8 - 18t^5 + 9t^2)dt$$

$$= \frac{9}{10}t^{10} - \frac{18}{7}t^7 + \frac{9}{4}t^4 \Big|_1^{2^{1/3}}$$

$$= \frac{90}{70}2^{1/3} - \frac{81}{140}$$

NOTES

3.11 EXTREME VALUES OF EXPLICIT FUNCTIONS AND STATIONARY VALUES OF IMPLICIT FUNCTIONS

Extrema of Functions of One Variable

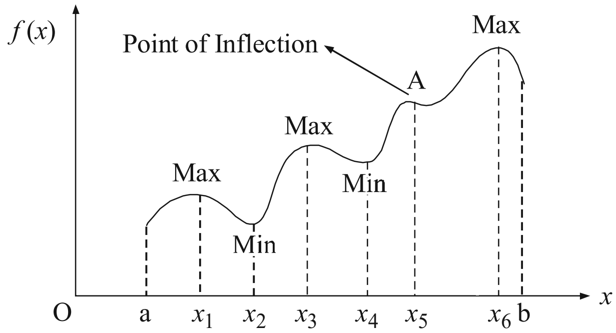
A function $y = f(x)$ has a maximum value at a point x_0 if for $|h|$ sufficiently small,

$$f(x_0 + h) < f(x_0)$$

Similarly, a function $y = f(x)$ has a minimum value at a point x_0 if

$$f(x_0 + h) > f(x_0)$$

Let $y = f(x)$ be a continuous function defined on the interval (a, b). The points x_1, x_2, x_3, x_4 and x_6 (not x_5) represent all the points of maxima and minima in $[a, b]$ (called the stationary or critical points). These include x_1, x_3 and x_6 as the points of maxima, and x_2 and x_4 as the points of minima.



Global Maxima: Since $f(x_6) > f(x)$ for all $x \neq x_6, f(x_6)$ is called the global maxima whereas $f(x_1), f(x_2)$ are called local or relative maxima. Observe $f(x_6) = \max\ \{f(x_1),\ f(x_3),\ f(x_6)\}$.

Global Minima: Since $f(x_2) < f(x)$ for all $x \neq x_2$,

$f(x_2)$ is called global minima whereas $f(x_4)$ is called a local minima.

The point A corresponding to $f(x_5)$ is a point of inflection.

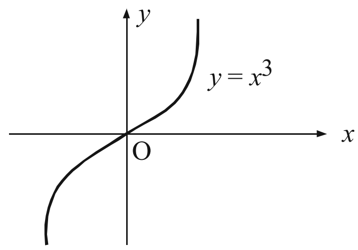
Notes:

1. A function may have more than one maximum values.
2. A function may have more than one minimum values.
3. A function may have no maximum or minimum values.

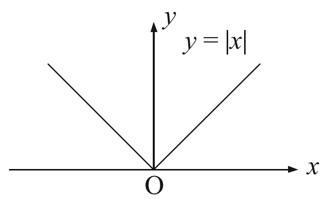
Necessary Conditions for Maximum and Minimum: If $f(x)$ be a maximum or a minimum at $x = c$ and if $f'(c)$ exists, then $f'(c) = 0$.

At the points of maximum and minimum, if the function $y = f(x)$ has a derivative, the tangent line to the curve $y = f(x)$ at each of these points is parallel to the x-axis.

From the above necessary condition, it follows immediately that if for all considered values of x, the function $f(x)$ has a derivative, then it can have an extremum (maximum or minimum) only at those values for which the derivative vanishes. The converse does not hold: it cannot be said that there definitely exists a maximum or minimum for every value at which the derivative vanish. The function $y = x^3$ at $x = 0$ has a derivative equal to zero ($y' = 3x^2$, $(y')_{x=0} = 0$) but at this point the function has neither a maximum nor a minimum.



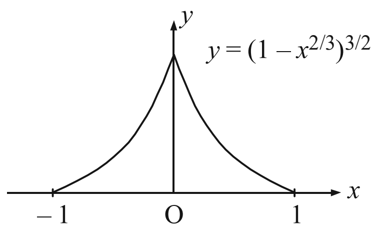
The function $y = |x|$ has no derivative at the point $x = 0$ but the function has a minimum value 0 at $x = 0$.



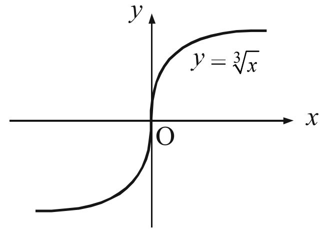
The function $y = (1 - x^{2/3})^{3/2}$ has no derivative at $x = 0$.

$$\left[\text{Note that } y' = \frac{3}{2}\left(1 - x^{2/3}\right)^{1/2}\left(\frac{-2}{3}x^{-1/3}\right) = -x^{-1/3}\left(1 - x^{2/3}\right)^{1/2} \text{ becomes infinite at } x = 0\right]$$

But the function has maximum value $y = 1$ at $x = 0$.



The function $y = \sqrt[3]{x}$ has no derivative at $x = 0$ (Note that $y \to \infty$ as $x \to 0$). At this point the function has neither a maximum nor a minimum.



So, a function can have an extremum only in two cases:

(*i*) At the points where the derivative exists and is zero.

(*ii*) At the points where the derivative does not exist.

Determination of Maxima and Minima

(*a*) If c be an interior point in the interval in which the function $f(x)$ is defined, $f'(c) = 0$ and $f''(c) \neq 0$, then $f(c)$ is

(*i*) a maximum at $x = c$ if $f''(c) < 0$.

(*ii*) a minimum at $x = c$ if $f''(c) > 0$.

(*b*) If c be an interior point of the interval of definition of the function $f(x)$ and if $f'(c) = f''(c) = ... = f^{n-1}(c) = 0$ and $f^n(c) \neq 0$, then

(*i*) if n be even, $f(c)$ is a maximum or minimum, according as $f^n(c)$ is negative or positive.

(*ii*) if n be odd, $f(c)$ is neither maximum nor minimum.

Example 3.25: Show that the maximum value of $\left(\frac{1}{x}\right)^x$ is $e^{1/e}$.

Solution: Let $f(x) = \left(\frac{1}{x}\right)^x$, then $\log f(x) = x \log \left(\frac{1}{x}\right) = -x \log x$

$$\therefore \quad \frac{f'(x)}{f(x)} = -(1 + \log x)\,[1 + \log x] \qquad ...(1)$$

or $\quad f'(x) = -f(x)\ (1 + \log x)$

From the necessary condition of extrema, we get $f'(x) = 0$

$\Rightarrow 1 + \log x = 0 \ (\because f(x) \neq 0)$

$$\Rightarrow \qquad x = \frac{1}{e}.$$

Differentiating Equation (1) with respect to x, we get

$$\frac{f''(x)}{f(x)} - \frac{[f'(x)]^2}{[f(x)]^2} = -1/x$$

$$\therefore \qquad f''(x) = \left[-\frac{1}{x} + \frac{[f'(x)]^2}{[f(x)]^2}\right] f(x)$$

$$\text{For } x = \frac{1}{e},\ f''\left(\frac{1}{e}\right) = [-e+0]\ f\left(\frac{1}{e}\right) = -e^{1/e} < 0$$

Hence $f(x) = (x)^{1/x}$ has a maximum value at $x = 1/e$ and the maximum value is $e^{1/e}$.

Example 3.26: Show that the maximum value of $\left(x + \frac{1}{x}\right)$ is less than its minimum value.

Solution: Let $f(x) = x + \frac{1}{x}$, then $f'(x) = 1 - \frac{1}{x^2}$ and $f''(x) = \frac{2}{x^3}$

$$f'(x) = 0 \text{ gives } 1 - \frac{1}{x^2} = 0 \Rightarrow x = \pm 1$$

For $x = 1$, $f''(1) = 2 > 0$. Hence at $x = 1$, $f(x)$ has a minimum value and the minimum value is 2.

For $x = -1$, $f''(-1) = -2 < 0$, Hence at $x = -1$, $f(x)$ has a maximum value and the maximum value is 0. This shows that the maximum value (0) of $\left(x + \frac{1}{x}\right)$ is less than its minimum value (2).

Extrema for Functions of Two Variables

A function $f(x, y)$ is said to have a maximum or a minimum value at the point (a, b) of the domain of $f(x, y)$, provided we can find a positive number δ such that for all values of x, y in $a - \delta < x < a + \delta$ and $b - \delta < y < b + \delta$, $(x \neq a, y \neq b)$

$$f(x, y) \lessgtr f(a, b)$$

i.e., if $f(a + h, b + k) - f(a, b) \lessgtr 0$ for $|h| < \delta$ and $|k| < \delta$, $f(a, b)$ is called an extreme value of $f(x, y)$ if it is either a maximum or a minimum.

Necessary Conditions for Maxima and Minima: If a function $f(x, y)$ has an extreme value (maximum or minimum) at (a, b) and if the first partial derivatives f_x and f_y exist at (a, b), then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Sufficient Condition for the Extremum of a Function *f* (*x*, *y*) at (*a*, *b*): If $f_x(a, b) = 0$, $f_y(a, b) = 0$ and $f_{xx}(a, b) = A$, $f_{xy}(a, b) = B$, $f_{yy}(a, b) = C$, then

1. $f(a, b)$ is a maximum value of $f(x, y)$ at (a, b) if $AC - B^2 > 0$ and $A < 0$.
2. $f(a, b)$ is a minimum value of $f(x, y)$ at (a, b) if $AC - B^2 > 0$ and $A > 0$.
3. $f(a, b)$ is neither a maximum nor a minimum value of $f(x, y)$ at (a, b) if $AC - B^2 < 0$.
4. The case is doubtful and needs further investigation if $AC - B^2 = 0$.

Saddle Point: A point (a, b) is said to be saddle point of a function $f(x, y)$ if $f(x, y)$ has neither a maximum nor a minimum at (a, b) though $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Critical Point: A point (a, b) is said to be a critical point of a function $f(x, y)$ if $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Example 3.27: Find the extreme values of,

$$f(x, y) = x^2 + xy + y^2 + ax + by$$

and determine whether the value you get is maximum or minimum.

Solution: Here, $f(x, y) = x^2 + xy + y^2 + ax + by$

$\therefore \quad f_x = 2x + y + a,\ f_y = x + 2y + b$

For maxima or minima of $f(x, y)$,

$f_x = 0$ and $f_y = 0$

i.e., $2x + y + a = 0$ and $x + 2y + b = 0$

Solving these two equations, we get

$$x = \frac{1}{3}(b - 2a) \text{ and } y = \frac{1}{3}(a - 2b)$$

Now, $f_{xx} = 2 = A, f_{yy} = 2 = C$ and $f_{xy} = B = 1$

$\therefore$ At the point $[\frac{1}{3}(b - 2a), \frac{1}{3}(a - 2b)]$,

$AC - B^2 = 4 - 1 = 3$

And $\quad A = 2$

Since $AC - B^2 > 0$ and $A > 0$, $f(x, y)$ is minimum at $[\frac{1}{3}(b - 2a), \frac{1}{3}(a - 2b)]$ and the minimum value of $f(x, y)$

$$= \frac{1}{9}(b-2a)^2 + \frac{1}{9}(b-2a)(a-2b) + \frac{1}{9}(a-2b)^2 + \frac{a}{3}(b-2a) + \frac{b}{3}(a-2b)$$

$$= \frac{1}{9}(-3b^2 - 3a^2 + 3ab) = \frac{1}{3}(ab - a^2 - b^2).$$

Example 3.28: Find all the maxima and minima of the function,

$$f(x, y) = x^3 + y^3 - 63(x + y) + 12xy$$

Solution: Here, $f(x, y) = x^3 + y^3 - 63(x + y) + 12xy$

$\therefore\ f_x = 3x^2 - 63 + 12y$ and $f_y = 3y^2 - 63 + 12x$

For the extreme of $f(x, y)$, $f_x = 0$ and $f_y = 0$

i.e., $\quad 3x^2 + 12y - 63 = 0 \quad$ and $\quad 3y^2 + 12x - 63 = 0$

i.e., $\quad x^2 + 4y - 21 = 0 \quad$ and $\quad y^2 + 4x - 21 = 0$

Subtracting 2nd from the 1st, we get

$$(x^2 - y^2) + 4(y - x) = 0$$

Or $\quad (x - y)(x + y - 4) = 0$

$\therefore \quad x - y = 0$ or $x + y - 4 = 0$

So we get two sets of equations

$$x^2 + 4y - 21 = 0, \qquad x - y = 0$$

And $\quad x^2 + 4y - 21 = 0, \quad x + y - 4 = 0$

NOTES

Solving these two sets of equations we get the following sets of points as root of $f_x = 0$ and $f_y = 0$

$$(3, 3), (-7, -7), (5, -1), (-1, 5)$$

Now, $A = f_{xx} = 6x$, $C = f_{yy} = 6y$ and $B = f_{xy} = 12$

At (3, 3), we have $A = 18$, $B = 12$ and $C = 18$.

So that $AC - B^2 = 18^2 - 12^2 > 0$ and $A > 0$

$\therefore$ $f(x, y)$ is minimum at (3, 3)

At (–7, –7), we have $A = -42$, $B = 12$, $C = -42$

So that $AC - B^2 = 42^2 - 12^2 > 0$ and $A < 0$

$\therefore$ $f(x, y)$ is maximum at (–7, –7)

At (5, –1), we have $A = 30$, $B = 12$ and $C = -6$

So that $AC - B^2 = -180 - 12^2 < 0$

$\therefore$ $f(x, y)$ is not extremum at (5, –1)

At (–1, 5), we have $A = -6$, $B = 12$ and $C = 30$

So that $AC - B^2 = -180 - 12^2 < 0$

$\therefore$ $f(x, y)$ is not extremum at (–1, 5).

3.12 LAGRANGE'S MULTIPLIERS METHOD

If we have to find the stationary value of a function of several variables which are not independent but interconnected by some relations then we try to convert the given function to one having least number of variables using the given conditions.

When such a procedure fails we use the method of Lagrange's multipliers which is described below.

Let $u = f(x, y, z)$ be the function whose maximum or minimum values are to be determined. Let the variables x, y, z be connected by the relation $v(x, y, z) = 0$.

For u to be a maximum or minimum it is necessary that

$$\frac{\partial u}{\partial x} = 0, \qquad \frac{\partial u}{\partial y} = 0, \qquad \frac{\partial u}{\partial z} = 0$$

$$\therefore \qquad \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz = 0 \qquad \ldots(3.12)$$

From $v(x, y, z) = 0$ we get

$$\frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy + \frac{\partial v}{\partial z}dz = 0 \qquad \ldots(3.13)$$

NOTES

Multiplying Equation (3.13) by the Lagrange multiplier λ and adding it to Equation (3.12) we get,

$$\left(\frac{\partial u}{\partial x}+\lambda\frac{\partial v}{\partial x}\right)dx+\left(\frac{\partial u}{\partial y}+\lambda\frac{\partial v}{\partial y}\right)dy+\left(\frac{\partial u}{\partial z}+\lambda\frac{\partial v}{\partial z}\right)dz=0 \quad \ldots(3.14)$$

This equation will be satisfied if we use the conditions,

$$\left(\frac{\partial u}{\partial x}+\lambda\frac{\partial v}{\partial x}\right)=0 \qquad \left(\frac{\partial u}{\partial y}+\lambda\frac{\partial v}{\partial y}\right)=0 \qquad \left(\frac{\partial u}{\partial z}+\lambda\frac{\partial v}{\partial z}\right)=0$$

Using the above conditions and $v(x, y, z) = 0$ we can find the value of λ and the values for the variables x, y and z which will give the extreme value of the function $u(x, y, z)$.

Note: If n constraints are given in the problem we have to use n multipliers namely $\lambda_1, \lambda_2, \ldots, \lambda_n$. The drawback of this method is that we cannot decide the nature of the stationary value. Sometimes physical considerations help us to decide whether u has attained a maximum or minimum value.

Example 3.29: Using Lagrange's multipliers method, find the extreme value of $x^2 + y^2 + z^2$ subject to the condition $ax + by + cz = p$.

Solution: Let $u = x^2 + y^2 + z^2$; $v = ax + by + cz - p$

$du = 0$, $dv = 0$ give

$$2x\,dx + 2y\,dy + 2z\,dz = 0 \quad \ldots(1)$$

$$a\,dx + b\,dy + c\,dz = 0 \quad \ldots(2)$$

Equation (1) + λ Equation (2) = 0 gives,

$$(2x+\lambda a)\,dx+(2y+\lambda b)\,dy+(2z+\lambda c)\,dz=0$$

For an extreme value of u we must have,

$$2x+\lambda a=0; \qquad 2y+\lambda b=0; \qquad 2z+\lambda c=0$$

$$\therefore \qquad \lambda = -\frac{2x}{a}=-\frac{2y}{b}=-\frac{2z}{c}$$

Or $$\frac{x}{a} = \frac{y}{b}=\frac{z}{c}$$

$$\frac{ax}{a^2}=\frac{by}{b^2}=\frac{cz}{c^2}=\frac{ax+by+cz}{a^2+b^2+c^2}=\frac{p}{a^2+b^2+c^2}$$

Hence, $$x=\frac{ap}{a^2+b^2+c^2}; \qquad y=\frac{bp}{a^2+b^2+c^2}; \qquad z=\frac{cp}{a^2+b^2+c^2}$$

Using these we get the extreme value of u as,

$$u = \frac{p^2}{a^2+b^2+c^2}.$$

Example 3.30: Find the maxima or minima of $x^m y^n z^p$ subject to the condition

$$ax + by + cz = p + q + r.$$

Solution: Let $u = x^m y^n z^p$; $v = ax + by + cz - p - q - r$

$$\log u = m \log x + n \log y + p \log z$$

$$\frac{1}{u} du = \frac{m}{x} dx + \frac{n}{y} dy + \frac{p}{z} dz$$

$du = 0$ gives $$\frac{m}{x} dx + \frac{n}{y} dy + \frac{p}{z} dz = 0 \quad \ldots(1)$$

$dv = 0$ gives $$a\,dx + b\,dy + c\,dz = 0 \quad \ldots(2)$$

Equation (1) + λ Equation (2) = 0 gives,

$$\left(\frac{m}{x} + a\lambda\right) dx + \left(\frac{n}{y} + \lambda b\right) dy + \left(\frac{p}{z} + \lambda c\right) dz = 0$$

For an extreme value of u, we have,

$$\frac{m}{x} + a\lambda = 0; \qquad \frac{n}{y} + \lambda b = 0; \qquad \frac{p}{z} + \lambda c = 0$$

$$\lambda = -\frac{m}{ax} = -\frac{n}{by} = -\frac{p}{cz}$$

(or) $$\frac{m}{ax} = \frac{n}{by} = \frac{p}{cz} = \frac{m+n+p}{ax+by+cz} = \frac{m+n+p}{p+q+r}$$

$$x = \frac{m(p+q+r)}{a(m+n+p)}; \qquad y = \frac{n(p+q+r)}{b(m+n+p)}; \qquad z = \frac{p(p+q+r)}{c(m+n+p)}$$

Using these values in u, we get the extreme value of u as,

$$u = \frac{m^m n^n p^p}{a^m b^n c^p} \left(\frac{p+q+r}{m+n+p}\right)^{m+n+p}.$$

3.13 DIFFERENTIAL FORMS AND STOKES' THEOREM

In the analysis of different mathematical fields, differential forms are defined as the specific method for multivariable calculus that is independent of coordinates. Differential forms use an integrated methodology for defining the integrands over curves, surfaces, solids and higher dimensional manifolds. The contemporary notion of differential forms was established and pioneered by Élie Cartan.

Fundamentally, considering the one variable calculus the expression $f(x)\,dx$ is an example of a *l-form* and can be integrated over an oriented interval $[a, b]$ in the domain of f and can be represented as,

$$\int_a^b f(x)\,dx.$$

Consider the following expression,

$f(x, y, z)\, dx \wedge dy + g(x, y, z)\, dz \wedge dx + h(x, y, z)\, dy \wedge dz$

This expression is referred as a **2-form** which has a surface integral over an oriented surface S and is represented as,

$$\int_S (f(x,y,z)\,dx \wedge dy + g(x,y,z)\,dz \wedge dx + h(x,y,z)\,dy \wedge dz)$$

The symbol '$\wedge$' denotes the exterior product and is occasionally termed as the **wedge product** of two differential forms.

Similarly, a **3-form** $f(x, y, z)\, dx \wedge dy \wedge dz$ represents a volume element which can be integrated over an oriented region of space. Generally, a k-form is considered as an object that may be integrated over a k-dimensional oriented manifold and is homogeneous of degree k in the coordinate differentials.

On an n-dimensional manifold, the top dimensional form or the 'n-form' is termed as a **volume form**.

A differential k-form can be integrated over an oriented manifold of dimension k. Additionally, a differential l-form is described as measuring an infinitesimal oriented length or l-dimensional oriented density; a differential 2-form is described as measuring an infinitesimal oriented area or 2-dimensional oriented density, and so on.

Integration of differential forms is distinctly defined only on oriented manifolds, for example of a l-dimensional manifold is an interval $[a, b]$ and intervals can be given an orientation as they are positively oriented if $a < b$ and negatively oriented otherwise.

If $a < b$ then the integral of the differential l-form $f(x)\, dx$ over the interval $[a, b]$ in conjunction with its natural or normal positive orientation is given as,

$$\int_a^b f(x)\,dx$$

Which is, considered as the negative of the integral of the similar differential form over the same interval when provided with the opposite orientation.

Specifically,

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$$

A standard explanation in one variable integration theory states that when the limits of integration are in the opposite order ($b < a$), then the increment dx is negative in the direction of integration.

NOTES

3.13.1 Stokes' Theorem

Stokes' theorem, also known as Kelvin–Stokes theorem is named after the Lord Kelvin and George Stokes. The Stokes' theorem is considered as the fundamental theorem for curls or simply the curl theorem and is a theorem in vector calculus on $\mathbb{R}^3$.

For a given vector field, the theorem relates the integral of the curl of the vector field over some surface to the line integral of the vector field around the boundary of the surface.

The classical definition of Stokes' theorem states that, "The line integral of a vector field over a loop is equal to the flux of its curl through the enclosed surface".

Stokes' theorem is considered as a unique and specific instance of the generalized Stokes' theorem. Specifically, a vector field on $\mathbb{R}^3$ can be considered as a *l*-form in which instance its curl is its exterior derivative defined as a 2-form.

Theorem 3.14 Stokes' Theorem: Let Σ be a smooth oriented surface in $\mathbf{R}^3$ with boundary $\partial\Sigma$.

If a vector field $\mathbf{A}$ is defined as,

$$\mathbf{A} = (P(x, y, z), Q(x, y, z), R(x, y, z))$$

This typically has continuous first order partial derivatives in a region containing Σ, then,

$$\iint_{\Sigma} (\nabla \times \mathbf{A}) \cdot \mathbf{da} = \oint_{\partial\Sigma} \mathbf{A} \cdot \mathbf{dl}.$$

More explicitly, the equality states that,

$$\iint_{\Sigma} \left(\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \mathrm{d}y\, \mathrm{d}z + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \mathrm{d}z\, \mathrm{d}x + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathrm{d}x\, \mathrm{d}y \right)$$
$$= \oint_{\partial\Sigma} \Big(P\, \mathrm{d}x + Q\, \mathrm{d}y + R\, \mathrm{d}z \Big).$$

The key challenge in a precise and accurate statement of Stokes' theorem is in defining the notion of a boundary.

Proofs

1. Parametrization of Integral

We first reduce the dimension by using the natural parametrization of the surface.

Let ψ and γ be in the section and consider by change of variables,

$$\oint_{\partial\Sigma} \mathbf{F}(\mathbf{x}) \cdot \mathrm{d}\mathbf{l} = \oint_{\gamma} \mathbf{F}(\boldsymbol{\psi}(\mathbf{y})) \cdot \mathrm{d}\boldsymbol{\psi}(\mathbf{y}) = \oint_{\gamma} \mathbf{F}(\boldsymbol{\psi}(\mathbf{y})) J_{\mathbf{y}}(\boldsymbol{\psi})\, \mathrm{d}\mathbf{y}$$

Here $J\psi$ stands for the Jacobian matrix of ψ.

Now consider that $\{\mathbf{e}_u, \mathbf{e}_v\}$ be an orthonormal basis in the coordinate directions of $\mathbf{R}_2$. Distinguishing that the columns of $J_{\mathbf{y}}\psi$ are precisely the partial derivatives of ψ at $\mathbf{y}$, we can expand the previous equation in coordinates as,

$$\begin{aligned}\oint_{\partial\Sigma} \mathbf{F}(\mathbf{x}) \cdot \mathrm{d}\mathbf{l} &= \oint_{\gamma} \mathbf{F}(\boldsymbol{\psi}(\mathbf{y}))J_{\mathbf{y}}(\boldsymbol{\psi})\mathbf{e}_u(\mathbf{e}_u \cdot \mathrm{d}\mathbf{y}) + \mathbf{F}(\boldsymbol{\psi}(\mathbf{y}))J_{\mathbf{y}}(\boldsymbol{\psi})\mathbf{e}_v(\mathbf{e}_v \cdot \mathrm{d}\mathbf{y}) \\ &= \oint_{\gamma} \left(\left(\mathbf{F}(\boldsymbol{\psi}(\mathbf{y})) \cdot \frac{\partial \boldsymbol{\psi}}{\partial u}(\mathbf{y}) \right) \mathbf{e}_u + \left(\mathbf{F}(\boldsymbol{\psi}(\mathbf{y})) \cdot \frac{\partial \boldsymbol{\psi}}{\partial v}(\mathbf{y}) \right) \mathbf{e}_v \right) \cdot \mathrm{d}\mathbf{y}\end{aligned}$$

2. Green's Theorem through the Product Rule

We first calculate the partial derivatives that appear in Green's theorem through the product rule:

$$\begin{aligned}\frac{\partial P_1}{\partial v} &= \frac{\partial(\mathbf{F} \circ \boldsymbol{\psi})}{\partial v} \cdot \frac{\partial \boldsymbol{\psi}}{\partial u} + (\mathbf{F} \circ \boldsymbol{\psi}) \cdot \frac{\partial^2 \boldsymbol{\psi}}{\partial v\,\partial u} \\ \frac{\partial P_2}{\partial u} &= \frac{\partial(\mathbf{F} \circ \boldsymbol{\psi})}{\partial u} \cdot \frac{\partial \boldsymbol{\psi}}{\partial v} + (\mathbf{F} \circ \boldsymbol{\psi}) \cdot \frac{\partial^2 \boldsymbol{\psi}}{\partial u\,\partial v}\end{aligned}$$

Appropriately, the second term vanishes in the difference and by equality of mixed partials. Therefore,

$$\begin{aligned}\frac{\partial P_1}{\partial v} - \frac{\partial P_2}{\partial u} &= \frac{\partial(\mathbf{F} \circ \boldsymbol{\psi})}{\partial v} \cdot \frac{\partial \boldsymbol{\psi}}{\partial u} - \frac{\partial(\mathbf{F} \circ \boldsymbol{\psi})}{\partial u} \cdot \frac{\partial \boldsymbol{\psi}}{\partial v} \\ &= \frac{\partial \boldsymbol{\psi}}{\partial u}(J_{\boldsymbol{\psi}(u,v)}\mathbf{F})\frac{\partial \boldsymbol{\psi}}{\partial v} - \frac{\partial \boldsymbol{\psi}}{\partial v}(J_{\boldsymbol{\psi}(u,v)}\mathbf{F})\frac{\partial \boldsymbol{\psi}}{\partial u} \qquad \text{(chain rule)} \\ &= \frac{\partial \boldsymbol{\psi}}{\partial u}\left(J_{\boldsymbol{\psi}(u,v)}\mathbf{F} - (J_{\boldsymbol{\psi}(u,v)}\mathbf{F})^{\mathsf{T}}\right)\frac{\partial \boldsymbol{\psi}}{\partial v}\end{aligned}$$

3. Proof through Differential Forms

The functions $\mathbf{R} \to \mathbf{R}^3$ can be identified with the differential l-forms on $\mathbf{R}^3$ through the map as follows,

$$F_1\mathbf{e}_1 + F_2\mathbf{e}_2 + F_3\mathbf{e}_3 \mapsto F_1\,\mathrm{d}x + F_2\,\mathrm{d}y + F_3\mathrm{d}z$$

Now we write the differential l-form that is associated or connected to a function $\mathbf{F}$ as $\omega_{\mathbf{F}}$. Then it can be calculated as,

$$\star\omega_{\nabla\times\mathbf{F}} = \mathrm{d}\omega_{\mathbf{F}}$$

Where * is referred as the Hodge star and {d} is the exterior derivative. Thus, by generalized Stokes' theorem,

$$\oint_{\partial\Sigma} \mathbf{F} \cdot \mathrm{d}\mathbf{l} = \oint_{\partial\Sigma} \omega_{\mathbf{F}} = \int_{\Sigma} \mathrm{d}\omega_{\mathbf{F}} = \int_{\Sigma} \star\omega_{\nabla\times\mathbf{F}} = \iint_{\Sigma} \nabla \times \mathbf{F} \cdot \mathrm{d}^2\mathbf{S}$$

3.14 JACOBIAN AND ITS PROPERTIES

NOTES

If $u(x, y)$ and $v(x, y)$ are two functions of x and y then the determinant $\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$

is called the Jacobian of u and v with respect to x and y and is denoted as $\dfrac{\partial(u, v)}{\partial(x, y)}$ or $|J|$.

In general if $u_1, u_2, \ldots, u_n$ are functions of $x_1, x_2, \ldots, x_n$ then,

$$\frac{\partial(u_1, u_2, \ldots u_n)}{\partial(x_1, x_2, \ldots x_n)} = \begin{vmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \vdots & & & \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{vmatrix}.$$

Properties of Jacobians

1. If u and v are functions of x and y then if,

$$J = \frac{\partial(u, v)}{\partial(x, y)} \quad \text{and} \quad J' = \frac{\partial(x, y)}{\partial(u, v)} \quad \text{then} \quad J J' = 1$$

$$J = \frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$$

$$J' = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

$$J J' = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial u} & \frac{\partial u}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial u}{\partial y}\cdot\frac{\partial y}{\partial v} \\ \frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial u} & \frac{\partial v}{\partial x}\cdot\frac{\partial x}{\partial v}+\frac{\partial v}{\partial y}\cdot\frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial u} & \frac{\partial u}{\partial v} \\ \frac{\partial v}{\partial u} & \frac{\partial v}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

2. If u and v are functions of r and s where r and s are functions of x and y, prove that

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial(u, v)}{\partial(r, s)}\cdot\frac{\partial(r, s)}{\partial(x, y)}.$$

$$\frac{\partial(u, v)}{\partial(r, s)}\cdot\frac{\partial(r, s)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial s} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial s} \end{vmatrix}\begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial u}{\partial s}\cdot\frac{\partial s}{\partial x} & \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial y}+\frac{\partial u}{\partial s}\cdot\frac{\partial s}{\partial y} \\ \frac{\partial v}{\partial r}\cdot\frac{\partial r}{\partial x}+\frac{\partial v}{\partial s}\cdot\frac{\partial s}{\partial x} & \frac{\partial v}{\partial r}\cdot\frac{\partial r}{\partial y}+\frac{\partial v}{\partial s}\cdot\frac{\partial s}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial(u, v)}{\partial(x, y)}.$$

Example 3.31: If $u = \frac{x^2}{y}$, $v = \frac{y^2}{x}$ find $\frac{\partial(x, y)}{\partial(u, v)}$.

Solution: Since it is easy to find $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ let us find $\frac{\partial(u, v)}{\partial(x, y)}$.

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{2x}{y} & -\frac{x^2}{y^2} \\ -\frac{y^2}{x^2} & \frac{2y}{x} \end{vmatrix}$$

$$= 4 - 1 = 3.$$

$$\therefore \quad \frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{\frac{\partial(u, v)}{\partial(x, y)}} = \frac{1}{3}.$$

NOTES

Example 3.32: If $u = x\,(1 - y)$, $v = xy\,(1 - z)$, $w = xyz$, prove that,

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \frac{1}{x^2 y}.$$

Solution:
$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} = \begin{vmatrix} 1-y & -x & 0 \\ y(1-z) & x(1-z) & -xy \\ yz & xz & xy \end{vmatrix}$$

$$= x^2 y.$$

$$\therefore \quad \frac{\partial(x, y, z)}{\partial(u, v, w)} = \frac{1}{x^2 y}.$$

Check Your Progress

11. State the theorem of change of variables.
12. Give the necessary and sufficient conditions for maximum and minimum.
13. How many multipliers are used in the Lagrange's multipliers method if there are n constraints?
14. State Stoke's theorem.
15. Define the term Jacobian.

3.15 ANSWERS TO 'CHECK YOUR PROGRESS'

1. If $z = f(x, y)$, then z is a dependent variable and x, y are independent variables. The function $z = f(x, y)$ is called a single-valued function if only one value of z is corresponded by each pair (x, y) for which the function is defined. If there is more than one value of z, the function is called a multi-valued function.
2. Let V and U be two vector spaces over the same field F, then a mapping $T : V \to U$ is called a homomorphism or a linear transformation if,

   $T(x + y) = T(x) + T(y)$ for all $x, y \in V$

   $T(\alpha x) = \alpha T(x)$, $\alpha \in F$
3. If f is differentiable at c, then f is continuous at c.
4. Let $f(x, y)$ be a function of two independent variables x and y, defined in the region R. The function $f(x, y)$ is a function of a single variable x when y is constant. Then the derivative of $f(x, y)$ (when exists) is called the partial derivative of $f(x, y)$ with respect to x.

5. A function $f(x, y)$ is said to be homogeneous of degree n in the variables x and y if it can be expressed in the form $x^n\ \phi(y/x)$ or in the form $y^n\ \phi(x/y)$.
6. A function $f(x, y)$ is said to be a harmonic function if $\nabla^2 f = \dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} = 0$.
7. Let u and v be two functions of x, both derivable at least upto n times, then the nth derivative of their product is given by,

$$\begin{aligned}(uv)_n &= \sum_{r=0}^{n} {}^n c_r\ u_{n-r}\ v_r \\ &= {}^n c_o\ u_n v + {}^n c_1\ u_{n-1}\ v_1 + \ldots + {}^n c_n\ u\ v_n \\ &= u_n\ v + n\ u_{n-1}\ v_1 + \frac{n(n-1)}{\lfloor 2}\ u_{n-2}\ v_2 + \ldots + u\ v_n\end{aligned}$$

   Where the suffixes of u and v denote the orders of differentiation of u and v with respect to x.
8. Let $f(x+h)$ be expandable into a power series in the variable h. Again the flat assumption is that this series can be differentiated term by term. Taylor's theorem states that,

$$f(a+h) = f(a) + h f'(a) + \frac{h^2}{2!} f''(a) + \ldots \infty$$

9. If the dependent variable y is expressed in terms of the independent variable x, we call y an explicit function of x and denote such a function by $y = f(x)$.
10. Let $f = (f_1, f_2, \ldots, f_n) \in C'$ on an open set S in R^n, and let $T = f(s)$. If the Jacobian determinant $J_f(a) \neq 0$ for some point a in S, then there are two open sets $X \subseteq S$ and $Y \subseteq T$ and a uniquely determined function g such that,
    (*a*) $a \in X$ and $f(a) \in Y$.
    (*b*) $Y = f(X)$.
    (*c*) F is one to one on X.
    (*d*) G is defined on Y, $g(Y) = X$, and $g[f(X)]$ for every $x \in X$.
    (*e*) $g \in C'$.
11. If (*i*) $f \in \Re\,[a, b]$, (*ii*) ϕ is derivable, strictly monotonic on $[\alpha, \beta]$ and maps it onto $[a, b]$, and (*iii*) $\phi' \in \Re\,[\alpha, \beta]$, then $\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(t)\phi'(t)dt$.
12. If $f(x)$ be a maximum or a minimum at $x = c$ and if $f'(c)$ exists, then $f'(c) = 0$.
13. If n constraints are given in the problem we have to use n multipliers namely $\lambda_1, \lambda_2, \ldots, \lambda_n$.

NOTES

14. Stokes' theorem, also known as Kelvin–Stokes theorem is named after the Lord Kelvin and George Stokes. The Stokes' theorem is considered as the fundamental theorem for curls or simply the curl theorem and is a theorem in vector calculus on $\mathbb{R}^3$. For a given vector field, the theorem relates the integral of the curl of the vector field over some surface to the line integral of the vector field around the boundary of the surface.

    The classical definition of Stokes' theorem states that, "The line integral of a vector field over a loop is equal to the flux of its curl through the enclosed surface".

    Stokes' theorem is considered as a unique and specific instance of the generalized Stokes' theorem. Specifically, a vector field on $\mathbb{R}^3$ can be considered as a *l*-form in which instance its curl is its exterior derivative defined as a 2-form.

15. If $u(x, y)$ and $v(x, y)$ are two functions of x and y then the determinant

    $$\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$$ is called the Jacobian of u and v with respect to x and y and is

    denoted as $\dfrac{\partial(u, v)}{\partial(x, y)}$ or $|J|$.

3.16 SUMMARY

- If $z = f(x, y)$, then z is a dependent variable and x, y are independent variables. The function $z = f(x, y)$ is called a single-valued function if only one value of z is corresponded by each pair (x, y) for which the function is defined. If there is more than one value of z, the function is called a multi-valued function.
- Let V and U be two vector spaces over the same field F, then a mapping $T : V \to U$ is called a homomorphism or a linear transformation if,

  $T(x + y) = T(x) + T(y)$ for all $x, y \in V$

  $T(\alpha x) = \alpha T(x)$, $\alpha \in F$
- Let $f{:}S \to R^m$ be a function defined on a set S in R^n with values in R^m. Let c be an interior point of S and let $B(c;r)$ be an n-ball lying in S. Let V be a point in R^n with $\|v\| < r$, so that $c+v \in B(c;r)$. Then the function f is said to be differentiable at c if there exists a linear function $T_c: R^n \to R^m$ such that, $f(c+v) = f(c) + T_c(v) + \|v\| E_c(v)$, where $E_c(v) \to 0$ as $v \to 0$
- Let $f(x, y)$ be a function of two independent variables x and y, defined in the region R. The function $f(x, y)$ is a function of a single variable x when

y is constant. Then the derivative of $f(x, y)$ (when exists) is called the partial derivative of $f(x, y)$ with respect to x.

- A function $f(x, y)$ is said to be homogeneous of degree n in the variables x and y if it can be expressed in the form $x^n\ \phi(y/x)$ or in the form $y^n\ \phi\ (x/y)$.
- A function $f(x, y)$ is said to be a harmonic function if,

  $$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

- Let $z = f(x, y)$ be a function of two independent variables x and y and f_x, f_y exist at (x, y), then $dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ is called the total differential of z.
- The nth order derivative of $f(x)$ with respect to x is denoted by y_n or $f^{(n)}\ (x)$ or $\frac{d^n y}{dx^n}$ or $y^{(n)}$ or $\frac{d^n}{dx^n}\{f(x)\}$ or $D^n f(x)$ whre $D \equiv \frac{d}{dx}$ and $y = f(x)$.
- Let u and v be two functions of x, both derivable at least upto n times, then the nth derivative of their product is given by,

  $$\begin{aligned}(uv)_n &= \sum_{r=0}^{n} {}^n c_r\ u_{n-r}\ v_r \\ &= {}^n c_o\ u_n v + {}^n c_1\ u_{n-1}\ v_1 + \ldots + {}^n c_n\ u\ v_n \\ &= u_n\ v + n\ u_{n-1}\ v_1 + \frac{n(n-1)}{\lfloor 2}\ u_{n-2}\ v_2 + \ldots + u\ v_n\end{aligned}$$

  Where the suffixes of u and v denote the orders of differentiation of u and v with respect to x.

- Let $f(x+h)$ be expandable into a power series in the variable h. Again the flat assumption is that this series can be differentiated term by term. Taylor's theorem states that, $f(a+h) = f(a) + h f'(a) + \frac{h^2}{2!} f''(a) + \ldots\ \infty$
- If the dependent variable y is expressed in terms of the independent variable x, we call y an explicit function of x and denote such a function by $y = f(x)$.
- Let $f = (f_1, f_2, \ldots, f_n) \in C'$ on an open set S in R^n, and let $T = f(s)$. If the Jacobian determinant $J_f(a) \neq 0$ for some point a in S, then there are two open sets $X \subseteq S$ and $Y \subseteq T$ and a uniquely determined function g such that,

  (*a*) $a \in X$ and $f(a) \in Y$.

  (*b*) $Y = f(X)$.

NOTES

(*c*) F is one to one on X.

(*d*) G is defined on Y, $g(Y) = X$, and $g[f(X)]$ for every $x \in X$.

(*e*) $g \in C'$.

- Let $f = (f_1,\ldots,f_n)$ be a vector valued function defined on an open set S in R^{n+k} with values in R^n. Suppose $f \in C^1$ on S. Let $(x_0;t_0)$ be a point in S for which the $n \times n$ determinant $\det[D_j f_i(x_0;t_0)] \neq 0$. Then there exists a k dimensional open set T_0 containing t_0 and, one and only one, vector valued function g, defined on T_0 and having values in R^n, such that

  (*a*) $g \in C^1$ on T_0

  (*b*) $g(t_0) = x_0$

  (*c*) $f(g(t;t)) = 0$ for every t in T_0

- If (*i*) $f \in \Re\,[a, b]$, (*ii*) ϕ is derivable, strictly monotonic on $[\alpha, \beta]$ and maps it onto $[a, b]$, and (*iii*) $\phi' \in \Re\,[\alpha, \beta]$, then

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(t)\phi'(t)dt.$$

- A function $y = f(x)$ has a maximum value at a point x_0 if for $|\,h\,|$ sufficiently small, $f(x_0 + h) < f(x_0)$.

  Similarly, a function $y = f(x)$ has a minimum value at a point x_0 if,

$$f(x_0 + h) > f(x_0)$$

- If we have to find the stationary value of a function of several variables which are not independent but interconnected by some relations then we try to convert the given function to one having least number of variables using the given conditions. When such a procedure fails we use the method of Lagrange's multipliers.

- Differential forms use an integrated methodology for defining the integrands over curves, surfaces, solids and higher dimensional manifolds. The contemporary notion of differential forms was established and pioneered by Élie Cartan.

- Fundamentally, considering the one variable calculus the expression $f(x)\,dx$ is an example of a *l-form* and can be integrated over an oriented interval $[a, b]$ in the domain of f and can be represented as,

$$\int_a^b f(x)\,dx.$$

- Integration of differential forms is distinctly defined only on oriented manifolds, for example of a l-dimensional manifold is an interval $[a, b]$ and intervals can be given an orientation as they are positively oriented if $a < b$ and negatively oriented otherwise.

- If $a < b$ then the integral of the differential l-form $f(x)\, dx$ over the interval $[a, b]$ in conjunction with its natural or normal positive orientation is given as,

$$\int_a^b f(x)\, dx$$

  Which is considered as the negative of the integral of the similar differential form over the same interval when provided with the opposite orientation.
- Stokes' theorem, also known as Kelvin–Stokes theorem is named after the Lord Kelvin and George Stokes. The Stokes' theorem is considered as the fundamental theorem for curls or simply the curl theorem and is a theorem in vector calculus on ”!^{3}.
- For a given vector field, the theorem relates the integral of the curl of the vector field over some surface to the line integral of the vector field around the boundary of the surface.
- The classical definition of Stokes' theorem states that, "The line integral of a vector field over a loop is equal to the flux of its curl through the enclosed surface".
- Stokes' theorem is considered as a unique and specific instance of the generalized Stokes' theorem. Specifically, a vector field on ”!^{3} can be considered as a l-form in which instance its curl is its exterior derivative defined as a 2-form.
- Stokes' Theorem: Let Σ be a smooth oriented surface in $\mathbf{R}^3$ with boundary $\partial\Sigma$.
- If $u\,(x, y)$ and $v\,(x, y)$ are two functions of x and y then the determinant $\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix}$ is called the Jacobian of u and v with respect to x and y and is denoted as $\dfrac{\partial(u, v)}{\partial(x, y)}$ or $|\,J\,|$.
- In general if $u_1, u_2, \ldots, u_n$ are functions of $x_1, x_2, \ldots, x_n$ then,

$$\frac{\partial(u_1, u_2, \ldots u_n)}{\partial(x_1, x_2, \ldots x_n)} = \begin{vmatrix} \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \vdots & & & \\ \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{vmatrix}.$$

NOTES

3.17 KEY TERMS

- **Domain of the function:** The set of values (x, y) for which a function is defined.
- **Region:** A connected open set.
- **Homogeneous function:** A function $f(x, y, z, \ldots)$ of degree n in the variables $x, y, z, \ldots$ if $f(tx, ty, tz, \ldots) = t^n(x, y, z, \ldots)$ for all values of t independent of $x, y, z, \ldots$
- **Explicit function:** The dependent variable y which can be expressed in terms of the independent variable x.

3.18 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define rectangular region and circular region.
2. What is the use of linear transformations?
3. Give the first order Taylor formula.
4. Define the second order partial derivative of a function of two variables.
5. State the Euler's theorem on homogeneous functions.
6. Find the nth derivative of e^{-ax}.
7. Write the significance of Taylor's theorem.
8. State the implicit function theorem.
9. What is the use of change of variables technique?
10. Define global maxima and global minima.
11. Give a drawback of Lagrange's multipliers method.
12. What is Jacobian used for?

Long-Answer Questions

1. Show that $\lim\limits_{x \to 0, y \to 0} \dfrac{2xy}{x^2 + y^2}$ does not exist.
2. Show that the function $f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$

   is continuous at (0, 0).
3. Show that any linear transformation $T: R \to R$ is of the form $T(x) = \alpha x$ for some $\alpha \in R$.

4. Explain the derivatives in an open subset of R^n.
5. If $u = \sqrt{xy}$, prove that $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -\frac{1}{4}(x^{-3/2}y^{1/2} + y^{-3/2}x^{1/2})$.
6. Find y_n when $y = e^x \log x$.
7. If $y = 2\cos x(\sin x - \cos x)$, show that $(y_{10})_0 = 2^{10}$.
8. State and prove Taylor's theorem.
9. If $f(x, y) = 0$ and $\phi(y, z) = 0$, show that $\frac{\partial f}{\partial y}\frac{\partial \varphi}{\partial z}\frac{dz}{dx} = \frac{\partial f}{\partial x}\frac{\partial \varphi}{\partial y}$.
10. State and prove implicit function theorem.
11. Evaluate $\int_0^{\pi/2} \sin 2x\, dx$ using change of variables technique.
12. Show that the function $f(x, y) = 4x^2y - y^2 - 8x^4$ is a maximum at (0, 0).
13. By Lagrange method, find the minimum distance of origin from the plane $3x + 2y + z = 12$.
14. If $ax^2 + by^2 = ab$, show that the extrema values of $u = x^2 + y^2 + xy$ are the roots of $4(u - a)(u - b) = ab$.
15. If $x = e^u \cos v$, $y = e^u \sin v$, verify the rule $JJ' = 1$.
16. State and prove Stoke's theorem.
17. Briefly discuss about the jacobians giving appropriate examples.

3.19 FURTHER READING

Rudin, Walter. 2017. *Real and Complex Analysis*, Third Edition. Noida: McGraw-Hill Education.

Gupta, S. L. and Nisha Rani. 2004. *Fundamental Real Analysis*, Fourth Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, First Edition. Cambridge (U.K.): Cambridge University Press.

Bartle, Robert G. and Donald R. Sherbert. 2014. *Introduction to Real Analysis*, Fourth Edition. New York: Wiley.

Trench, William F. 2002. *Introduction to Real Analysis*, London: Pearson.

Loeb, Peter A. 2016. *Real Analysis*, Basel (Switzerland): Birkhäuser.

Royden, Halsey. 2015. *Real Analysis*, Fourth Edition. Noida: Pearson Education India.

NOTES

UNIT 4 LEBESGUE MEASURABILITY AND INTEGRATION OF NON-NEGATIVE FUNCTIONS

Structure

4.0 INTRODUCTION

In the measure theory, the term Lebesgue measure is named after French mathematician Henri Lebesgue. Fundamentally, the Lebesgue measure is defined as the standard method used to assign a measure to subsets of n-dimensional Euclidean space. Additionally, for $n = 1, 2$ or 3, the Lebesgue measure coincides with the standard universal measure of length, area or volume. Generally, in the mathematical analysis the measure is also termed as the n-dimensional volume, n-volume or merely only the volume. In real analysis, the Lebesgue measure is specifically used to define the Lebesgue integration.

In mathematical analysis, the 'Real Analysis' only considers and evaluates the functions of a real variable and numbers, while the 'Measure Theory' exclusively considers and evaluates the concept of a measure, which is the method for 'Measuring' that how big a given set is. Sets to which the Lebesgue measure can be assigned are termed as the 'Lebesgue Measurable'; characteristically the measure of the Lebesgue measurable set A can be denoted by λ(A). A measurable set is specifically defined as a set to which the extension or expansion can possibly be accomplished, this extension or expansion is assumed to be the measure.

Consequently, the Jordan measure, the Borel measure and the Lebesgue measure can be defined using the sets which are characteristically measurable according to the Jordan, Borel and Lebesgue, respectively. In mathematical analysis, a non-measurable set is defined as a set to which a significant meaningful 'Volume' cannot be assigned. Mathematically, in formal set theory the existence of these sets is interpreted for providing evidence about the notions of length, area and volume.

The integral of a non-negative function of a single variable is specifically defined as the area between the graph of that function and the X-axis. Fundamentally, the Lebesgue integral can be used for extending the integral to a bigger class or group of functions. Additionally, it can also extend or broaden the domains for defining and approximating these functions. Many years before the 20th century, the mathematicians were already aware of the theory that when the non-negative functions have a smooth adequate graph, basically the continuous functions on the closed bounded intervals, then the specific area under the curve can be defined as the integral and can be then uniquely computed with the help of the approximation techniques specifically on the region through polygons.

The Lebesgue integral is named after Henri Lebesgue (1875–1941), typically Lebesgue defined and established the integral in the year 1904. Principally, the Lebesgue integral functions have a significant role in the theory and derivation of probability, real analysis, and numerous other fields in mathematics. Mathematically, as per the Lebesgue explanation the term Lebesgue integration specifies either the general theory of integration of a function with respect to a general measure or the specific instance of integration of a function typically defined on a sub-domain of the real line with respect to the Lebesgue measure. The Riemann integral specifies that by partitioning the domain of an assigned function, one can approximate or estimate the assigned function by means of piecewise constant functions in each sub-interval. On the contrary, the Lebesgue integral are specifically used to partition the range of that function.

The Lebesgue differentiation theorem states that, "For almost every point, the value of an integrable function is the limit of infinitesimal averages taken about the point". The theorem is explicitly used in the approximation of real analysis.

In the mathematical analysis, a function of Bounded Variation (BV) also termed as BV function, is considered as a real valued function whose total variation is bounded or finite. Considering a continuous function of a single variable, which has bounded variation signifies that the distance along the direction of the Y-axis ignoring the contribution of motion along X-axis, travelled by a point moving along the graph has a finite value. Similarly, consider a continuous function of several variables, the connotation and implication of the definition is equivalent, except that the considered continuous path cannot be the whole graph of the given function, but can be every intersection of the graph itself with a hyperplane (for several variables) and plane (for functions of two variables) parallel to a fixed X-axis and to the Y-axis.

The key objective of the Lebesgue integral is to provide an integral notion in which the limits of integrals hold moderate assumptions. Basically, there is no assurance that every function is the Lebesgue integrable, but it is possible that improper integrals exist for functions that are not Lebesgue integrable.

In this unit, you will study about the Lebesgue outer measure, measurable sets, regularity, measurable functions, Borel and Lebesgue measurability, non-measurable sets, integration of non-negative functions, the general integral, integration of series, Riemann and Lebesgue integrals, the four derivatives, functions of bounded variation, Lebesgue differentiation theorem, and differentiation and integration.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Define the Lebesgue outer measure
- Know about the measurable sets, regularity, and measurable functions
- Understand the Borel and Lebesgue measurability and non-measurable sets
- Define the integration of non-negative functions
- Elaborate on the general integral
- Explain about the integration of series
- Analyse the Reimann and Lebesgue integrals
- Comprehend on the four derivatives
- Interpret the functions of bounded variation
- Discuss the Lebesgue differentiation theorem
- Understand the differentiation and integration

4.2 LEBESGUE OUTER MEASURE

In the measure theory, the term **Lebesgue measure** is named after French mathematician Henri Lebesgue. This measure was described by Henri Lebesgue in the year 1901, and in the year 1902 by the description **Lebesgue integral**. Both the Lebesgue measure and the Lebesgue integral were published in his dissertation thesis in the year 1902.

Fundamentally, the Lebesgue measure is defined as the standard method used to assign a measure to subsets of n-dimensional Euclidean space. Additionally, for $n = 1, 2$ or 3, the Lebesgue measure coincides with the standard universal measure of length, area or volume. Generally, in the mathematical analysis the measure is also termed as the n-dimensional volume, n-volume or merely only the volume. In real analysis, the Lebesgue measure is specifically used to define the Lebesgue integration.

In mathematical analysis, the 'Real Analysis' only considers and evaluates the functions of a real variable and numbers, while the 'Measure Theory' exclusively considers and evaluates the concept of a measure, which is the method for 'Measuring' that how big a given set is. Sets to which the Lebesgue measure can be assigned are termed as the 'Lebesgue Measurable'; characteristically, the measure of the Lebesgue measurable set A can be denoted by $\lambda(A)$.

Lebesgue Outer Measure

Assume that there is an **outer measure** λ on a set X which is a measure then it will be considered as **additive**. Specifically, for given any two sets $A, B \subseteq X$ we can state that both $A \cap B$ and $A \cap B^c$ are disjoint in conjunction with $(A \cap B) \cup (A \cap B^c) = A$ and accordingly we can state that,

$$\lambda\,(A) = \lambda\,(A \cap B) + \lambda\,(A \cap B^c)$$

This notation essentially may not hold for all A and B but it is very significant as it specifies the following definition.

Definition 1: Let λ be an outer measure on a set X. Then $E \subseteq X$ is said to be measurable with respect to λ or λ-measurable if,

$$\lambda\,(A) = \lambda\,(A \cap E) + \lambda\,(A \cap E^c) \text{ for all } A \subseteq X$$

This specifies that taking each and every feasible 'Test Set' as 'A', the measures of the parts of A that uniquely fall within and without E can be checked that whether these defined measures feasibly add up to A or not.

Because λ is subadditive, hence we can specify that,

$$\lambda\,(A) \leq \lambda\,(A \cap E) + \lambda\,(A \cap E^c)$$

Consequently, to check measurability we need only verify that,

$$\lambda\,(A) \geq \lambda\,(A \cap E) + \lambda\,(A \cap E^c) \text{ for all } A \subseteq X \qquad \dots(4.1)$$

Let $\mathcal{M} = \mathcal{M}\,(\lambda)$ denote the collection of λ-measurable sets.

Theorem 4.1: Show that $\mathcal{M}$ is a field.

Proof: Trivially, we can state that ϕ and X are in $\mathcal{M}$.

Now consider any $E_1, E_2 \in \mathcal{M}$ and also any test set as $A \subseteq X$.

Then,

$$\lambda(A) = \lambda(A \cap E_1) + \lambda(A \cap E_1^c)$$

On applying the measurability definition for E_2 with the test set as $A \cap E_1^c$ we obtain,

$$\begin{aligned}\lambda\,(A \cap E_1^c) &= \lambda\,((A \cap E_1^c) \cap E_2) + \lambda\,((A \cap E_1^c) \cap E_2^c)\\ &= \lambda\,(A \cap E_1^c \cap E_2) + \lambda\,(A \cap (E_1 \cup E_2)^c)\end{aligned}$$

When the above equations are combined then we have,

$$\lambda\,(A) = \lambda\,(A \cap E_1) + \lambda\,(A \cap E_1^c \cap E_2) + \lambda\,(A \cap (E_1 \cup E_2)^c) \quad \dots(4.2)$$

On the Right Hand Side (RHS) to the first two terms of Equation (4.2), we use the subadditive property of λ. Now for the sets we obtain,

$$\begin{aligned}(A \cap E_1) \cup (A \cap E_1^c \cap E_2) &= A \cap (E_1 \cup (E_1^c \cap E_2))\\ &= A \cap ((E_1 \cup E_1^c) \cap (E_1 \cup E_2))\\ &= A \cap (X \cap (E_1 \cup E_2))\\ &= A \cap (E_1 \cup E_2)\end{aligned}$$

Subsequently,

$$\lambda\,(A \cap E_1) + \lambda\,(A \cap E_1^c \cap E_2) \geq \lambda\,(A \cap (E_1 \cup E_2))$$

We substitute this in Equation (4.2) to obtain,

$\lambda(A) \geq \lambda(A \cap (E_1 \cup E_2)) + \lambda(A \cap (E_1 \cup E_2)^c)$

Consequently, the Equation (4.1) has been authenticated and verified for the notation $E_1 \cup E_2$, i.e., $E_1 \cup E_2 \in \mathcal{M}$.

Note that the given definition of λ-measurable sets exhibits symmetric because $E \in \mathcal{M}$ if and only if $E^c \in \mathcal{M}$.

Therefore, $\quad E_1 \setminus E_2 = E_1 \cap E_2^c = (E_1^c \cup E_2)^c \in \mathcal{M}$

Hence proved that $\mathcal{M}$ is a field.

Definition 2: Let E be a subset of $\mathbb{R}$ and also assume that $\{I_k\}$ is a sequence of open intervals. Then we can define the Lebesgue outer measure of E as follows,

$$\mu^*(\mathrm{E}) = \inf \left\{ \sum_{k=1}^{\infty} \ell(I_k) : E \subseteq \bigcup_{k=1}^{\infty} I_k \right\}$$

Remember that $0 \leq \mu^*(\mathrm{E}) \leq \infty$.

Theorem 4.2: The Lebesgue outer measure holds the following properties:

(1) If $E_1 \subseteq E_2$, then $\mu^*(E_1) \leq \mu^*(E_2)$.

(2) The Lebesgue outer measure of any countable set is zero.

(3) The Lebesgue outer measure of the empty set is zero.

(4) Lebesgue outer measure is invariant under translation, that is,

$\mu^*(E + x_0) = \mu^*(E)$

(5) Lebesgue outer measure is countably subadditive, i.e.,

$$\mu^*\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$$

Proof

For (1) – Property (1) is trivial.

For (2) and (3) – To prove Property (2) and Property (3) consider the following:

Assume that $E = \{x_k : k \in Z^+\}$ is a countably infinite set.

Let $\varepsilon > 0$ and also assume that ε_k is a sequence of positive numbers such that,

$$\sum_{k=1}^{\infty} \epsilon_k = \frac{\epsilon}{2}$$

Because,

$$E \subseteq \bigcup_{k=1}^{\infty} (x_k - \epsilon_k, x_k + \epsilon_k)$$

This specifies that $\boldsymbol{\mu^*(E)} \leq \varepsilon$.

Consequently, $\mu^*(E) = 0$. Because if $\varnothing \subseteq E$, then $\mu^*(\varnothing) = 0$.

For (4) – To prove Property (4) consider the following:

Assume that if each covering of E through open intervals is capable of generating a cover of $E + x_0$ by means of open intervals through the same length, then $\mu*(E + x_0) \leq \mu*(E)$. Also, $\mu*(E + x_0) \geq \mu*(E)$ because E is considered a translation of $E + x_0$. Consequently, $\mu*(E + x_0) = \mu*(E)$.

For (5) – To prove Property (5) consider the following:

Assume that,

$$\sum_{i=1}^{\infty} \mu^*(E_i) = \infty \ \{I_k^i\}$$

Then the given statement is insignificant or trivial.

Assume that the sum is finite and also let $\varepsilon > 0$.

Then for each i, there exists a sequence of the form $\{I_k^i\}$ of open intervals such that there is,

$$E_i \subseteq \bigcup_{k=1}^{\infty} I_k^i \quad \text{and} \quad \sum_{k=1}^{\infty} \ell(I_k^i) < \mu^*(E_i) + \tfrac{\epsilon}{2^i}$$

Consider that $\{I_k^i\}$ is a double indexed sequence of open intervals such that,

$$\bigcup_{i=1}^{\infty} E_i \subseteq \bigcup_{i=1}^{\infty} \bigcup_{k=1}^{\infty} I_k^i$$

And,

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \ell(I_k^i) < \sum_{i=1}^{\infty}(\mu^*(E_i) + \frac{\epsilon}{2^i}) = \sum_{i=1}^{\infty} \mu^*(E_i) + \epsilon$$

Consequently,

$$\mu^*(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} \mu^*(E_i) + \epsilon$$

The result follows because $\varepsilon > 0$ was random or arbitrary.

Example 4.1: Calculate the outer measure of the set of irrational numbers in the interval [0,1].

Solution: Assume that A be the set of irrational numbers in [0, 1].

Because $A \subseteq [0, 1]$, then $\mu*(A) \leq 1$.

Let Q be the set of rational numbers in [0, 1]. Note that $[0, 1] = A \cup Q$.

By Theorem 4.2 Property (5) and solving, we obtain

$1 \leq \mu*(A) + \mu*(Q)$

However, if Q is countable, then by Theorem 4.2 Property (2), $\mu*(Q) = 0$.

Therefore, $\mu*(A) = 1$.

4.2.1 Measurable Sets

In mathematics, the term measure can be defined as a simplification and generalization of typical conventional notions, such as mass, distance/length, area, volume, probability of events, etc.

The modern measure theory was characteristically given by Émile Borel, Henri Lebesgue, Nikolai Luzin, Johann Radon, Constantin Carathéodory, and Maurice Fréchet.

In mathematical analysis, the 'Real Analysis' only considers and evaluates the functions of a real variable and numbers, while the 'Measure Theory' exclusively considers and evaluates the concept of a measure, which is the method for 'Measuring' that how big a given set is. Sets to which the Lebesgue measure can be assigned are termed as the 'Lebesgue Measurable'; characteristically the measure of the Lebesgue measurable set A can be denoted by $\lambda(A)$. A measurable set is specifically defined as a set to which the extension or expansion can possibly be accomplished, this extension or expansion is assumed to be the measure. Consequently, the Jordan measure, the Borel measure and the Lebesgue measure can be defined using the sets which are characteristically measurable according to the Jordan, Borel and Lebesgue, respectively.

Definition 1: Let X be a set and Σ a σ-algebra over X. A function μ from Σ to the extended real number line is called a measure if it satisfies the following properties:

Non-Negativity: For all E in Σ, we have $\mu(E) \geq 0$.

Null Empty Set: $\mu(\emptyset) = 0$.

Countable Additivity or σ-Additivity: For all countable collections $\{E_k\}_{k=1}^{\infty}$ of pairwise disjoint sets in Σ,

$$\mu\left(\bigsqcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k)$$

If at least one set E has finite measure, then the constraint that $\mu(\emptyset) = 0$ is realized spontaneously. Certainly, through countable additivity,

$$\mu(E) = \mu(E \cup \emptyset) = \mu(E) + \mu(\emptyset)$$

And consequently, $\mu(\emptyset) = 0$.

Considering that the condition of non-negativity is ignored but the second and third of these conditions are fulfilled and μ takes the values $\pm\infty$, then μ is termed as a **signed measure**.

The pair (X, Σ) is termed as a **measurable space**, the members of Σ are termed as the **measurable sets**.

A measurable set X is known as a null set if $\mu(X) = 0$. A subset of a null set is described as a negligible set. A negligible set must not be measurable, but every measurable negligible set is certainly and inevitably a null set. A measure is termed complete if every negligible set is measurable.

A measure can be extended to a complete or perfect by means of considering the σ-algebra of subsets Y which vary through a negligible set from a measurable set X, i.e., the symmetric difference of X and Y is contained in a null set, such that $\mu(Y)$ can be defined to equal $\mu(X)$.

If the μ-measurable function f takes values on $[0, \infty]$ then,

$$\mu\{x : f(x) \geq t\} = \mu\{x : f(x) > t\}$$

NOTES

For almost all $t \in \mathbb{R}$ with respect to the Lebesgue measure. This property is typically used with Lebesgue integral.

Both $\mu\ \{x : f(x) > t\}$ and $\mu\ \{x : f(x) \geq t\}$ are defined as the monotonically non-increasing functions of t therefore both of them are continuous almost everywhere, with relation to the Lebesgue measure.

If $\mu\ \{x : f(x) > t\} = \infty$ for all t, then by the additivity and non-negativity,

$$\mu\ \{x : f(x) \geq t\} = \mu\ \{x : f(x) > t\} + \mu\ \{x : f(x) = t\} = \infty$$

being as essential.

Definition 2: A set $E \subseteq \mathbb{R}$ is Lebesgue measurable if for each set $A \subseteq \mathbb{R}$, the equality $\mu^*\ (A) = \mu^*\ (A \cap E) + \mu^*\ (A \cap \overline{E})$ is satisfied. If E is a Lebesgue measurable set, then the Lebesgue measure of E is its Lebesgue outer measure and will be written as $\mu(E)$.

Because the Lebesgue outer measure satisfies the subadditivity property, therefore we continually have $\mu^*\ (A) \leq \mu^*\ (A \cap E) + \mu^*\ (A \cap \overline{E})$ and can confirm the reverse inequality.

Note: There is a set E that divides A into two mutually exclusive sets, $A \cap E$ and $A \cap \overline{E}$, if and only if $u^*_{x_2}\ (A) = u^*_{x_2}\ (A \cap E) + \mu^*\ (A \cap \overline{E})$ holds, then the set E is termed as the Lebesgue measurable.

Example 4.2: Assume that E has measure zero where $E \subseteq \mathbb{R}$. Prove that the set $E^2 = \{x^2 : x \in E\}$ has measure zero.

Solution: Let $E_n = E \cap (-n, n) \subseteq E$.

Then, $E_n^{\ 2} = E^2 \cap (0, n^2) \subseteq E^2$

And, $E^2 = \bigcup\limits_{n=1}^{\infty} E_n^2.$

Because E has measure zero, then E_n has measure zero. Let $\varepsilon > 0$. Assume there exists a sequence of intervals (a_k, b_k) such that,

$$E_n \subseteq \bigcup_{k=1}^{\infty} (a_k, b_k) \quad \text{and} \quad \sum_{k=1}^{\infty} |b_k - a_k| < \frac{\epsilon}{2n}.$$

For simplicity, consider only the situation that $0 < a_k < b_k$.

Since,

$$\mu(a_k^2, b_k^2) = |b_k^2 - a_k^2| = |b_k + a_k| \cdot |b_k - a_k| \leq 2n|b_k - a_k|$$

Then,

$$\mu(E_n^2) \leq \sum_{k=1}^{\infty} \mu(a_k^2, b_k^2) \leq \sum_{k=1}^{\infty} 2n|b_k - a_k| = \epsilon.$$

It implies that the measure of $E_n^{\ 2}$ is zero, which indicates that the measure of E^2 is also zero.

Hence proved.

4.2.2 Regularity and Measurable Functions

Specifically, in the measure theory, a **measurable function** is a function between the underlying sets of two measurable spaces that preserves the structure of the spaces, the preimage of any measurable set is measurable. In real analysis, measurable functions are used in the definition of the Lebesgue integral. In probability theory, a measurable function on a probability space is termed as a random variable.

In real analysis and measure theory, the **regularity theorem** for **Lebesgue measure** is defined as an acquired result which states that Lebesgue measure on the real line is a **regular measure**. Usually, this indicates about the real line and states that every Lebesgue measurable subset is 'Approximately Open' and 'Approximately Closed'.

Statement of the Theorem

Lebesgue measure on the real line **R** is referred as a **regular measure**, i.e., for all Lebesgue measurable subsets A of **R**, and $\varepsilon > 0$, there exist subsets C and U of **R** such that,

- C is Closed
- U is Open
- $C \subseteq A \subseteq U$
- The Lebesgue Measure of $U \setminus C$ is strictly Less Than ε.

Additionally, when A has finite Lebesgue measure, then C is considered to be compact, i.e., by the Heine–Borel theorem it is closed and bounded.

Corollary: The Structure of Lebesgue Measurable Sets

If A is a 'Lebesgue Measurable Subset of **R**', then there exists a 'Borel Set B' and a 'Null Set N' such that A is the 'Symmetric Difference of B and N' and is given as,

$$\boldsymbol{A = B \,\Delta\, N = (B \setminus N) \cup (N \setminus B)}$$

Definition: Assume that (X, Σ) and (Y, T) be measurable spaces, signifying or implying that X and Y are the sets with respective σ-algebras Σ and T. A function $f: X \to Y$ is said to be measurable if for every $E \in T$ the pre-image of E under f is in Σ; that is, for all $E \in T$.

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

That is, $\sigma(f) \subseteq \Sigma$, where $\sigma(f)$ is the σ-algebra typically generated by f.

If $f: X \to Y$ is considered as a measurable function, then we have the equation of the form,

$$f: (X, \Sigma) \to (Y, T)$$

For emphasizing the dependency on the σ-algebras Σ and T.

Distinguished Classes of the Measurable Functions

Following are the three significant and distinguished classes of the measurable functions in real analysis:

Class 1: By definition the 'Random Variables' are the measurable functions specifically defined on probability spaces.

NOTES

Class 2: When (X, Σ) and (Y, T) are the Borel spaces, then a measurable function $f: (X, \Sigma) \to (Y, T)$ is also termed as a Borel function. Continuous functions are Borel functions but not all Borel functions are continuous. However, a measurable function is almost a continuous function as per the Luzin's theorem. If a Borel function occurs as a section of a map $Y \to X$ then it is known as a Borel section.

Class 3: A Lebesgue measurable function is a measurable function $f: (\mathbb{R}, \mathcal{L}) \to (\mathbb{C}, \mathcal{B}_{\mathbb{C}}) \to$ where '$\mathcal{L}$' is defined as the σ-algebra of Lebesgue measurable sets and $\mathcal{B}_{\mathbb{C}}$ is the Borel algebra on the complex numbers $\mathbb{C}$. Typically, the Lebesgue measurable functions are considered useful in mathematical analysis since they can be integrated. For the condition, $f: X \to \mathbb{R}$, f is considered as the Lebesgue measurable iff and only iff $\{f > \alpha\} = \{x \in X : f(x) > \alpha\}$ is uniquely measurable for all $\alpha \in \mathbb{R}$.

Additionally, this condition is also considered as equivalent for any of the specified $\{f \geq \alpha\}$, $\{f < \alpha\}$, $\{f \leq \alpha\}$ being measurable either for all 'α' or the preimage of any open set being measurable.

Continuous functions, monotone functions, step functions, semicontinuous functions, Riemann integrable functions and functions of bounded variation all are considered as the Lebesgue measurable. A function $f: X \to \mathbb{C}$ is measurable iff and only iff the real and imaginary parts are measurable.

4.2.3 Borel and Lebesgue Measurability

In mathematical analysis and in particular in the measure theory, a **Borel measure** on a topological space is defined as a measure for all open sets and consequently on all Borel sets.

Definition: Consider that X be a locally compact Hausdorff space and also consider that $\mathcal{B}(X)$ be the smallest σ-algebra which contains or includes the open sets of X; then this is established as the '**σ-Algebra of Borel Sets**'. Further, the '**Borel Measure'** is specified as any measure μ defined precisely on the σ-algebra of Borel sets. Some of the mathematicians define that μ is locally finite which implies that $\mu(C) < \infty$ for every compact set C. When a **Borel measure** μ is both **inner regular** and **outer regular**, then it is termed as a '**Regular Borel Measure**'. If μ is both inner regular and outer regular, and is also locally finite, then in this condition it is known as a Radon measure.

On the Real Line

Characteristically, the real line $\mathbb{R}$ with its normal topology is defined as a locally compact Hausdorff space, therefore a Borel measure can be defined on it. In this instance, $\mathcal{B}(\mathbb{R})$ is referred as the smallest σ-algebra that comprises of the open intervals of $\mathbb{R}$. Though there can be several Borel measures μ, we define the preferred option of Borel measure which assigns $\mu((a, b]) = b - a$ for every half-open interval $(a, b]$ and is therefore occasionally termed as the Borel measure on $\mathbb{R}$. This specific measure is considered as the restriction to the Borel σ-algebra of the Lebesgue measure λ, which is characterized and explained as a complete measure and is defined on the Lebesgue σ-algebra.

Principally, the Lebesgue σ-algebra is essentially stated as the completion of the Borel σ-algebra, which implies that it is the smallest σ-algebra that comprises of all the Borel sets and has a complete measure on it. Additionally, the Borel measure and the Lebesgue measure overlap or coincide on the Borel sets, i.e., $\lambda(E) = \mu(E)$ for every Borel measurable set, where μ is the Borel measure as already discussed.

Borel Function

Definition: The map $f: X \to Y$ between two topological spaces is termed as the '**Borel or Borel Measurable**' if $f^{-1}(A)$ is a Borel set for any open set A as per the σ-algebra of Borel sets of X is the smallest σ-algebra containing the open sets. When the target Y is taken as the real line, then it is sufficiently assumed that $f^{-1}(]a, \infty[)$ is Borel for any $a \in \mathbb{R}$. Considering the two topological spaces X and Y and also the corresponding Borel σ-algebras $\mathcal{B}(X)$ and $\mathcal{B}(Y)$ we can define that the Borel measurability of the function $f: X \to Y$ is then equivalent to the measurability of the map f realized as map between the measurable spaces $(X, \mathcal{B}(X))$ and $(Y, \mathcal{B}(Y))$.

Lebesgue Measure

Consider that $\Omega = (0, 1]$.

Assume that $\mathcal{F}_0$ contains the empty set and all sets which have finite unions of the intervals of the form $(a, b]$. A conventional and characteristic element of this set is given as the form,

$$F = (a_1, b_1] \cup (a_2, b_2] \cup \ldots\ldots \cup (a_n, b_n]$$

Where, $0 \le a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ and $n \in \mathbb{N}$.

Lemma: The following three lemmas can be considered.

(1) $\mathcal{F}_0$ is an algebra.

(2) $\mathcal{F}_0$ is not a σ-algebra.

(3) $\sigma(\mathcal{F}_0) = \mathcal{B}$.

Proof: Following proofs are derived for the above mentioned three lemmas.

(1) By definition it can be stated that $\Phi \in \mathcal{F}_0$. Also, $\Phi^C = (0, 1] \in \mathcal{F}_0$. Consider the complement of $(a_1, b_1] \cup (a_2, b_2]$ is $(0, a_1] \cup (b_1, a_2] \cup (b_2, 1]$, which also belongs to $\mathcal{F}_0$. Additionally, it can be defined that the union of several finite sets each of which are also the finite unions of the intervals of the form $(a, b]$, is too also a set which can be taken as the union of finite number of intervals and therefore belongs to $\mathcal{F}_0$.

(2) Remember that $\left(0, \frac{n}{n+1}\right] \in \mathcal{F}_0$ for every n, but there is also $\bigcup_{n=1}^{\infty} \left(0, \frac{n}{n+1}\right] = (0, 1) \notin \mathcal{F}_0$.

(3) Initially, the null set is evidently referred to as a Borel set. Following, we have previously observed that every interval of the form $(a, b]$ is termed as a Borel set. Consequently, every element of $\mathcal{F}_0$ (except the null set) is considered as a finite union of such intervals and is also considered as a Borel set. Therefore, $\mathcal{F}_0 \subseteq \mathcal{B}$. This specifies or implies that $\sigma(\mathcal{F}_0) \subseteq \mathcal{B}$.

For every $F \in \mathcal{F}_0$ of the form,

$F = (a_1, b_1] \cup (a_2, b_2] \cup \ldots\ldots \cup (a_n, b_n]$

The function $\mathbb{P}_0 : \mathcal{F}_0 \to [0, 1]$ can be specifically defined such that,

$\mathbb{P}_0 (\Phi) = 0$

And, $\mathbb{P}_0 (F) = \sum_{i=1}^{n} (b_i - a_i).$

Remember that $\mathbb{P}_0 (\Omega) = \mathbb{P}_0 ((0, 1]) = 1.$

Additionally, if $(a_1, b_1], (a_2, b_2], \ldots\ldots (a_n, b_n]$ are disjoint sets, then by implying the finite additivity of $\mathbb{P}_0$ we have the equations of the form,

$$\mathbb{P}_0 \left(\bigcup_{i=1}^{n} ((a_i, b_i]) \right) = \sum_{i=1}^{n} \mathbb{P}_0 ((a_i, b_i])$$
$$= \sum_{i=1}^{n} (b_i - a_i)$$

This specifies and implies that $\mathbb{P}_0$ is also countably additive on $\mathcal{F}_0$ too.

Subsequently, it can be stated that there exists a unique probability measure $\mathbb{P}$ on $((0, 1] , \mathcal{B})$ which is also equivalent as $\mathbb{P}_0$ on $\mathcal{F}_0$. This unique and distinctive probability measure on (0, 1] is termed as the **Lebesgue measure** or **uniform measure**.

4.2.4 Non-Measurable Sets

In mathematical analysis, a **non-measurable set** is considered as a set, which is not assigned any significant 'Volume'. The mathematical existence of such unique sets is interpreted for providing information and evidences about the notions and basic concepts of length, area and volume in the conventional set theory. According to the Zermelo–Fraenkel set theory, the axiom of choice entails that non-measurable subsets of $\mathbb{R}$ exist.

In mathematical analysis, a non-measurable set is defined as a set to which a significant meaningful 'Volume' cannot be assigned. Mathematically, in formal set theory the existence of these sets is interpreted for providing evidence about the notions of length, area and volume.

The notion and concept of a non-measurable set has been historically led by the Félix Édouard Justin Émile Borel and the Andrey Nikolaevich Kolmogorov for formulating probability theory on sets which are significantly constrained or restrictrd to be measurable. Characteristically, the measurable sets on the line are considered as the iterated countable unions and intersections of intervals, termed as the Borel sets, are referred as the plus-minus null sets. These sets are sufficiently adequate to involve every conceivable or feasible definition of a set that are used in standard mathematical analysis and solutions, but it needs exceptionally unique formulations to prove that the sets are measurable.

The measure of the union of two disjoint sets to be the sum of the measure of the two sets. A **measure** with this natural property is termed as the **finitely additive**. Though a finitely additive measure is adequate and appropriate for most perception of area and is also analogous or equivalent to Riemann integration but it is not appropriate for solving probability problems, because the predictable and conventional contemporary behaviours of sequences of events or random variables claim for countable additivity.

Theorem 4.3: Any measurable subset $A \subset \mathbb{R}$ with $\lambda(A) > 0$ contains a non-measurable subset.

Proof: The simple method to prove the given theorem is used as a standard result in measure theory, considering that λ denotes the Lebesgue measure on $\mathbb{R}$.

It is appropriate to assume that $A \subset (0, 1)$.

Consider that if $A \subset \mathbb{R}$ takes the positive measure, then there is some $n \in \mathbb{Z}$ such that $A \cap (n, n+1)$ also holds positive measure and consequently by means of translation invariance it accordingly solves for,

$$A' = A \cap (n, n+1) - n \subset (0, 1)$$

Therefore, if $N \subset A'$ is considered as a non-measurable set, then $N + n \subset A \cap (n, n+1) \subset A$ is the required non-measurable set.

Check Your Progress

1. Define Lebesgue measure.
2. What is Lebesgue outer measure?
3. State about the measurable space.
4. When be a measure can extended as complete?
5. What is a measurable function?
6. Define the term Borel measurable.

4.3 INTEGRATION OF NON-NEGATIVE FUNCTIONS

The integral of a **non-negative** common measurable function can be defined as an appropriate supremum of approximations by means of simple functions and the integral of a measurable function (not necessarily positive) is defined as the difference of two integrals of non-negative measurable functions.

The integral of a non-negative function of a single variable is specifically defined as the area between the graph of that function and the X-axis. Fundamentally, the Lebesgue integral can be used for extending the integral to a bigger class or group of functions. Additionally, it can also extend or broaden the domains for defining and approximating these functions. Many years before the 20th century, the mathematicians were already aware of the theory that when the non-negative functions have a smooth adequate graph, basically the continuous functions on the closed bounded intervals, then the specific area under the curve can be defined as the integral and can be then uniquely computed with the help of the approximation techniques specifically on the region through polygons.

Indicator Functions

NOTES

For assigning a value to the integral of the indicator function 1_S of a measurable set S which is consistent with the given measure μ, the only satisfactory option is to set:

$$\int 1_S \, d\mu = \mu(S)$$

Note that the result may possibly be equal to $+\infty$, unless μ is a finite measure.

Let f be a non-negative measurable function on E, which helps in attaining the value $+\infty$, alternatively f takes non-negative values in the extended real number line. Subsequently, we obtain the equation of the form,

$$\int_E f\, d\mu = \sup\left\{ \int_E s\, d\mu : 0 \le s \le f,\ s \text{ simple} \right\}$$

This integral coincides or overlaps with the previous one normally defined on the set of simple functions, when E is considered as a segment $[a, b]$.

We have defined the integral of f for any non-negative extended real-valued measurable function on E.

For some specific functions, the integral $\int_E f\, d\mu$ is considered as infinite.

For a non-negative measurable function f, assume that $s_n(x)$ be the simple function whose value is $k/2^n$ whenever $k/2^n \le f(x) < (k+1)/2^n$, for k being a non-negative integer less than 4^n. This can be directly proved that,

$$\int f\, d\mu = \lim_{n\to\infty} \int s_n\, d\mu$$

If the limit on the right hand side exists as an extended real number.

In this section we will use the measure space (X, F, μ).

Definition: Let s be a non-negative F measurable function such that, $s = \sum_{i=1}^{N} a_i X_{A_i}$ with disjoint F measurable sets A_i, $\bigcup_{i=1}^{N} A_i = X$ and $a_i \ge 0$. For any $E \in F$ define the integral of f over E to be, $I_E(s) = \sum_{i=1}^{N} a_i \mu(A_i \cap E)$ with the convention that if $a_i = 0$ and $\mu(A_i \cap E) = +\infty$ then $0 \times (+\infty) = 0$. Therefore, the area under $s \equiv 0$ in R is zero.

Example 4.3: Consider that $([0, 1], \mathcal{L}, \mu)$. Define,

$$f(x) = \begin{cases} 1 & \text{if } x \text{ Rational} \\ 0 & \text{if } x \text{ Irrational} \end{cases}$$

This is a simple function with $A_1 = Q \cap [0,1] \in L$ and A_0 the set of irrationals in $[0,1]$ which is in L as the complement of A_1. Thus, f is measurable and is given as,

$$\begin{aligned} I_{[0,1]}(f) &= 1\mu(\mathbb{Q} \cap [0,1]) + 0\mu(\mathbb{Q}^C \cap [0,1]) \\ &= 0 \end{aligned}$$

Since, the Lebesgue measure of a countable set is zero.

Lemma 1: If $E_1 \subseteq E_2 \subseteq E_3 \ldots$ are in F and $E = \bigcup_{n=1}^{\infty} E_n$ then,

$$\lim_{n\to\infty} \mu(E_n) = \mu(E).$$

and we say that we have an increasing sequence of sets.

Proof: If there exists an n such that $\mu(E_n) = +\infty$ then $E_n \subseteq E$ implies $\mu(E) = +\infty$ and the result follows.

So assume that $\mu(E_n) < +\infty$ for all $n \geq 1$. Then, $E = E_1 \cup \bigcup_{n=2}^{\infty} (E_n \setminus E_{n-1})$ is a disjoint union. Note that $E_{n-1} \subseteq E_n$ implies that $E_n = (E_n \setminus E_{n-1}) \cup E_{n-1}$, which is a disjoint union. So $\mu(E_n) = \mu(E_n \setminus E_{n-1}) + \mu(E_{n-1})$. Because the measures are finite, this can be rearranged as follows:

$$\mu(E_n \setminus E_{n-1}) = \mu(E_n) - \mu(E_{n-1}). \text{ So,}$$

$$\begin{aligned} \mu(E) &= \mu(E_1) + \sum_{n=2}^{\infty} \mu(E_n \setminus E_{n-1}) \\ &= \mu(E_1) + \lim_{N\to\infty} \sum_{n=1}^{N} (\mu(E_n) - \mu(E_{n-1})) \end{aligned}$$

(By the Definition of Infinite Sum)

$$= \lim_{N\to\infty} \mu(E_N)$$

Theorem 4.4: Let s and t be two simple non-negative F measurable functions on (X, F, μ) and $E, F \in F$. Then,

1. $I_E(cs) = cI_E(s)$ for all $c \in R$.
2. $I_E(s + t) = I_E(s) + I_E(t)$.
3. If $s \leq t$ on E then $I_E(s) \leq I_E(t)$.
4. If $F \subseteq E$ then $I_F(s) \leq I_E(s)$.
5. If $E_1 \subseteq E_2 \subseteq E_3 \subseteq \ldots$ and $E = \bigcup_{k=1}^{\infty} E_k$ then $\lim_{k\to\infty} I_{E_k}(s) = I_E(s)$.

Proof: As per the Lemma given above we can state that,

$$s = \sum_{i=1}^{M} a_i X_{A_i} = \sum_{i=1}^{M} \sum_{j=1}^{N} a_i XC_{i\,j}$$

And

$$t = \sum_{j=1}^{N} b_j XB_j = \sum_{i=1}^{M} \sum_{j=1}^{N} b_j XC_{ij}$$

With $C_{i,j} = A_i \cap B_j \in F$.

NOTES

1. Note that $cs = \sum_{i=1}^{M} ca_i XA_i$ and so,

$$I_E(cs) = \sum_{i=1}^{M} ca_i \mu(A_i)$$

$$= c\sum_{i=1}^{M} a_i \mu(A_i) = cI_E(s)$$

2. Then $s + t = \sum_{i=1}^{M}\sum_{j=1}^{N}(a_i + b_j)XC_{ij}$. Therefore,

$$I_E(s+t) = \sum_{i=1}^{M}\sum_{j=1}^{N}(a_i + b_j)\mu(C_{ij} \cap E)$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} a_i \mu(C_{ij} \cap E) + \sum_{i=1}^{M}\sum_{j=1}^{N} b_i \mu(C_{ij} \cap E)$$

$$= \sum_{i=1}^{M} a_i \mu\left(\bigcup_{j=1}^{N}(C_{ij} \cap E)\right) + \sum_{j=1}^{N} b_i \mu\left(\bigcup_{i=1}^{M}(C_{ij} \cap E)\right)$$

$$= \sum_{i=1}^{M} a_i \mu(A_i \cap E) + \sum_{j=1}^{N} b_j \mu(B_j \cap E)$$

$$= I_E(s) + I_E(t)$$

3. Given any $1 \le i \le M$, $1 \le j \le N$ for which $C_{ij} \cap E \ne \phi$ we obtain for any $x \in C_{ij} \cap E$ such that,

$$a_i = s(x) \le t(x) = b_j. \text{ So,}$$

$$I_E(s) = \sum_{i=1}^{M}\sum_{j=1}^{N} a_i \mu(C_{ij} \cap E)$$

$$\le \sum_{i=1}^{M}\sum_{j=1}^{N} b_j \mu(C_{ij} \cap E)$$

$$= I_E(t)$$

4. By monotonicity of μ we have,

$$I_F(s) = \sum_{i=1}^{M} a_i \mu(A_i \cap F)$$

$$\le \sum_{i=1}^{M} a_i \mu(A_i \cap E)$$

$$= I_E(s)$$

5. If $E_1 \subseteq E_2 \subseteq E_3 \subseteq ....$ and $E = \bigcup_{k=1}^{\infty} E_k$ then we have $\lim_{n\to\infty} \mu(E_k) = \mu(E)$. Thus,

$$\lim_{k\to\infty} I_{E_k}(s) = \lim_{k\to\infty}\sum_{i=1}^{M} a_i \mu(A_i \cap E_k)$$

$$= \sum_{i=1}^{M} a_i \lim_{k\to\infty} \mu(A_i \cap E_k)$$

NOTES

$$= \sum_{i=1}^{M} a_i \mu(A_i \cap E)$$

$$= I_E(s)$$

Definition: If $f: X \to R^+$ is a non-negative F measurable function and $E \in F$, then the integral of f over E is given as,

$$\int_E f d\mu = \sup\{I_E(s) : s \text{ is a simple } F\text{-measurable function, } 0 \le s \le f\}$$

But, if $E \ne X$ then only f is defined on some domain containing E.

Let $I(f, E)$ denote the set,

$\{I_E(s) : s$ is a simple F-measurable function, $0 \le s \le f\}$

So the integral equals $\sup I(f, E)$.

Note: The integral exists for all non-negative F measurable functions, even though it may be infinite.

If $\int_E f d\mu = \infty$ then the integral is defined.

If $\int_E f d\mu < \infty$ then f is μ-integrable or summable on E.

Theorem 4.5: For a non-negative F measurable simple function t, we have $\int_E t d\mu = I_E(t)$.

Proof: Given any simple F measurable function, $0 \le s \le t$ we have $I_E(s) \le I_E(t)$ by Theorem 4.4.

Let $I_E(t)$ is an upper bound for $I(t, E)$ for which $\int_E t d\mu$ is the least of all upper bounds.

Hence,

$$\int_E t d\mu \le I_E(t)$$

Also, $\int_E t d\mu \ge I_E(s)$ for all simple F measurable functions, $0 \le s \le t$ and therefore is greater than $I_E(s)$ for any particular s, namely $s = t$. Hence, $\int_E t d\mu \ge I_E(t)$.

Example 4.4: If $f \equiv k$, i.e., a constant, then $\int_E f d\mu = I_E(f) = k_\mu(E)$.

Theorem 4.6: Consider that all sets are in F and all functions are non-negative and F measurable.

1. For all $c \ge 0$,

$$\int_E cf d\mu = c \int_E f d\mu \qquad \ldots(4.3)$$

NOTES

2. If $0 \le g \le h$ on E then,

$$\int_E g\,d\mu \le \int_E h\,d\mu$$

3. If $E_1 \subseteq E_2$ and $f \ge 0$ then,

$$\int_{E_1} f\,d\mu \le \int_{E_2} f\,d\mu$$

Proof:

1. If $c = 0$ then both the right hand side and left hand side of Equation (4.3) are 0. Assume $c > 0$.

   If $0 \le s \le cf$ is a simple F measurable function then also is $0 \le \frac{1}{c}s \le f$.

Thus,

$$\int_E f\,d\mu \ge I_E\left(\frac{1}{c}s\right) = \frac{1}{c}I_E(s)$$

By Theorem 4.4 (1).

Hence, $c\int_E f\,d\mu$ is an upper bound for $I(cf, E)$ for which $\int_E cf\,d\mu$ is the least upper bound. Thus, $c\int_E f\,d\mu \ge \int_E cf\,d\mu$.

Considering the observation that if $0 \le s \le f$ is a simple F measurable function then also is $0 \le cs \le cf$ and we obtain,

$$\int_E cf\,d\mu \ge I_E(cs) \qquad \text{By the definition of } \int_E$$

$$= cI_E(s) \qquad \text{By Theorem 4.4(1).}$$

Hence, $\frac{1}{c}\int_E (cf)\,d\mu$ is an upper bound for $I(f, E)$ for which $\int_E f\,d\mu$ is the least upper bound. Therefore, $\frac{1}{c}\int_E (cf)\,d\mu \ge \int_E f\,d\mu$, or, $\int_E (cf)\,d\mu \ge c\int_E f\,d\mu$.

On combining both inequalities, we obtain the result.

2. Let $0 \le s \le g$ be a simple F measurable function. Then, since $g \le h$ we trivially have $0 \le s \le h$ in which the $I_E(s) \le \int_E h\,d\mu$ by the definition of integral $\int_E$.

   Thus, $\int_E h\,d\mu$ is an upper bound for $I(g, E)$. As in (1) we get

   $$\int_E h\,d\mu \ge \int_E g\,d\mu.$$

3. Let $0 \le s \le f$ be a simple, F measurable function. Then,

   $I_{E1}(s) \le I_{E2}(s)$ By Theorem 4.4(3)

NOTES

$\leq \int_{E_2} f d\mu$ By the definition of $\int_{E_2}$

So $\int_{E_2} f d\mu$ is an upper bound for $I(f, E_1)$ and hence is greater than the least of all upper bounds. consequently, $\int_{E_2} f d\mu \geq \int_{E1} f d\mu$.

Lemma 1: Let $E \in F, f \geq 0$ is F measurable and $\int_E f d\mu < \infty$. Set, $A = \{x \in E : f(x) = +\infty\}$. Then, $A \in F$ and $\mu(A) = 0$.

Proof: Since f is F measurable, therefore $f^{-1}(\{\infty\}) \in F$ and also $A = E \cap f^{-1}(\{\infty\}) \in F$. Define,

$$s_n(x) = \begin{cases} n & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A \end{cases}$$

Since $A \in F$, we infer that s_n is an F measurable simple function. Also, $s_n \leq f$ and so,

$$\begin{aligned} n\mu(A) &= I_E(s_n) && \text{by definition of } I_E \\ &\leq \int_E \ f d\mu && \text{by definition of } \int_E \\ &< \infty && \text{by assumption} \end{aligned}$$

Which is true for all $n \geq 1$ specifying that $\mu(A) = 0$.

Lemma 2: If f is F measurable and non-negative on $E \in F$ and $\mu(E) = 0$, then $\int_E f d\mu = 0$.

Proof: Let $0 \leq s \leq f$ be a simple F measurable function. Therefore, $s = \sum_{n=1}^{N} a_n X_{A_n}$ for some $a_n \geq 0, A_n \in F$. Then $I_E(s) = \sum_{n=1}^{N} a_n \mu(A_n \cap E)$. But μ is monotone which specifies that $\mu(A_n \cap E) \leq \mu(E) = 0$ for all n and hence $I_E(s) = 0$ for all such simple functions. Consequently, $I(f, E) = \{0\}$ and so $\int_E f d\mu = \sup I(f, E) = 0$.

Lemma 3: If $g \geq 0$ and $\int_E g d\mu = 0$, then $\mu\{x \in E : g(x) > 0\} = 0$.

Proof: Let $A = \{x \in E : g(x) > 0\}$ and $A_n = \{x \in E : g(x) > 1/n\}$.

Then, the sets $A_n = E \cap \{x : g(x) > 1/n\} \in Fb$ satisfy $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$with $A = \bigcup_{n=1}^{\infty} A_n$.

By Lemma 1, $\mu(A) = \lim_{n \to \infty} \mu(A_n)$.Using,

$$s_n(x) = \begin{cases} \dfrac{1}{n} & \text{if} \quad x \in A_n \\ 0 & \text{otherwise} \end{cases}.$$

Therefore, $s_n \leq g$ on A_n we have,

NOTES

$$\frac{1}{n}\mu(A_n) = I_{A_n}(s_n)$$

$$\leq \int_{A_n} gd\mu \quad \text{by the definition of } \int_{A_n}$$

$$\leq \int_{E_n} gd\mu \qquad \text{By Theorem 4.6(3)}$$

$$= 0 \qquad \text{By assumption}$$

So $\mu(A_n) = 0$ for all n and hence $\mu(A) = 0$.

Definition: If a property P holds on all points in $E \setminus A$ for some set A with $\mu(A) = 0$ then P is said to hold almost everywhere (μ) on E. It is possible that P holds on some of the points of A or that the set of points on which P does not hold is non-measurable. But, if μ is a complete measure, such as the Lebesgue-Stieltjes measure μ_F, then the condition is simple. Assume that a property P holds almost everywhere (μ) on E. The definition says that the set of points, say D, on which P does not hold, can be covered by a set of measure zero, i.e., there exists $A : D \subseteq A$ and $\mu(A) = 0$.

However, if μ is complete then D will be defined as measurable of measure zero.

Lemma 4: If $g \geq 0$ and $\int_E gd\mu = 0$ then $g = 0$ almost everywhere (μ) on E.

Theorem 4.7: If g, h: $X \to R^+$ are F measurable functions and $g \leq h$ almost everywhere (μ) then, $\int_E gd\mu \leq \int_E hd\mu$.

Proof: By assumption there exists a set $D \subseteq E$, of measure zero, such that for all $x \in E/D$ we have $g(x) \leq h(x)$. Let $0 \leq s \leq g$ be a simple F measurable function written as,

$$s = \sum_{i=1}^{N} a_i X_{A_i} \text{, with } \bigcup_{i=1}^{N} A_i = E$$

A simple F measurable function can be defined as,

$$s^*(x) = \begin{cases} s(x) & \text{if} \quad x \notin D \\ 0 & \text{if} \quad x \in D \end{cases}$$

$$= \sum_{i=1}^{N} a_i X_{A_i} \cap D^C$$

Then, for $x \in E/D$ we have $s^*(x) = s(x) \leq g(x) \leq h(x)$, while for $x \in D$ we have $s^*(x) = 0 \leq h(x)$. Thus, $s^*(x) \leq h(x)$ for all $x \in E$.

Remember that, $A_i = (A_i \cap D^c) \cup (A_i \cap D)$, a disjoint union in which $\mu(A_i) = \mu(A_i \cap D^c) \cup \mu(A_i \cap D) = \mu(A_i)$. But $A_i \cap D \subseteq D$ and so $\mu(A_i \cap D) \leq \mu(D) = 0$. Thus, $\mu(A_i) = \mu(A_i \cap D^c)$. Consequently,

$$I_E(s^*) = \sum_{i=1}^{N} a_i \mu(A_i \cap D^C)$$

$$= \sum_{i=1}^{N} a_i \mu(A_i)$$

$$= I_E(s)$$

Therefore, $I_E(s) = I_E(s^*) \leq \int_E hd\mu$ by the Definition of Integral $\int_E$.

Thus, $\int_E hd\mu$ is an upper bound for $I(g, E)$ while $\int_E gd\mu$ is the least of all upper bounds for $I(g, E)$. Hence, $\int_E hd\mu \geq \int_E gd\mu$.

Corollary: If $g, h: X \to R^+$ are F measurable with $g = h$ almost everywhere (μ) on E then,

$$\int_E gd\mu = \int_E hd\mu \, .$$

Proof: By assumption there exists a set $D \subseteq E$ of measure zero such that for all $x \in E/D$ we have $g(x) = h(x)$. In particular, for these x we have $g(x) \leq h(x)$ and $h(x) \leq g(x)$. Therefore, $g \leq h$ almost everywhere (μ) on E and $h \leq g$ almost everywhere (μ) on E. Hence, the result follows from two applications of Theorem 4.7.

Therefore, a function may have its values changed on a set of measure zero without changing the value of its integral. Particularly, we may assume that a non-negative integrable function has finite value.

4.4 THE GENERAL INTEGRAL

A general integral can be defined as a relation between the variables in the equation including one arbitrary function such that the equation is satisfied when the relation is substituted in it, for every alternative of the arbitrary function.

An integral can be defined as the distributional integral of functions of one real variable, i.e., more general as compared to the Lebesgue integral which permits the integration of functions with distributional values everywhere or nearly everywhere.

Define the positive part f^+ and negative part f^- of a function as,

$f^+ = \max(f, 0)$

$f^- = \max(-f, 0)$

Also,

$f = f^+ - f^-$

$|f| = f^+ + f^-$

Definition: A measurable function f is said to be integrable over E if f^+ and f^- are both integrable over E. Therefore, we can define,

$$\int_E f = \int_E f^+ - \int_E \overline{f}$$

NOTES

Theorem 4.8: Let f and g be integrable over E. Then,

(*i*) The function $f + g$ is integrable over E and $\int_E (f+g) = \int_E f + \int_E g$.

(*ii*) If $f \leq g$ almost everywhere then, $\int_E f \leq \int_E g$.

(*iii*) If A and B are disjoint measurable sets contained in E, then

$$\int_{A \cup B} f = \int_A f + \int_B f.$$

Proof: From the definition, it follows that the functions f^+, f^-, g^+, g^- are all integrable. If $h = f + g$, then $h = (f^+ - f^-) + (g^+ - g^-)$ and hence $h = (f^+ + g^+) - (f^- + g^-)$. Since, $f^+ + g^+$ and $f^- + g^-$ are integrable therefore we then have, the following equation:

$$\begin{aligned}
\int_E h &= \int_E [(f^+ + g^+) - (\overline{f} + \overline{g})] \\
&= \int_E (f^+ + g^+) - \int_E (\overline{f} + \overline{g})] \\
&= \int_E f^+ + \int_E g^+ - \int_E \overline{f} - \int_E \overline{g} \\
&= (\int_E f^+ - \int_E \overline{f}) + (\int_E g^+ - \int_E \overline{g})
\end{aligned}$$

That is,

$$\int_E (f+g) = \int_E f + \int_E g$$

Proof of (*ii*) follows from Part (*i*) and the fact that the integral of a non-negative integrable function is non-negative.

For the Proof of (*iii*) we have,

$$\begin{aligned}
\int_{A \cup B} f &= \int f \chi_{A \cup B} \\
&= \int f \chi_A + \int f \chi_B \\
&= \int_A f + \int_B f
\end{aligned}$$

Now, $f + g$ is not defined at points where $f = \infty$ and $g = -\infty$, and where $f = -\infty$ and $g = \infty$. However, the set of such points must have measure equal to 0, since f and g are integrable. Consequently, the integrability and the value of $\int (f+g)$ is independent of the choice of values in these ambiguous conditions.

Theorem 4.9: Let f be a measurable function over E. Then f in integrable over E iff $|f|$ is integrable over E. Furthermore, if f is integrable, then

$$|\int_E f| \leq \int_E |f|.$$

Proof: If f is integrable then both f^+ and f^- are also integrable. But $|f| = f^+ + f^-$. Hence, integrability of f^+ and f^- implies the integrability of $|f|$.

Moreover, if f is integrable, then since $f(x) \leq |f(x)| = f(x)$, the property of measurable function states that if $f \leq g$ almost everywhere then, $\int f \leq \int g$ implies that,

$$\int f \leq \int |f| \qquad \ldots(4.4)$$

On the other hand, since $-f(x) \leq |f(x)|$, we have

$$-\int f \leq \int |f| \qquad \ldots.(4.5)$$

From Equations (4.4) and (4.5) we have,

$$|\int f| \leq \int |f|$$

Conversely, suppose f is measurable and suppose $|f|$ is integrable. Since, $0 \leq f^+(x) \leq |f(x)|$, it follows that f^+ is integrable. Similarly, f^- is also integrable then f is also integrable.

Lemma: Let f be integrable. Then given $\varepsilon > 0$, there exists $\delta > 0$ such that $|\int_A f| < \varepsilon$ whenever A is a measurable subset of E with $mA < \delta$.

Proof: When f is non-negative, the lemma is proved. Now for arbitrary measurable function f we have $f = f^+ - f^-$. Therefore, given $\varepsilon > 0$ there exists $\delta_1 > 0$ such that,

$$\int_A f^+ < \frac{\varepsilon}{2}$$

When $mA < \delta_1$. Similarly there exists $\delta_2 > 0$ such that,

$$\int_A f^- < \frac{\varepsilon}{2}$$

When $mA < \delta_2$. Thus, when $mA < \delta = \min(\delta_1, \delta_2)$, we have the following equation:

$$|\int_A f \leq \int_A |f| = \int_A f^+ + \int_A \overline{f} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Hence, the lemma is proved.

4.5 INTEGRATION OF SERIES

A **series** formalizes the inaccurate notion of taking the sum of an endless sequence of numbers. The contemporary notion that to assign a value to a series can be avoided by considering the inaccurate notion of adding an 'Infinite' number of terms. As an alternative, the finite sum of the first n terms of the sequence, known as a partial sum, is considered and the concept of a limit is used to the sequence of partial sums as n increases or expands without bound. The series is assigned the value of this limit if it exists.

Consider an integer N and a function f defined on the unbounded interval $[N, \infty)$, on which it is monotonic decreasing. Then the infinite series,

$$\sum_{n=N}^{\infty} f(n)$$

Converges to a real number if and only if the improper integral

$$\int_N^{\infty} f(x)\,dx$$

is finite. In particular, if the integral diverges, then the series diverges as well.

Proof: The proof basically uses the comparison test, comparing the term $f(n)$ with the integral of f over the intervals $[n-1, n)$ and $[n, n+1)$, respectively. The monotonous function f is continuous almost everywhere. To evaluate this, let $D = \{x \in [N, \infty) \mid f$ is discontinuous at $x\}$.

For every $x \in D$ exists by the density of $\mathbb{Q}$ a $c(x) \in \mathbb{Q}$ so that $c(x) \in \left[\lim_{y \downarrow x} f(y), \lim_{y \uparrow x} f(y)\right]$.

Since, f is a monotonic decreasing function, we know that

$$f(x) \leq f(n) \quad \text{for all } x \in [n, \infty)$$

And

$$f(n) \leq f(x) \quad \text{for all } x \in [N, n].$$

Hence, for every integer $n \geq N$,

$$\int_n^{n+1} f(x)\,dx \leq \int_n^{n+1} f(n)\,dx = f(n) \tag{4.6}$$

And, for every integer $n \geq N+1$,

$$f(n) = \int_{n-1}^{n} f(n)\,dx \leq \int_{n-1}^{n} f(x)\,dx. \tag{4.7}$$

By summation over all n from N to some larger integer M, we get the following Equation from Equation (4.6)

$$\int_N^{M+1} f(x)\,dx = \sum_{n=N}^{M} \underbrace{\int_n^{n+1} f(x)\,dx}_{\leq f(n)} \leq \sum_{n=N}^{M} f(n)$$

And from Equation (4.7)

$$\sum_{n=N}^{M} f(n) \leq f(N) + \sum_{n=N+1}^{M} \underbrace{\int_{n-1}^{n} f(x)\,dx}_{\geq f(n)} = f(N) + \int_{N}^{M} f(x)\,dx.$$

Combining these two estimates we obtain the following yields:

$$\int_{N}^{M+1} f(x)\,dx \leq \sum_{n=N}^{M} f(n) \leq f(N) + \int_{N}^{M} f(x)\,dx.$$

Remark: If the improper integral is finite, then the proof also gives the lower and upper bounds.

$$\int_{N}^{\infty} f(x)\,dx \leq \sum_{n=N}^{\infty} f(n) \leq f(N) + \int_{N}^{\infty} f(x)\,dx$$

for the infinite series.

4.6 REIMANN AND LEBESGUE INTEGRALS

The **Lebesgue integral** is named after Henri Lebesgue (1875–1941), typically Lebesgue defined and established the integral in the year 1904. Principally, the Lebesgue integral functions have a significant role in the theory and derivation of probability, real analysis, and numerous other fields in mathematics. Mathematically, as per the Lebesgue explanation the term Lebesgue integration specifies either the general theory of integration of a function with respect to a general measure or the specific instance of integration of a function typically defined on a sub-domain of the real line with respect to the Lebesgue measure. The **Riemann integral** specifies that by partitioning the domain of an assigned function, one can approximate or estimate the assigned function by means of piecewise constant functions in each sub-interval. On the contrary, the Lebesgue integral are specifically used to partition the range of that function.

The key objective of the Lebesgue integral is to provide an integral notion in which the limits of integrals hold moderate assumptions. Basically, there is no assurance that every function is the Lebesgue integrable, but it is possible that improper integrals exist for functions that are not Lebesgue integrable.

Any function which is **Riemann integrable** is also Lebesgue integrable and positively with the same values for the two integrals. This can be easily proved. First, we recall one definition of Riemann integrability. This definition is different from other, but is considered to be equivalent. Let $f : A \to R$ be a bounded function on a bounded rectangle $A \subseteq R^m$. Consider R-valued functions that are simple with respect to a rectangular partition of A, also known as step functions. Step functions are obviously both Riemann and Lebesgue integrable with the same values for the integral. The lower and upper Riemann integrals for f are given as,

$$\mathscr{L}(f) = \sup\left\{\int_A l\; d\lambda : \text{Step Function } l \leq f\right\}$$

$$\mathscr{U}(f) = \inf\left\{\int_A u\; d\lambda : \text{Step Function } u \geq f\right\}$$

NOTES

If we have $\mathscr{L}(f) \leq \mathscr{U}(f)$; then we state that f is Riemann integrable if $\mathscr{L}(f) = \mathscr{U}(f)$, and the Riemann integral of f is defined as $\mathscr{L}(f) = \mathscr{U}(f)$

Equivalently, f is Riemann integrable when there exists sequence of lower simple functions $l_n \leq f$ and upper simple functions $u_n \geq f$ such that,

$$\lim_{n\to\infty} \int_A l_n = \mathscr{L}(f) = \mathscr{U}(f) = \lim_{n\to\infty} \int_A u_n .$$

Theorem 4.10: (Riemann Integrability Implies Lebesgue Integrability): Let $A \subset R^m$ be a bounded rectangle. If $f : A \to R$ is properly Riemann integrable, then it is also Lebesgue integrable with respect to Lebesgue measure with the same value for the integral.

Proof: Take a sequence l_n and u_n. Let $L = \sup_n ln$ and $U = \inf_n u_n$. Clearly, these are measurable functions, and we have,

$$l_n \leq L \leq f \leq U \leq u_n$$

Taking Lebesgue integral and defining, limits we can write,

$$\lim_{n\to\infty} \int l_n \leq \int L \leq \int U \leq \lim_{n\to\infty} \int u_n$$

Here, the limits on the two sides are the same, because the Riemann and Lebesgue integrals for l_n and u_n coincide. Therefore, $\int (U - L) = 0$. Then $U = L$ almost everywhere and U or L equals f almost everywhere. Since Lebesgue measure is complete hence f is a Lebesgue measurable function.

Finally, the Lebesgue integral $\int f$, now exists and is squeezed in between the two limits on the left and on the right defining that both equal the Riemann integral of f.

4.7 THE FOUR DERIVATIVES

In real analysis, the derivative of a function of a real variable measures the sensitivity for changing the function value (output value) with respect to a change in its argument (input value). Derivatives are a fundamental tool of calculus. For example, the derivative of the position of a moving object with respect to time is the object's velocity and this measures how quickly the position of the object changes when time advances.

The derivative of a function of a single variable at a selected input value, when it exists, is the slope of the tangent line to the graph of the function at that point. The tangent line is the best linear approximation of the function near that input value. For this reason, the derivative is often described as the "Instantaneous Rate of Change", the ratio of the instantaneous change in the dependent variable to that of the independent variable.

Derivatives can be generalized to functions of several real variables. In this generalization, the derivative is reinterpreted as a linear transformation whose graph is (after an appropriate translation) the best linear approximation to the graph of the original function. The Jacobian matrix is the matrix that represents this linear transformation with respect to the basis given by the choice of independent and dependent variables. It can be calculated in terms of the partial derivatives with respect to the independent variables. For a real valued function of several variables, the Jacobian matrix reduces to the gradient vector.

Differentiation and Integration

The process of finding a derivative is called **differentiation**. The reverse process is called antidifferentiation. The fundamental theorem of calculus relates antidifferentiation with **integration**. Differentiation and integration constitute the two fundamental operations in single variable calculus.

A function of a real variable $y = f(x)$ is differentiable at a point a of its domain, if its domain contains an open interval I containing a and the limit exists.

$$L = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

This defines that for every positive real number ε (even very small), there exists a positive real number δ such that, for every h there is $|h| < \delta$ and $h \neq 0$ then $f(a+h)$ is defined, and

$$\left| L - \frac{f(a+h) - f(a)}{h} \right| < \varepsilon,$$

If the function f is differentiable at a, that is, if the limit L exists, then this limit is called the derivative of f at a, and denoted by $f'(a)$ (read as "f prime of a") or $\frac{df}{dx}(a)$ (read as "The derivative of f with respect to x at a", "dy by dx at a", or "dy over dx at a").

Let f be a function that has a derivative at every point in its domain. We can then define a function that maps every point x to the value of the derivative of f at x. This function is written f' and is called the derivative function or the derivative of f. Let f be a differentiable function and let f' be its derivative. The derivative of f' (if it has one) is written f'' and is called the second derivative of f. Similarly, the derivative of the second derivative, if it exists, is written f''' and is called the third derivative of f. Continuing this process, one can define, if it exists, the nth derivative as the derivative of the $(n-1)$th derivative. These repeated derivatives are called higher order derivatives. The nth derivative is also called the derivative of order n.

A function f need not have a derivative for example if it is not continuous. Similarly, even if f does have a derivative, then it may not have a second derivative. For example, let

$$f(x) = \begin{cases} +x^2, & \text{if } x \geq 0 \\ -x^2, & \text{if } x \leq 0. \end{cases}$$

NOTES

Calculation shows that f is a differentiable function whose derivative at x is given by,

$$f'(x) = \begin{cases} +2x, & \text{if } x \geq 0 \\ -2x, & \text{if } x \leq 0. \end{cases}$$

$f'(x)$ is twice the absolute value function at x and it does not have a derivative at zero. A function can have a kth derivative for each non-negative integer k but not for $(k+1)$th derivative. A function that has k successive derivatives is called k times differentiable. If in addition the kth derivative is continuous, then the function is said to be of differentiability class C^k. This is a stronger condition than having k derivatives. A function that has infinitely many derivatives is called infinitely differentiable or smooth.

4.8 FUNCTIONS OF BOUNDED VARIATION

In the mathematical analysis, a function of Bounded Variation (BV) also termed as BV function, is considered as a real valued function whose total variation is bounded or finite. Considering a continuous function of a single variable, which has bounded variation signifies that the distance along the direction of the Y-axis ignoring the contribution of motion along X-axis, travelled by a point moving along the graph has a finite value. Similarly, consider a continuous function of several variables, the connotation and implication of the definition is equivalent, except that the considered continuous path cannot be the whole graph of the given function, but can be every intersection of the graph itself with a hyperplane (for several variables) and plane (for functions of two variables) parallel to a fixed X-axis and to the Y-axis.

Definition 1: Let S be a non-empty set of real numbers. Then,

1. The set S is bounded above if there is a number M such that $M \geq x$ for all $x \in S$. The number M is called an upper bound of S.
2. The set S is bounded below if there exists a number m such that $m \leq x$ for all $x \in S$.
3. The set S is bounded if it is bounded above and below. Equivalently S is bounded if there exists a number r such that $|x| \leq r$ for all $x \in S$. The number r is called a bound for S.

Definition 2: Let, S be a non-empty set of real numbers.

1. Suppose that S is bounded above. The number β is the supremum of S if β is an upper bound of S and there is no number less than β that is an upper bound of S. We write,

   $\beta = \sup S$
2. Suppose that S is bounded below. A number α is the infimum of S if α is a lower bound of S and there is no number greater than α that is a lower bound of S. We write,

   $\alpha = \inf S$

NOTES

Theorem 4.11: Let S be a non-empty set of real numbers that is bounded above, and let b be an upper bound of S. Then the following are equivalent:

1. $b = \sup S$.
2. For all $\varepsilon > 0$ there exists an $x \in S$ such that $|b - x| < \varepsilon$.
3. For all $\varepsilon > 0$ there exists an $x \in S$ such that $x \in (b - \varepsilon, b]$.

We often refer to sup S as the Least Upper Bound (LUB) of S and to inf S as the Greatest Lower Bound (GLB) of S.

Axiom: Every non-empty set of real numbers that is bounded above has a least upper bound.

Definition 3: A partition of an interval $[a, b]$ is a set of points $\{x_0, x_1, ..., x_n\}$ such that $a = x_0 < x_1 < x_2 ... < x_n = b$.

Definition 4: Let $f : \lceil a, b \rceil \to \mathbb{R}$ be a function and let $[c, d]$ be any closed subinterval of $[a, b]$. If the set,

$$s = \left\{ \sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) | : \{x_i : 1 \le i \le n\} \text{ is a partition of } [c, d] \right\}$$

is bounded then the variation of f on $[c, d]$ is defined to be $V(f, \lceil c, d \rceil) = \sup S$. If S is bounded then the variation of f is said to be ∞. A function f is of bounded variation on $[c, d]$ if $V(f, [c, d])$ is finite.

Example 4.5: If f is constant on $[a, b]$, then f is of bounded variation on $[a, b]$. Consider the constant function $f(x) = c$ on $[a, b]$. Let,

$$\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) |$$

is zero for every partition of $[a, b]$. Thus, $V(f, [a, b])$ is zero.

Theorem 4.12: If f is increasing on $[a, b]$, then f is of bounded variation on $[a, b]$ and $V(f, [a, b]) = f(b) - f(a)$.

Proof: Let $\{x_i : 1 \le i \le n\}$ be a partition of $[a, b]$. Consider,

$$\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) | = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) = f(b) - f(a)$$

This sum is the same for every partition of $[a, b]$. Therefore, we have $V(f, [a, b]) = f(b) - f(a) < \infty$. Thus, f is of bounded variation on $[a, b]$.

Similarly, if f is decreasing on $[a, b]$ then $V(f, [a, b]) = f(a) - f(b)$.

Theorem 4.13: Between any two distinct real numbers there is a rational number and an irrational number.

Example 4.6: Show that, the function f defined by,

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is Irrational} \\ 1 & \text{if } x \text{ is Rational} \end{cases}$$

is not of bounded variation on any interval.

Solution: Let $n \in \mathbb{Z}$ and $n > 0$. Let $[a, b]$ be a closed interval in $\mathbb{R}$. We construct a partition $P = \{x_0, x_1, \ldots, x_{n+2}\}$ of $[a, b]$ such that $V(f,[a, b]) \geq \sum_{i=1}^{n+2} |f(x_i) - f(x_{i-1})| > n$ as follows. Recall that by definition $x_0 = a$. We know that between any two real numbers there is a rational number and an irrational number. Take x_1 to be an irrational number between a and b. Then take x_2 to be an irrational number between x_1 and b. Continuing in this manner and taking x_{2i+1} to be an irrational number between x_{2i} and b, and x_{2i} to be a rational number between x_{2i-1} and b, then finally $x_{n+2} = b$. Thus, a partition is created that begins with a and then alternates between rational and irrational numbers until it finally ends with b. Now consider the sum $\sum_{i=1}^{n+2} |f(x_i) - f(x_{i-1})|$, which is the variation of f on $[a, b]$. Thus,

$$\begin{aligned} V(f,[a, b]) &\geq \sum_{i=1}^{n+2} |f(x_i) - f(x_{i-1})| \\ &\geq \sum_{i=2}^{n+2} |f(x_i) - f(x_{i-1})| \\ &= |f(x_2) - f(x_1)| + \ldots + |f(x_{n+1}) - f(x_n)| \\ &= |1-0| + |0-1| + \ldots + |1-0| \\ &= 1 + 1 + 1 + \ldots + 1 \\ &= n \end{aligned}$$

Consequently, $V(f,[a, b])$ is arbitrarily large and hence $V(f,[a, b]) = \infty$.

4.8.1 Algebraic Properties of Functions of Bounded Variations

Theorem 4.14: Let f and g be functions of bounded variation on $[a, b]$ and let k be a constant. Then,

1. f is bounded on $[a, b]$.
2. f is of bounded variation on every closed subinterval of $[a, b]$.
3. kf is of bounded variation on $[a, b]$.
4. $f + g$ and $f - g$ are of bounded variation on $[a, b]$.
5. fg is of bounded variation on $[a, b]$.
6. If $1/g$ is bounded on $[a, b]$, then f/g is of bounded variation on $[a, b]$.

Lemma: Let $f : [a, b] \to \mathbb{R}$ be a function. Let $\{x_i : 0 \leq i \leq n\}$ be a partition of $[a, b]$ and let $\{y_i : 0 \leq i \leq m\}$ be a partition of $[a, b]$ such that $\{x_i : 0 \leq i \leq n\} \subseteq \{y_i : 0 \leq i \leq m\}$. Then,

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \leq \sum_{i=1}^{m} |f(y_i) - f(y_{i-1})|$$

Proof: Start the calculation by adding one point to the partition $\{x_i : 0 \le i \le n\}$ which gives the desired result.

Let $\{x_i : 0 \le i \le n\}$ and $\{y_i : 0 \le i \le m\}$ be partitions as in the statement of the lemma. Suppose $y \in \{y_i : 0 \le i \le m\}$. If $y = x_j$ for some j then the sum does not change. Thus we assume that $y \ne x_j$ for all j. In this condition y falls between two points x_{k-1} and x_k in $\{x_i : 0 \le i \le n\}$ for some k. We take the sum $\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) |$ and write it as follows:

$$\sum_{i=1}^{k-1} | f(x_i) - f(x_{i-1}) | + | f(x_k) - f(x_{k-1}) | + \sum_{i=k+1}^{n} | f(x_i) - f(x_{i-1}) |$$

Take $| f(x_k) - f(x_{k-1}) |$. We know that,

$$\begin{aligned} | f(x_k) - f(x_{k-1}) | &= | f(x_k) - f(x_{k-1}) + f(y) - f(y) | \\ &\le | f(x_k) - f(y_i) | + | f(y_i) - f(x_{k-1}) | \end{aligned}$$

by the triangle inequality. We relabel the partition with the extra point as $\{x_i : 0 \le i \le n+1\}$. Thus, since all the addends are positive, we can write,

$$\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) | \le \sum_{i=1}^{n+1} | f(x_i) - f(x_{i-1}) |$$

Because there are atmost a finite number of the y_i, the desired result follows by induction.

Proof of Theorem 4.14: To prove (2) we begin by assuming that f is of bounded variation on $[a, b]$. Thus $V(f,[a, b]) = \sup\{\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) |\} = r$ where r is a real number. Let $[c, d]$ be a closed interval of $[a, b]$ and $\{x_i : 1 \le i \le n\}$ be a partition of $[c, d]$. Then extend this partition to $[a, b]$ by adding the point a and b and relabeling. Subsequently, $\{x_i : 0 \le i \le n+2\}$ is a partition of $[a, b]$ such that $x_1 = c$, $x_{n+1} = d$. Then,

$$\sum_{i=2}^{n+1} | f(x_i) - f(x_{i-1}) | \le | f(x_1) - f(a) | + \sum_{i=2}^{n+1} | f(x_i) - f(x_{i-1}) | + | f(b) - f(x_n) | \le r$$

Because the original partition of $[c, d]$ was arbitrary, we can conclude that $r \ge V(f,[c, d])$.

To prove (3), we assume that $\{x_i : 1 \le i \le n\}$ be a partition of $[a, b]$. Consider,

$$\sum_{i=1}^{n} | kf(x_i) - kf(x_{i-1}) | = | k | \sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) |$$

$$\le | k | V(f,[a, b])$$

NOTES

Because the partition was arbitrary, kf is of bounded variation. Further we can observe that $V(kf,[a, b]) = |k| V(f,[a, b])$.

To prove (4) we assume that $\{x_i : 1 \leq i \leq n\}$ be a partition of [a, b]. By repeated use of the triangle inequality, we obtain the equations of the form.

$$\sum_{i=1}^{n} | f(x_i) + g(x_i) - f(x_{i-1}) - g(x_{i-1}) |$$

$$\leq \sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) | + \sum_{i=1}^{n} | g(x_i) - g(x_{i-1}) |$$

$$\leq V(f,[a, b]) + V(g,[a, b])$$

Notice that, $V(f,[a, b]) + V(g,[a, b])$ is finite and the partition taken was arbitrary. Thus by the least upper bound axiom, $f + g$ is of bounded variation.

To prove that $f - g$ is of bounded variation on [a, b], we simply consider that $f - g = f + (-1)\, g$. Since $(-1)g$ is of bounded variation on [a, b], $f - g$ is of bounded variation on [a, b].

To prove (6) we assume that f and g are of bounded variation on [a, b] and that $1/g$ is bounded on [a, b]. Thus we know that there exists number M such that for all $x \in [a, b]$, $|1/g(x)| \leq M$. Now we have to show that $1/g$ is of bounded variation on [a, b]. We begin by taking $\{x_i : 0 \leq i \leq n\}$ as an arbitrary partition of [a, b] and consider the usual sum,

$$\sum_{i=1}^{n} \left| \frac{1}{g(x_i)} - \frac{1}{g(x_{i-1})} \right| = \sum_{i=1}^{n} \left| \frac{g(x_{i-1}) - g(x_i)}{g(x_i) g(x_{i-1})} \right|$$

$$\leq M^2 \sum_{i=1}^{n} | g(x_i) - g(x_{i-1}) |$$

$$\leq M^2 . V(g,[a, b])$$

Because the partition was arbitrary, we can state that the sum is bounded above by $M^2 . V(g,[a, b])$ and therefore by the least upper bound axiom, $1/g$ is of bounded variation.

Theorem 4.15: Let $f : \lceil a, b \rceil \to \mathbb{R}$ be a function and let $c \in (a, b)$. If f is of bounded variation on [a, c] and [c, b] then f is of bounded variation on [a, b] and $V(f,[a, b]) = V(f,[a, c]) + V(f,[c, b])$.

4.8.2 Functions of Bounded Variation as a Difference of Two Increasing Functions

Theorem 4.16: If $f : \lceil a, b \rceil \to \mathbb{R}$ is a function of bounded variation then there exist two increasing functions, f_1 and f_2 such that $f = f_1 - f_2$.

We define an increasing function f such that if $x_1 < x_2$ then $f(x_1) \leq f(x_2)$.

Lemma 1: For a function f, $V(f,[a, b]) = 0$ if and only if f is constant on $[a, b]$.

Proof: Suppose that f is constant. Then f is monotone function. Also, $V(f,[a, b]) = f(b) - f(a)$.

However, $f(b) = f(a)$ and so $V(f,[a, b]) = 0$.

We will prove the reverse by contraposition. Suppose that f is not constant on $[a, b]$. We can prove that $V(f,[a, b]) \neq 0$. Since f is not constant on $[a, b]$ there exist an x_1 and an x_2 such that both x_1 and x_2 are between a and b such that $f(x_1) \neq f(x_2)$. If we take these two points as a partition of $[a, b]$ we have,

$$V(f,[a, b]) \geq |f(x_1) - f(a)| + |f(x_2) - f(x_1)| + |f(b) - f(x_2)|$$

However, we know that $|f(x_1) - f(x_2)| > 0$. Since each other addend is at least 0, we state that the sum must be greater than 0 and thus $V(f,[a, b]) > 0$ and hence $V(f,\lceil a, b \rceil \neq 0$.

Lemma 2: If f is a function of bounded variation on $[a, b]$ and $x \in [a, b]$ then the function $g(x) = V(f,[a, x])$ is an increasing function.

Proof: We introduce x_1 and x_2 such that $x_1 < x_2$. We can show that $g(x_1) \leq g(x_2)$. Because f is of bounded variation on $[a, b]$,

$$\begin{aligned} V(f, [a, x_2]) &= V(f, [a, x_1]) + V(f, [x_1, x_2]) \\ V(f, [a, x_2]) - V(f, [a, x_1]) &= V(f, [x_1, x_2]) \\ g(x_2) - g(x_1) &= V(f, [x_1, x_2]) \end{aligned}$$

Since $V(f, [x_1, x_2]) \geq 0$ we define that $g(x_2) \geq g(x_1)$. Furthermore, by the above lemma, we have equality only if f is constant on $[x_1, x_2]$.

Proof of Theorem 4.16: We define $f_1 = V(f, [a, x])$ for $x \in (a, b)$ and $f_1(a) = 0$. This function can be increased by Lemma 2. Define f_2 as $f_2(x) = f_1(x) - f(x)$. Then $f = f_1 - f_2$. Now, show that f_2 is increasing.

Suppose that $a < x < y < b$. Using Theorem 4.14, we can write

$$\begin{aligned} f_1(y) - f_1(x) &= V(f, [x, y]) \\ &\geq |f(y) - f(x)| \\ &\geq f(y) - f(x) \end{aligned}$$

From this we see that

$$\begin{aligned} f_1(y) - f_1(x) &\geq f(y) - f(x) \\ f_1(y) - f(y) &\geq f_1(x) - f(x) \\ f_2(y) &\geq f_2(x) \end{aligned}$$

This shows that f_2 is increasing on $[a, b]$ and hence completes the proof.

Corollary: If $f:\lceil a, b \rceil \to \mathbb{R}$ is of bounded variation on $[a, b]$ then f is the difference of two strictly increasing functions.

Proof: We know from Theorem 4.16 that f can be written as the difference of two increasing functions. We call these functions f_1 and f_2 and write $f = f_1 - f_2$ where f_1 and f_2 are increasing.

Create two new functions, $g_1(x) = f_1(x) + x$ and $g_2(x) = f_2(x) + x$. Because both f_i and x are increasing functions, hence their sum is also increasing. However, since x is a strictly increasing function, the result of this addition is also a strictly increasing function. Thus we write,

$$f(x) = f_1(x) - f_2(x) = (f_1(x) + x) - (f_2(x) + x) = g_1(x) - g_2(x)$$

where g_1 and g_2 are strictly increasing functions.

4.9 LEBESGUE DIFFERENTIATION THEOREM

In real analysis, the Lebesgue differentiation theorem is a theorem of real analysis, which states that for almost every point, the value of an integrable function is the limit of infinitesimal averages taken about the point. The theorem is named After Henri Lebesgue.

For a Lebesgue integrable real or complex valued function f on $\mathbf{R}^n$, the indefinite integral is a set function which maps a measurable set A to the Lebesgue integral of $f \cdot \mathbf{1}_A$, where $\mathbf{1}_A$ denotes the characteristic function of the set A. It is usually written as,

$$A \mapsto \int_A f \, \mathrm{d}\lambda,$$

with λ the n–dimensional Lebesgue measure. The derivative of this integral at x is defined to be,

$$\lim_{B \to x} \frac{1}{|B|} \int_B f \, \mathrm{d}\lambda,$$

Where $|B|$ denotes the volume (i.e., the Lebesgue measure) of a ball B centered at x, and $B \to x$ means that the diameter of B tends to 0.

Proof: The theorem in its stronger form, that almost every point is a Lebesgue point of a locally integrable function f, can be proved as a consequence of the weak–L^1 estimates for the Hardy–Littlewood maximal function. The proof given below follows the standard derivation that can be found in Benedetto & Czaja (2009), Stein & Shakarchi (2005), Wheeden & Zygmund (1977) and Rudin (1987).

Since the statement is local in character, f can be assumed to be zero outside some ball of finite radius and hence integrable. It is then sufficient to prove that the set,

$$E_\alpha = \left\{ x \in \mathbf{R}^n : \limsup_{|B| \to 0,\, x \in B} \frac{1}{|B|} \left| \int_B f(y) - f(x) \, \mathrm{d}y \right| > 2\alpha \right\}$$

has measure 0 for all $\alpha > 0$.

Let ε > 0 be given. Using the density of continuous functions of compact support in $L^1(R^n)$, one can find such a function g satisfying,

$$\|f-g\|_{L^1} = \int_{\mathbf{R}^n} |f(x)-g(x)|\,\mathrm{d}x < \varepsilon.$$

It is then helpful to rewrite the main difference as,

$$\frac{1}{|B|}\int_B f(y)\,\mathrm{d}y - f(x) = \Big(\frac{1}{|B|}\int_B (f(y)-g(y))\,\mathrm{d}y\Big) + \Big(\frac{1}{|B|}\int_B g(y)\,\mathrm{d}y - g(x)\Big) + (g(x)-f(x)).$$

The first term can be bounded by the value at x of the maximal function for $f-g$, denoted here by $(f-g)^*(x)$:

$$\frac{1}{|B|}\int_B |f(y)-g(y)|\,\mathrm{d}y \le \sup_{r>0}\frac{1}{|B_r(x)|}\int_{B_r(x)} |f(y)-g(y)|\,\mathrm{d}y = (f-g)^*(x).$$

The second term disappears in the limit since g is a continuous function, and the third term is bounded by $|f(x)-g(x)|$. For the absolute value of the original difference to be greater than 2α in the limit, at least one of the first or third terms must be greater than α in absolute value. However, the estimate on the Hardy–Littlewood function says that,

$$\Big|\{x:(f-g)^*(x)>\alpha\}\Big| \le \frac{A_n}{\alpha}\|f-g\|_{L^1} < \frac{A_n}{\alpha}\varepsilon,$$

for some constant A_n depending only upon the dimension n. The Markov inequality (also called Tchebyshev's inequality) says that,

$$\Big|\{x:|f(x)-g(x)|>\alpha\}\Big| \le \frac{1}{\alpha}\|f-g\|_{L^1} < \frac{1}{\alpha}\varepsilon$$

Where,

$$|E_\alpha| \le \frac{A_n+1}{\alpha}\varepsilon.$$

Since ε was arbitrary, it can be taken to be arbitrarily small and the theorem follows.

A special case of the Lebesgue differentiation theorem is the Lebesgue density theorem, which is equivalent to the differentiation theorem for characteristic functions of measurable sets. The density theorem is usually proved using a simpler method. The Vitali covering lemma is vital to the proof of this theorem; its role lies in proving the estimate for the Hardy–Littlewood maximal function. The theorem also holds if balls are replaced, in the definition of the derivative, by families of sets with diameter tending to zero satisfying the Lebesgue's regularity condition, defined as family of sets with bounded eccentricity. This follows since the same substitution can be made in the statement of the Vitali covering lemma.

NOTES

Check Your Progress

7. Define the integration of non-negative functions.
8. State the general integral.
9. What do you understand by the integration of series?
10. Define the Reimann and Lebesgue integrals.
11. What are the derivatives?
12. Define the functions of bounded variation.
13. State the Lebesgue differentiation theorem.

4.10 ANSWERS TO 'CHECK YOUR PROGRESS'

1. In the measure theory, the term Lebesgue measure is named after French mathematician Henri Lebesgue. Fundamentally, the Lebesgue measure is defined as the standard method used to assign a measure to subsets of n-dimensional Euclidean space. In real analysis, the Lebesgue measure is specifically used to define the Lebesgue integration. Sets to which the Lebesgue measure can be assigned are termed as the 'Lebesgue Measurable'; characteristically the measure of the Lebesgue measurable set A can be denoted by $\lambda(A)$.
2. The Lebesgue outer measure can be defined assuming that there is an outer measure λ on a set X which is a measure then it will be considered as additive. Specifically, for given any two sets $A, B \subseteq X$ it can be stated that both $A \cap B$ and $A \cap B^c$ are disjoint in conjunction with $(A \cap B) \cap (A \cap B^c) = A$ and accordingly we can state that,

   $$\lambda (A) = \lambda (A \cap B) + \lambda (A \cap B^c)$$
3. The pair (X, Σ) is termed as a measurable space, the members of Σ are termed as the measurable sets. A measurable set X is known as a null set if $\mu (X) = 0$. A subset of a null set is described as a negligible set. A negligible set must not be measurable, but every measurable negligible set is certainly and inevitably a null set. A measure is termed complete if every negligible set is measurable.
4. A measure can be extended to a complete or perfect by means of considering the σ-algebra of subsets Y which vary through a negligible set from a measurable set X, i.e., the symmetric difference of X and Y is contained in a null set, such that $\mu (Y)$ can be defined to equal $\mu (X)$.
5. Specifically, in the measure theory, a measurable function is a function between the underlying sets of two measurable spaces that preserves the structure of the spaces, the preimage of any measurable set is measurable. In real analysis, measurable functions are used in the definition of the Lebesgue integral. In probability theory, a measurable function on a probability space is termed as a random variable.

6. The map $f : X \to Y$ between two topological spaces is termed as the 'Borel or Borel Measurable' if $f^{-1}(A)$ is a Borel set for any open set A as per the σ-algebra of Borel sets of X is the smallest σ-algebra containing the open sets.

7. Let s be a non-negative F measurable simple function so that, $s = \sum_{i=1}^{N} a_i X_{A_i}$ with disjoint F measurable sets A_i, $\bigcup_{i=1}^{N} A_i = X$ and $a_i \geq 0$. For any $E \in F$ define the integral of f over E to be, $I_E(s) = \sum_{i=1}^{N} a_i \mu(A_i \cap E)$ with the convention that if $a_i = 0$ and $\mu(A_i \cap E) = +\infty$ then $0 \times (+\infty) = 0$. Therefore, the area under $s \equiv 0$ in R is zero.

8. A measurable function f is said to be integrable over E if f^+ and f^- are both integrable over E. In this condition we define, $\int_E f = \int_E f^+ - \int_E \overline{f}$

9. Consider an integer N and a function f defined on the unbounded interval $[N, \infty)$, on which it is monotonic decreasing. Then the infinite series,

$$\sum_{n=N}^{\infty} f(n)$$

Converges to a real number if and only if the improper integral,

$$\int_N^{\infty} f(x)\, dx$$

is finite. In particular, if the integral diverges, then the series diverges as well.

10. The Lebesgue integral is named after Henri Lebesgue (1875–1941), typically Lebesgue defined and established the integral in the year 1904. Principally, the Lebesgue integral functions have a significant role in the theory and derivation of probability, real analysis, and numerous other fields in mathematics. The Riemann integral specifies that by partitioning the domain of an assigned function, one can approximate or estimate the assigned function by means of piecewise constant functions in each sub-interval. On the contrary, the Lebesgue integral are specifically used to partition the range of that function. The key objective of the Lebesgue integral is to provide an integral notion in which the limits of integrals hold moderate assumptions.

11. In real analysis, the derivative of a function of a real variable measures the sensitivity to change of the function value (output value) with respect to a change in its argument (input value). Derivatives are a fundamental tool of calculus. For example, the derivative of the position of a moving object with respect to time is the object's velocity, this measures how quickly the position of the object changes when time advances.

12. Let, S be a non-empty set of real numbers.
    (a) Suppose that S is bounded above. The number β is the supremum of S if β is an upper bound of S and there is no number less than β that is an upper bound of S. We write, $\beta = \sup S$.
    (b) Suppose that S is bounded below. A number α is the infimum of S if α is a lower bound of S and there is no number greater than α that is a lower bound of S. We write, $\alpha = \inf S$.
13. In real analysis, the Lebesgue differentiation theorem is a theorem of real analysis, which states that for almost every point, the value of an integrable function is the limit of infinitesimal averages taken about the point. The theorem is named after Henri Lebesgue.

4.11 SUMMARY

- In the measure theory, the term Lebesgue measure is named after French mathematician Henri Lebesgue. This measure was described by Henri Lebesgue in the year 1901, and in the year 1902 by the description Lebesgue integral.
- Both the Lebesgue measure and the Lebesgue integral were published in his dissertation thesis in the year 1902.
- Fundamentally, the Lebesgue measure is defined as the standard method used to assign a measure to subsets of n-dimensional Euclidean space. Additionally, for $n = 1, 2$ or 3, the Lebesgue measure coincides with the standard universal measure of length, area or volume.
- Generally, in the mathematical analysis the measure is also termed as the n-dimensional volume, n-volume or merely only the volume. In real analysis, the Lebesgue measure is specifically used to define the Lebesgue integration.
- Sets to which the Lebesgue measure can be assigned are termed as the 'Lebesgue Measurable'; characteristically the measure of the Lebesgue measurable set A can be denoted by $\lambda(A)$.
- Lebesgue outer measure: Assume that there is an outer measure λ on a set X which is a measure then it will be considered as additive. Specifically, for given any two sets $A, B \subseteq X$ we can state that both $A \cap B$ and $A \cap B^c$ are disjoint in conjunction with $(A \cap B) \cup (A \cap B^c) = A$ and accordingly we can state that,

  $$\lambda (A) = \lambda (A \cap B) + \lambda (A \cap B^c)$$
- Sets to which the Lebesgue measure can be assigned are termed as the 'Lebesgue Measurable'; characteristically, the measure of the Lebesgue measurable set A can be denoted by $\lambda (A)$.
- The pair (X, Σ) is termed as a measurable space, the members of Σ are termed as the measurable sets.
- A measurable set X is known as a null set if $\mu (X) = 0$. A subset of a null set is described as a negligible set. A negligible set must not be measurable, but every measurable negligible set is certainly and inevitably a null set.

- A measure is termed complete if every negligible set is measurable.
- A measure can be extended to a complete or perfect by means of considering the σ-algebra of subsets Y which vary through a negligible set from a measurable set X, i.e., the symmetric difference of X and Y is contained in a null set, such that $\mu(Y)$ can be defined to equal $\mu(X)$.
- In real analysis, measurable functions are used in the definition of the Lebesgue integral.
- In real analysis and measure theory, the regularity theorem for Lebesgue measure is defined as an acquired result which states that Lebesgue measure on the real line is a regular measure. Usually, this indicates about the real line and states that every Lebesgue measurable subset is 'Approximately Open' and 'Approximately Closed'.
- In mathematical analysis and in particular in the measure theory, a Borel measure on a topological space is defined as a measure for all open sets and consequently on all Borel sets.
- Consider that X be a locally compact Hausdorff space and also consider that $\mathcal{B}(X)$ be the smallest σ-algebra which contains or includes the open sets of X; then this is established as the 'σ-Algebra of Borel Sets'.
- The 'Borel Measure' is specified as any measure μ defined precisely on the σ-algebra of Borel sets. Some of the mathematicians define that μ is locally finite which implies that $\mu(C) < \infty$ for every compact set C.
- When a Borel measure μ is both inner regular and outer regular, then it is termed as a 'Regular Borel Measure'. If μ is both inner regular and outer regular, and is also locally finite, then in this condition it is known as a Radon measure.
- Characteristically, the real line $\mathbb{R}$ with its normal topology is defined as a locally compact Hausdorff space, therefore a Borel measure can be defined on it. In this instance, $\mathcal{B}(\mathbb{R})$ is referred as the smallest σ-algebra that comprises of the open intervals of $\mathbb{R}$.
- Principally, the Lebesgue σ-algebra is essentially stated as the completion of the Borel σ-algebra, which implies that it is the smallest σ-algebra that comprises of all the Borel sets and has a complete measure on it.
- Characteristically, the measurable sets on the line are considered as the iterated countable unions and intersections of intervals, termed as the Borel sets, are referred as the plus-minus null sets.
- The measure of the union of two disjoint sets to be the sum of the measure of the two sets. A measure with this natural property is termed as the finitely additive.
- Let s be a non-negative F measurable simple function so that, $s = \sum_{i=1}^{N} a_i X_{A_i}$ with disjoint F measurable sets A_i, $\bigcup_{i=1}^{N} A_i = X$ and $a_i \geq 0$. For any $E \in F$ define the integral of f over E to be, $I_E(s) = \sum_{i=1}^{N} a_i \mu(A_i \cap E)$ with the

NOTES

convention that if $a_i = 0$ and $\mu(A_i \cap E) = +\infty$ then $0 \times (+\infty) = 0$. Therefore, the area under $s \equiv 0$ in R is zero.

- If $f : X \to R^+$ is a non-negative F measurable function, $E \in F$, then the integral of f over E is,

  $$\int_E f d\mu = \sup\{I_E(s) : s \text{ is a simple } F\text{-measurable function, } 0 \le s \le f\}$$

  But, if $E \neq X$ we need only that f is defined on some domain containing E.

- If a property P holds on all points in $E \setminus A$ for some set A with $\mu(A) = 0$ then P is said to hold almost everywhere (μ) on E. It is possible that P holds on some of the points of A or that the set of points on which P does not hold is non-measurable. But, if μ is a complete measure, such as the Lebesgue-Stieltjes measure μ_F, then the situation is simpler. Assume that a property P holds almost everywhere (μ) on E. The definition says that the set of points, D say, on which P does not hold, can be covered by a set of measure zero, i.e., there exists $A : D \subseteq A$ and $\mu(A) = 0$.

- A measurable function f is said to be integrable over E if f^+ and f^- are both integrable over E. In this condition we define, $\int\limits_E f = \int\limits_E f^+ - \int\limits_E \overline{f}$

- Any function which is Riemann integrable is Lebesgue integrable as well and positively with the same values for the two integrals. First, we recall one definition of Riemann integrability.

- Consider an integer N and a function f defined on the unbounded interval $[N, \infty)$, on which it is monotonic decreasing. Then the infinite series

  $$\sum_{n=N}^{\infty} f(n)$$

  Converges to a real number if and only if the improper integral

  $\int_N^{\infty} f(x)\, dx$ is finite. In particular, if the integral diverges, then the series diverges as well.

- In real analysis, the derivative of a function of a real variable measures the sensitivity to change of the function value (output value) with respect to a change in its argument (input value). Derivatives are a fundamental tool of calculus. For example, the derivative of the position of a moving object with respect to time is the object's velocity, this measures how quickly the position of the object changes when time advances.

- The derivative of a function of a single variable at a is taken as input value, when it exists, is the slope of the tangent line to the graph of the function at that point. The tangent line is the best linear approximation of the function near that input value. For this reason, the derivative is often described as the "Instantaneous Rate of Change", the ratio of the instantaneous change in the dependent variable to that of the independent variable.

- The process of finding a derivative is called differentiation. The reverse process is called antidifferentiation. The fundamental theorem of calculus relates antidifferentiation with integration. Differentiation and integration constitute the two fundamental operations in single variable calculus.
- Let, S be a non-empty set of real numbers.
  (a) Suppose that S is bounded above. The number β is the supremum of S if β is an upper bound of S and there is no number less than β that is an upper bound of S. We write, $\beta = \sup S$.
  (b) Suppose that S is bounded below. A number α is the infimum of S if α is a lower bound of S and there is no number greater than α that is a lower bound of S. We write, $\alpha = \inf S$.
- In real analysis, the Lebesgue differentiation theorem is a theorem of real analysis, which states that for almost every point, the value of an integrable function is the limit of infinitesimal averages taken about the point. The theorem is named after Henri Lebesgue.
- A special case of the Lebesgue differentiation theorem is the Lebesgue density theorem, which is equivalent to the differentiation theorem for characteristic functions of measurable sets. The density theorem is usually proved using a simpler method. The Vitali covering lemma is vital to the proof of this theorem; its role lies in proving the estimate for the Hardy–Littlewood maximal function.

4.12 KEY TERMS

- **Lebesgue outer measure:** Assume that there is an outer measure λ on a set X which is a measure then it will be considered as additive. Specifically, for given any two sets $A, B \subseteq X$ we can state that both $A \cap B$ and $A \cap B^c$ are disjoint in conjunction with $(A \cap B) \cup (A \cap B^c) = A$ and accordingly we can state that,

  $\lambda\,(A) = \lambda\,(A \cap B) + \lambda\,(A \cap B^c)$
- **Borel measure:** The 'Borel Measure' is specified as any measure μ defined precisely on the σ-algebra of Borel sets.
- **Regular Borel measure:** When a Borel measure μ is both inner regular and outer regular, then it is termed as a 'Regular Borel Measure'.
- **Integration of non-negative functions:** Let s be a non-negative F measurable simple function so that, $s = \sum_{i=1}^{N} a_i X_{A_i}$ with disjoint F measurable sets A_i, $\bigcup_{i=1}^{N} A_i = X$ and $a_i \geq 0$.
- **General integral:** A measurable function f is said to be integrable over E if f^+ and f^- are both integrable over E. In this condition we define, $\int_E f = \int_E f^+ - \int_E \overline{f}$

- **Integration of series:** Consider an integer N and a function f defined on the unbounded interval [N, ¥), on which it is monotonic decreasing. Then the infinite series $\sum_{n=N}^{\infty} f(n)$
- **Reimann and Lebesgue integrals:** Any function which is Riemann integrable is Lebesgue integrable as well and positively with the same values for the two integrals.
- **Derivatives:** The derivative of a function of a real variable measures the sensitivity to change of the function value (output value) with respect to a change in its argument (input value). Derivatives are a fundamental tool of calculus.
- **Lebesgue differentiation theorem:** The Lebesgue differentiation theorem is a theorem of real analysis, which states that for almost every point, the value of an integrable function is the limit of infinitesimal averages taken about the point. The theorem is named after Henri Lebesgue.

4.13 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Why is Lebesgue outer measure used?
2. What are the measurable sets?
3. Define the terms regularity and measurable functions.
4. State about the Borel and Lebesgue measurability.
5. Why are non-measurable sets used?
6. Define integration of non-negative functions.
7. What is the general integral?
8. What is the significance of the integration of series?
9. State about the Riemann and Lebesgue integrals.
10. What are the derivatives?
11. Define the term functions of bounded variation.
12. State the Lebesgue differentiation theorem.

Long-Answer Questions

1. Explain the Lebesgue measure and Lebesgue outer measure giving examples.
2. Briefly discuss the measurable sets giving appropriate examples.
3. Describe regularity and measurable functions with the help of examples.
4. Explain the three significant and distinguished classes of the measurable functions in real analysis.

NOTES

5. Discuss in detail the Borel and Lebesgue measurability giving relevant examples.
6. Elaborate on the non-measurable sets.
7. Briefly explain the integration of non-negative functions giving theorems and proofs.
8. What is general integral? Explain giving examples.
9. Briefly discuss about the integration of series in real analysis.
10. Discuss the significance of Riemann and Lebesgue integrals giving appropriate examples.
11. Describe the derivatives on the basis of differentiation and integration.
12. Elaborate on the functions of bounded variation giving examples.
13. State and prove the Lebesgue differentiation theorem.
14. Calculate the outer measure of the set of irrational numbers in the interval [0,1].
15. Assume that E has measure zero where $E \subseteq \mathbb{R}$. Prove that the set $E^2 = \{x^2 : x \in E\}$ also has measure zero.
16. Prove that any measurable subset $A \subset \mathbb{R}$ with $\lambda(A) > 0$ contains a non-measurable subset.

4.14 FURTHER READING

Rudin, Walter. 2017. *Real and Complex Analysis*, Third Edition. Noida: McGraw-Hill Education.

Gupta, S. L. and Nisha Rani. 2004. *Fundamental Real Analysis*, Fourth Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, First Edition. Cambridge (U.K.): Cambridge University Press.

Bartle, Robert G. and Donald R. Sherbert. 2014. *Introduction to Real Analysis*, Fourth Edition. New York: Wiley.

Trench, William F. 2002. *Introduction to Real Analysis*, London: Pearson.

Loeb, Peter A. 2016. *Real Analysis*, Basel (Switzerland): Birkhäuser.

Royden, Halsey. 2015. *Real Analysis*, Fourth Edition. Noida: Pearson Education India.

UNIT 5 MEASURES AND OUTER MEASURES

NOTES

Structure

5.0 INTRODUCTION

In the measure theory, the concept of a measure is a generalization of common notions, such as mass, distance/length, area, volume, etc. The perception behind this concept dates back to Ancient Greece when Archimedes tried to calculate the area of a circle. The foundations of modern measure theory were the significant theories and notations of Émile Borel, Henri Lebesgue, Nikolai Luzin, Johann Radon, Constantin Carathéodory and Maurice Fréchet.

Although, measures can be defined on arbitrary collections of sets, the most natural domain of a measure is a σ-ring. It generalizes the intuitive notions of length, area, and volume. The earliest and most important examples are Jordan measure and Lebesgue measure, but other examples are Borel measure, probability measure, complex measure, and Haar measure.

An outer measure or exterior measure is a function defined on all subsets of a given set with values in the extended real numbers satisfying some additional technical conditions. The theory of outer measures was first introduced by Constantin Carathéodory to provide an abstract basis for the theory of measurable sets and countably additive measures. Carathéodory's work on outer measures found many applications in measure basically in theoretic set theory because the outer measures are used in the proof of the fundamental Carathéodory's extension theorem, and was used in an essential method by Hausdorff to define a dimension,

NOTES

such as metric invariant now called Hausdorff dimension. Outer measures are commonly used in the field of geometric measure theory.

A measure space is a fundamental object of measure theory that typically studies generalized notions of volumes. It contains an underlying set, the subsets of this set that are feasible for measuring the 'σ-Algebra' and the method that is used for measuring the 'Measure'. A measurable space consists of the first two components without a specific measure.

In mathematics, Jensen's inequality is named after the Danish mathematician Johan Jensen which relates the value of a convex function of an integral to the integral of the convex function. Principally, the Jensen's inequality was proved by Jensen in 1906. Given its generality, the inequality appears in many forms depending on the context. In its simplest form the inequality states that the convex transformation of a mean is less than or equal to the mean applied after convex transformation; it is a simple corollary that the opposite is true of concave transformations. The classical form of Jensen's inequality involves several numbers and weights. The inequality can be stated quite generally using either the language of measure theory or equivalently the probability.

In mathematical analysis, Hölder's inequality, named after Otto Hölder, is a fundamental inequality between integrals and an indispensable tool to study and analyse the L^p spaces. Hölder's inequality is used to prove the Minkowski inequality which is the triangle inequality in the space L^p (μ). The Minkowski inequality establishes that the L^p spaces are normed vector spaces. The Minkowski inequality is named after the German mathematician Hermann Minkowski.

Convergence in measure is either of two distinct mathematical concepts both of which generalize the concept of convergence in probability. On a finite measure space, both the notions are equivalent. Otherwise, convergence in measure can refer to either global convergence in measure or local convergence in measure. In the mathematical field of analysis, the uniform convergence is a mode of convergence of functions which are stronger than pointwise convergence.

In this unit, you will study about the measures and outer measures, extension of a measure, uniqueness of extension, completion of a measure, measure spaces, integration with respect to a measure, the L^p-spaces, convex functions, Jensen's inequality, Hölder and Minkowski inequalities, completeness of L^p, convergence in measure, and almost uniform convergence.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the measures and outer measures
- Define the extension of a measure
- Understand the uniqueness of extension
- Explain the completion of a measure
- Elaborate on the measure spaces
- Comprehend on the integration with respect to a measure

- Define the L^P-spaces, completeness of L^P and convex functions
- State the Jensen's inequality
- Analyse the Hölder and Minkowski inequalities
- Define the convergence in measure
- Explain the almost uniform convergence

NOTES

5.2 MEASURES AND OUTER MEASURES

In the measure theory, the concept of a measure is a generalization of common notions, such as mass, distance/length, area, volume, etc. The perception behind this concept dates back to Ancient Greece when Archimedes tried to calculate the area of a circle. The foundations of modern measure theory were the significant theories and notations of Émile Borel, Henri Lebesgue, Nikolai Luzin, Johann Radon, Constantin Carathéodory and Maurice Fréchet.

An outer measure or exterior measure is a function defined on all subsets of a given set with values in the extended real numbers satisfying some additional technical conditions. The theory of outer measures was first introduced by Constantin Carathéodory to provide an abstract basis for the theory of measurable sets and countably additive measures.

Definition: An outer measure $\mu*$ is an extended real valued set function defined on all subsets of a space X having the following properties:

(a) $\mu*\phi = 0$

(b) $A \subset B \Rightarrow \mu*A \leq \mu*B$ (Monotonicity)

(c) $E \subset \sum_{i=1}^{\infty} E_i \Rightarrow \mu * E \leq \sum_{i=1}^{\infty} \mu * E_i$ (Subadditivity)

The outer measure $\mu*$ is said to be finite if $\mu*X < \infty$.

As per the Lebesgue measure we can state that a set E is measurable with respect to $\mu*$ if for every set A we have,

$$\mu * A = \mu * (A \cap E) + \mu * (A \cap E^C)$$

Because $\mu*$ is subadditive therefore in order to explain that E is measurable, we just have to prove that,

$$\mu * A \geq \mu * (A \cap E) + \mu * (A \cap E^C), \text{ for every } A.$$

When $\mu*A = \infty$, then this inequality holds trivially. Consequently, this can be proved for sets A with $\mu*A$ finite.

Theorem 5.1: The class β of $\mu*$-measurable sets are σ-algebra. If $\bar{\mu}$ is restricted to β, then $\bar{\mu}$ is a complete measure on β.

Proof: It is obvious that the empty set is measurable. Using the definition of measurability in E and E^c, we have that E^c is measurable whenever E is measurable. Now, consider that E_1 and E_2 be measurable sets, then by the measurability of E_2, we can state that,

NOTES

$$\mu^* A = \mu^*(A \cap E_2) + \mu^*(A \cap E_2{}^C)$$

And by the measurability of E_1,

$$\mu^* A = \mu^*(A \cap E_2) + \mu^*(A \cap E_2{}^C \cap E_1) + \mu^*(A \cap E_1^C \cap E_2{}^C)$$

Now, since

$$A \cap [E_1 \cup E_2] = [A \cap E_2] \cup [A \cap E_1 \cap E_2{}^C]$$

We have,

$$\mu^*(A \cap [E_1 \cup E_2]) \leq \mu^*(A \cap E_2) + \mu^*(A \cap E_2{}^C \cap E_1)$$

By using the subadditivity properly, we can state that,

$$\mu^* A \geq \mu^*(A \cap [E_1 \cup E_2]) + \mu^*(A \cap E_1^C \cap E_2{}^C)$$

This implies that $E_1 \cup E_2$ is measurable. Therefore, we can state that the union of two measurable sets is measurable. But by induction, the union of any finite number of measurable sets is measurable. Hence, β is an algebra of sets. Suppose, $E = \cup E_i$, where $<E_i>$ is a disjoint sequence of measurable sets and hence,

$$G_n = \bigcup_{i=1}^{n} E_i$$

Then G_n is measurable, and

$$\mu^* A = \mu^*(A \cap G_n) + \mu^*(A \cap G_n{}^C) \;\; \geq \mu^*(A \cap G_n) + \mu^*(A \cap E^C)$$

Because $E^c \subset G_n{}^c$.

Now, $G_n \cap E_n = E_n$ and $G_n \cap E_n{}^c = G_{n-1}$, and by the measurability of E_n we have,

$$\mu^*(A \cap G_n) = \mu^*(A \cap E_n) + \mu^*(A \cap G_{n-1})$$

By induction we have, $\mu^*(A \cap G_{n-1}) = \mu^*(A \cap E_{n+1}) + \mu^*(A \cap E_{n-2})$.

Also, we can state that,

$$\mu^*(A \cap G_n) = \sum_{i=1}^{n} \mu^*(A \cap E_i)$$

And so,

$$\mu^* A \geq \mu^*(A \cap E^C) = \sum_{i=1}^{\infty} \mu^*(A \cap E_i) \;\; \geq \mu^*(A \cap E^C) + \mu^*(A \cap E)$$

Since, $A \cap E \subset \bigcup_{i=1}^{\infty} (A \cap E_i)$

Thus, E is measurable.

Since, the union of any sequence of sets in an algebra which can be replaced by a disjoint union of sets in an algebra, it follows that β is a σ-algebra.

NOTES

Let us now prove that $\bar{\mu}$ is finitely additive. Let E_1 and E_2 be disjoint measurable sets.

Then, the measurability of E_2 implies that,

$$\bar{\mu}(E_1 \cup E_2) = \mu^*(E_1 \cup E_2)$$

$$= \bar{\mu}^*([E_1 \cup E_2] \cap E_2 + \mu^*([E_1 \cup E_2] \cap E_2^C)$$

$$= \mu^* E_2 + \mu^* E_1$$

Consequently, finite additivity follows by the induction.

If E is the disjoint union of the measurable sets $\{E_i\}$, then

$$\overline{\mu_i} E \geq \bar{\mu}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \bar{\mu} E_i$$

And so,

$$\bar{\mu}\, E \geq \sum_{i=1}^{\infty} \bar{\mu}\, E_i$$

But,

$\bar{\mu}\, E \leq \sum_{i=1}^{\infty} \bar{\mu}\, E_i$ by the subadditivity of $\mu*$. Hence, $\bar{\mu}$ is countably additive.

So $\bar{\mu}$ is a measure since it is non-negative and $\bar{\mu}\,\phi = \mu*\phi = 0$.

Example 5.1: If E_1 and E_2 are measurable then prove that $E_1 \cap E_2$.

Solution: As per the definition of measurability, we can stat that,

A subset E of X is called measurable whenever,

$\mu\,(A) = \mu\,(A \cap E) + \mu\,(A \cap E^c)$ holds for all A subset of X.

To be a measure it satisfies the following properties:

1. $\mu\,(\emptyset) = 0$
2. $\mu\,(A) \leq \mu\,(B)$ if $A \subset B$, i.e., μ is monotone.
3. $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu(E_i)$ holds for every sequence of subsets E_i of X, i.e., μ is subadditive.

5.3 EXTENSION OF A MEASURE

A measure on an algebra is defined as a non-negative extended real valued set function μ which is typically defined on an algebra $\boldsymbol{A}$ of sets such that,

(a) $\mu\phi = 0$

(b) If $<A_i>$ is a disjoint sequence of sets in $\boldsymbol{A}$ whose union is also in $\boldsymbol{A}$, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu A_i$$

NOTES

Therefore, a measure on an algebra $\boldsymbol{A}$ is a measure iff $\boldsymbol{A}$ is a σ-algebra.

We can construct an outer measure μ∗ and prove that the measure $\bar{\mu}$ is an extension of measure μ defined on an algebra.

The extension measure $\mu{*}E = \inf \sum_{i=1}^{\infty} \mu A_i$, where $<A_i>$ ranges over all sequence from $\boldsymbol{A}$ such that,

$$E \subset \bigcup_{i=1}^{\infty} A_i$$

Lemma 1: If $A \in \boldsymbol{A}$ and if $<A_i>$ is any sequence of sets in $\boldsymbol{A}$, such that $A \subset \bigcup_{i=1}^{\infty} A_i$, then show that $\mu A \le \sum_{i=1}^{\infty} \mu A_i$.

Proof: Consider that $B_n = A \cap A_n \cap A^C{}_{n-1} \cap ... \cap A_i^C$. Then $B_n \in \boldsymbol{A}$ and $B_n \subset A_n$. But since A is the disjoint union of the sequence $<B_n>$, by countable additivity we have,

$$\mu\, A = \sum_{n=1}^{\infty} \mu B_n \le \sum_{n=1}^{\infty} \mu\, A_n$$

Corollary: If $A \in \boldsymbol{A}$, then prove that $\mu{*}A = \mu A$.

Consequently, we have

$$\mu\, A \le \sum_{n=1}^{\infty} \mu A_n < \mu{*}\, A + \varepsilon$$

Or,

$$\mu\, A \le \mu{*}\, A + \varepsilon$$

Subsequently, because ε is arbitrary, therefore, we have

$$\mu\, A \le \mu{*}\, A$$

Also, by definition

$$\mu{*}\, A \le \mu\, A$$

Therefore,

$\mu{*}A = \mu A$. Hence proved.

Lemma 2: The set function μ∗ is an outer measure.

Proof: From the given definition, μ∗ is a monotone non-negative set function defined for all sets and $\mu{*}\phi = 0$. Now we have to prove that it is countably subadditive. Let $E \subseteq \bigcup_{i=1}^{\infty} E_i$. If $\mu{*}E_i = \infty$ for any i, then we have,

$\mu{*}E \le \sum \mu{*}E_i = \infty$.

NOTES

If $\mu * E_i \neq \infty$, then given $\varepsilon > 0$, and there exists for each i a sequence $< A_{ij} >_{j=1}^{\infty}$ of sets in $\boldsymbol{A}$ such that $E_i \subset \bigcup_{j=1}^{\infty} A_{ij}$ and

$$\sum_{j=1}^{\infty} \mu A_{ij} < \mu * E_i + \frac{\varepsilon}{2^i}$$

Then,

$$\mu * E \leq \sum_{i,j} \mu A_{ij} < \sum_{i=1}^{\infty} \mu * E_i + \varepsilon$$

Since ε is an arbitrary positive number, therefore we have,

$$\mu * E \leq \sum_{i=1}^{\infty} \mu * E_i$$

Hence it proves that $\mu*$ is subadditive.

Lemma 3: If $A \in \boldsymbol{A}$, then A is measurable with respect to $\mu*$.

Proof: Consider that E be an arbitrary set of finite outer measure and ε be a positive number, then there is a sequence $<A_i>$ from $\boldsymbol{A}$, such that $E \subset \cup A_i$, and $\Sigma \mu\, A_i < \mu * E + \varepsilon$.

By the additivity of μ on $\boldsymbol{A}$, we have

$$\mu(A_i) = \mu(A_i \cap A) + \mu(A_i \cap A^C)$$

Hence,

$$\mu * E + \varepsilon > \sum_{i=1}^{\infty} \mu(A_i \cap A)$$

$$\sum_{i=1}^{\infty} \mu(A_i \cap A^C) > \mu * (E \cap A) + \mu * (E \cap A^C)$$

Because,

$$E \cap A \subset \cup (A_i \cap A)$$

And

$$E \cap A^C \subset \cup (A_i \cap A^C)$$

Since ε is an arbitrary positive number, therefore we have

$$\mu * E \geq \mu * (\varepsilon \cap A) + \mu * (E \cap A^C)$$

Hence proved that A is $\mu*$-measurable.

Note: The outer measure $\mu*$ which we have defined above is known as the outer measure induced by μ.

Notation: For a given algebra $\boldsymbol{A}$ of sets we use $\boldsymbol{A}_\sigma$ to denote those sets which are countable unions of sets of $\boldsymbol{A}$ and use $\boldsymbol{A}_{\sigma\delta}$ to denote those sets which are countable intersection of sets in $\boldsymbol{A}_\sigma$.

NOTES

Theorem 5.2: Let μ be a measure on an algebra $\boldsymbol{A}$, $\mu*$ be the outer measure induced by μ and E be any set. Then for $\varepsilon > 0$, there exists a set $A \in \boldsymbol{A}_\sigma$ with $E \subset A$ and $\mu * A \le \mu * E + \varepsilon$.

There is also a set $B \in \boldsymbol{A}_{\sigma\delta}$ with $E \subset B$ and $\mu*E = \mu*B$.

Proof: From the definition of $\mu*$ there is a sequence $<A_i>$ from $\boldsymbol{A}$ such that $E \subset \cup A_i$ and,

$$\sum_{i=1}^{\infty} \mu A_i \le \mu * E + \varepsilon \qquad \ldots (5.1)$$

Take, $A = \cup A_i$

Then, $\mu * A \le \Sigma \mu * A_i$

$$= \Sigma \mu A_i \qquad \ldots (5.2)$$

Because $\mu*$and μ agree on members of $\boldsymbol{A}$ by the above mentioned corollary.

hence, Equations (5.1) and (5.2) imply that

$\mu * A \le \mu * E + \varepsilon$

which proves the first part of the Theorem 5.2.

To prove the second statement of the Theorem 5.2, consider that for each positive integer n there is a set A_n in $\boldsymbol{A}_\sigma$, such that, $E \subset A_n$ and

$\mu * A_n < \mu * E + \dfrac{1}{n}$ (From First Part Proved Above)

Let $B = \cap A_n$. Then, $B \in \boldsymbol{A}_{\sigma\delta}$ and $E \subset B$. Since $B \subset A_n$, therefore

$\mu * B < \mu * A_n \le \mu * E + \dfrac{1}{n}$

Since n is arbitrary, then by monotonicity, $\mu*B \le \mu*E$.

Hence proved that $\mu*B = \mu*E$.

5.4 UNIQUENESS OF EXTENSION

In the measure theory of real analysis, the Carathéodory's extension theorem states that, "Any premeasure defined on a given ring R of subsets of a given set Ω can be extended to a measure on the σ-algebra generated by R, and this **extension** is **unique** if the premeasure is **σ-finite"**. The Carathéodory's extension theorem is named after the Greek mathematician Constantin Carathéodory. Accordingly, any premeasure on a ring that contains all intervals of real numbers can be typically extended to the Borel algebra of the set of real numbers, and this exceptionally effective conclusion of measure theory indicates to the Lebesgue measure.

The Carathéodory's extension theorem is also occasionally termed as the Carathéodory-Fréchet extension theorem, the Carathéodory–Hopf extension theorem, the Hopf extension theorem and the Hahn–Kolmogorov extension theorem.

Fundamentally, the most simple statement of the Carathéodory's extension theorem is often termed as the Hahn–Kolmogorov theorem.

Consider that Σ_0 be an algebra of subsets of a set X, and also consider that there is a function,

$$\mu_0 : \Sigma_0 \to [0, \infty]$$

This is finitely additive, and specifies that,

$$\mu_0\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mu_0(A_n)$$

This is for any positive integer N and for $A_1, A_2, \ldots, A_N$ disjoint sets in Σ_0.

We assueme that this function satisfies the Sigma (σ) additivity assumption of the form,

$$\mu_0\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu_0(A_n)$$

This is for any disjoint family $\{A_n : n \in \mathbb{N}\}$ of elements of Σ_0 such that $\cup_{n=1}^{\infty} A_n \in \Sigma_0$.

The function μ_0 which conforms or obeys these two properties are termed as the premeasures.

Subsequently, we can state that the μ_0 extends to a specific measure, which is defined on the σ-algebra Σ generated by Σ_0; i.e., there exists a measure of the form,

$$\mu : \Sigma \to [0, \infty]$$

such that its constraint or limitation to Σ_0 coincides with μ_0.

When μ_0 is σ-finite, then the extension is defined as unique.

The Carathéodory's extension theorem considered very significant as it helps in constructing a measure by defining it on a small algebra of sets, so that its sigma additivity can be verified. Additionally, this theorem also ensures its extension to a σ-algebra.

Theorem 5.3 Unique Extension Theorem: Any set function P defined on a field $\mathcal{F}_0$ of sets and satisfying the properties of a probability measure on $\mathcal{F}_0$ extends uniquely to a probability measure on the σ-field $\mathcal{F}$ generated by $\mathcal{F}_0$.

Proof: Suppose that there is a field $\mathcal{F}_0$ of sets on a space Ω, such as the finite unions of open sets.

Let P be a measure on $\mathcal{F}_0$.

Where $\mathcal{F}$ be the σ-field typically generated by $\mathcal{F}_0$.

NOTES

We have to define that there exists a unique extension of P from $\mathcal{F}$ to $\mathcal{F}_0$.

To prove the existence of unique extension, we will initially define the outer measure as an extension of P, as follows:

Assume that there is a given measure P on $\mathcal{F}_0$, then we will define its outer measure $P*$ on all subsets of Ω by,

$$P^*(A) = \inf\left\{\sum_{i=1}^{\infty} P(A_i) |\ A \subset \bigcup_{i=1}^{\infty} A_i;\ A_i \in \mathcal{F}_0\right\}$$

Specifically, the lowest sum of measures of a collection of $\mathcal{F}_0$ sets which contain A.

Additionally, the inner measure can also be defined as one minus the largest sum of measures of a collection of $\mathcal{F}_0$ sets which are contained in A^c, and defined as,

$$P_*(A) = \sup\left\{1 - \sum_{i=1}^{\infty} P(A_i) |\ A^c \subset \bigcup_{i=1}^{\infty} A_i;\ A_i \in \mathcal{F}_0\right\}$$

But this is equivalent to the following definition of inner measure:

$$P_*(A) = 1 - P^*(A^c)$$

The measure P can be extended on $\mathcal{F}_0$ to a collection $\mathcal{F}$ of feasibly as many sets as possible.

Consequently, we can define that $\mathcal{F}$ be the collection of sets $A \subset \Omega$ having the same inner measure and outer measure, and then subsequently we can define that $P(A)$.

Specifically, for all sets A such that.

$$P^*(A) = P_*(A)$$

$$P^*(A) + P^*(A^c) = 1$$

Subsequently, we can define such sets to be in $\mathcal{F}$ and define,

$$P(A) = P^*(A) = P_*(A)$$

5.5 COMPLETION OF A MEASURE

The term **complete measure** or more specifically a **complete measure space** is defined as a specific measure space wherein every single subset of every single null set is measurable, i.e., having measure zero. More appropriately, we can state that a measure space (X, Σ, μ) is termed as complete if and only if,

$$\boldsymbol{S \subseteq N \in \Sigma} \quad \text{and} \quad \boldsymbol{\mu(N) = 0 \Rightarrow S \in \Sigma}$$

The term **completeness** can be essentially illustrated by considering the typical product space problems. Assume that given is previously constructed Lebesgue measure on the real line, then in order to denote this measure space we

NOTES

use the notations ($\mathbb{R}$, B, λ). Further, some two-dimensional Lebesgue measure λ^2 can be constructed on the plane $\mathbb{R}^2$ as a product measure. Simply and certainly, this can be accomplished by taking the σ-algebra on $\mathbb{R}^2$ to be $B \otimes B$, which is the considered as the smallest σ-algebra that contains all the measurable 'Rectangles' $A_1 \times A_2$ for $A_1, A_2 \in B$.

While this approach does define a measure space, it has a flaw. Since every singleton set has one-dimensional Lebesgue measure zero,

$$\lambda^2 (\{0\} \times A) \leq \lambda (\{0\}) = 0$$

This implies for 'Any' subset A of $\mathbb{R}$. Even though, assume that A is a non-measurable subset of the real line, for example the Vitali set. In mathematics, a Vitali set was found by Giuseppe Vitali in 1905, basically, it is an elementary example of a set of real numbers that is not Lebesgue measurable.

Then we can state that the λ^2-measure of $\{0\} \times A$ is not defined, however,

$$\{0\} \times A \subseteq \{0\} \times \mathbb{R}$$

Remember that the given larger set also have λ^2-measure zero. Consequently, as defined above this 'Two-Dimensional Lebesgue Measure' is not complete, hence some completion procedure is essential.

Constructing a Complete Measure

Consider that a possibly incomplete measure space (X, Σ, μ) is given, then of this measure space there is an extension (X, Σ_0, μ_0), which is complete. The smallest of the extension, i.e., the smallest σ-algebra Σ_0 is termed as the completion of the measure space.

Using the following assumptions or statement the completion can be constructed:

- Let Z be the set of all the subsets of the zero-μ-measure subsets of X, instinctively those elements of Z which are already not in Σ are specifically the ones which prevent completeness from holding true.
- Let Σ_0 be the σ-algebra created or produced by Σ and Z, i.e., the smallest σ-algebra that contains every element of Σ and of Z.
- Let μ has an extension μ_0 to Σ_0, which is unique if μ is σ-finite, then it is called the outer measure of μ, given by the infimum.

$$\mu_0 (C) := \inf \{\mu (D) \mid C \subseteq D \in \Sigma\}$$

Then (X, Σ_0, μ_0) is referred as a complete measure space and is termed as the completion of (X, Σ, μ).

In the above given construction it can be explained that every member of Σ_0 is of the form $A \cup B$ for some $A \in \Sigma$ and some $B \in Z$, and

$$\mu_0 (A \cup B) = \mu (A)$$

Additionally, the Borel measure when defined on the Borel σ-algebra specifically created or produced by the open intervals of the real line is not complete, and therefore the above defined completion procedure has to be used for defining the complete Lebesgue measure. This can be explained and exemplified through the fact that the set of all Borel sets over the reals holds the equivalent cardinality

NOTES

as the reals. Even though the Cantor set is a Borel set, i.e., it has measure zero and its power set holds cardinality which is strictly or precisely greater than that of the reals. Consequently, we can state that there is a subset of the Cantor set which is not contained in the Borel sets. Therefore, the Borel measure is not complete.

The n-dimensional Lebesgue measure is defined as the completion of the n-fold product of the one-dimensional Lebesgue space with itself. It is similarly referred as the completion of the Borel measure as in the one-dimensional condition.

5.6 MEASURE SPACES

In mathematics, the term **measure space** is defined as a fundamental and essential object of measure theory which analyses the universal generalized and simplified notions of volumes. Characteristically, it comprises of an underlying set, which is referred as the subsets of this set that are feasible and sufficient for measuring the **'σ-Algebra'** and the method that is used for measuring the **'Measure'**. One significant example of a measure space can be given as a probability space.

Typically, a measurable space comprises of the first two components without a specific measure.

Definition 1: A measure space is defined as a triple $(X, \mathcal{A}, \mu)$.

Where, X is a Set.

$\mathcal{A}$ is a σ-Algebra on the Set X

μ is a Measure on $(X, \mathcal{A})$

Significant Key Classes of Measure Spaces

Extremely significant key classes of measure spaces are specifically defined by means of the following properties of their associated and corelated measures:

- Probability spaces, a measure space where the measure is a probability measure.
- Finite measure spaces, where the measure is a finite measure.
- The σ-finite measure spaces, where the measure is a σ-finite measure.

Additional class of measure spaces are defined as the complete measure spaces.

A measure space is characteristically defined as a measurable space that possesses a non-negative measure. The typical examples of measure spaces include n-dimensional Euclidean space with Lebesgue measure and the unit interval with Lebesgue measure, i.e., probability.

The Lebesgue integral depends ultimately on the idea of measure. In particular, the mathematical framework requires a set, a σ-algebra of subsets alongwith a set function that assigns a non-negative number (called its measure) to each set in the σ-algebra.

Definition 2: Suppose Ω is a set and $\mathcal{A}$ a σ-algebra of subsets of Ω. A measure, μ on $\mathcal{A}$ is a set function having domain $\mathcal{A}$ satisfying the following:

NOTES

(a) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$.

(b) $\mu(\phi) = 0$.

(c) If $A_1, A_2, \ldots$ are in $\mathcal{A}$, with $A_i \cap A_j = \phi$ for $i \neq j$, then

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$$

Here, the pair (Ω, A) is termed as a measurable space and the triple (Ω, A, μ) or $(X, \mathcal{A}, \mu)$ is called the measure space.

5.7 INTEGRATION WITH RESPECT TO A MEASURE

For explaining integration with respect to a measure, we first define the integral of a non-negative function with respect to a measure. Then we write a real valued function particularly as the difference of two non-negative functions for defining the integral of a real valued function with respect to a measure.

Consider the following definition to explain the integration with respect to a measure.

Definition $\mathcal{S}$-Partition: Assume that $\mathcal{S}$ is a σ-algebra on a set X. Characteristically, an $\mathcal{S}$-partition of X is defined as a finite collection $A_1, \ldots, A_m$ of disjoint sets in S such that $A_1 \cup \cdots \cup A_m = X$.

Implementing the convention that $0 \cdot \infty$ and $\infty \cdot 0$ should both be interpreted to be 0.

Further, consider an arbitrary measure and therefore X must not be a subset of **R**. More significantly, for the condition when X is a closed interval $[a, b]$ in **R** and μ is Lebesgue measure on the Borel subsets of $[a, b]$, then the sets $A_1, \ldots, A_m$ do not need to be subintervals of $[a, b]$ as they are in the lower Riemann sum, they should only be Borel sets.

In mathematics, specifically in the real analysis, the integral of a non-negative function of a single variable can be simply interpreted as the area between the graph of that function and the X-axis. The Lebesgue integral extends the integral to a larger class of functions. It also extends the domains on which these functions can be defined. The Lebesgue integral is named after Henri Lebesgue (1875–1941), who introduced the Lebesgue integral in the year 1904. It is also a pivotal part of the axiomatic theory of probability.

The integral of a positive function f between limits a and b can be interpreted as the area under the graph of f. However, Riemann integration does not interact accurately by taking limits of sequences of functions, because producing such limiting processes are difficult for analyses. The Lebesgue integral perfectly explains how and when it is possible to take limits under the integral sign through the monotone convergence theorem and dominated convergence theorem.

Consider a measure space (E, X, μ) where E is a set, X is a σ-algebra of subsets of E, and μ is a non-negative measure on E defined on the sets of X.

NOTES

For example, E can be Euclidean n-space $\mathbb{R}^n$ or some Lebesgue measurable subset of it, X is the σ-algebra of all Lebesgue measurable subsets of E, and μ is the Lebesgue measure. As per the mathematical theory of probability, a probability measure μ, can be considered which satisfies $\mu(E) = 1$.

Lebesgue's theory defines integrals for a class of functions called measurable functions. A real valued function f on E is measurable if the preimage of every interval of the form (t, ∞), i.e., any Borel set is in X:

$$\{x \mid f(x) > t\} \in X \quad \forall t \in \mathbb{R}.$$

We can explain this to state that this can be quivalent to the preimage of any Borel subset of $\mathbb{R}$ be in X. The set of measurable functions is closed under algebraic operations, but more significantly it is closed under various kinds of pointwise sequential limits:

$$\sup_{k\in\mathbb{N}} f_k, \quad \liminf_{k\in\mathbb{N}} f_k, \quad \limsup_{k\in\mathbb{N}} f_k$$

These are measurable if the original sequence $(f_k)_k$, where $k \in \mathbb{N}$, consists of measurable functions.

There are several approaches for defining an integral:

$$\int_E f\, d\mu = \int_E f(x)\ d\mu(x)$$

This is for measurable real valued functions f defined on E.

Consider another important case of the measure space $(\mathbf{R}, B, \lambda)$, where λ is the Lebesgue measure, or its subsets $([a, b], B, \lambda)$. The following are basic examples of λ integrable functions and of functions which are not.

Example 5.2. To prove that a given measurable function f is integrable with respect to the Lebesgue measure on a subset $X \subset \mathbf{R}$ or for a more general measure space, then the most common technique is to find a "Simple" comparison function g which is known to be integrable and for which it is known that,

$$|f(x)| \leqslant g(x), \qquad x \in X.$$

Generally, more than one comparison function can be used, for instance we can find the disjoint subsets X_1, X_2 such that $X = X_1 \cup X_2$, and functions g_1, g_2 are integrable on X_1 and X_2, respectively, therefore,

$$|f(x)| \leqslant \begin{cases} g_1(x) & \text{if } x \in X_1, \\ g_2(x) & \text{if } x \in X_2. \end{cases}$$

This can be applied with infinitely many subsets. For instance, consider,

$$X = [1, +\infty[, \qquad f(x) = x^{-\nu}$$

Where $v > 0$. Then f is λ-integrable on X if and only if $v > 1$. Actually note that,

$$0 \leqslant f(x) \leqslant n^{-\nu}, \quad x \in [n, n+1[, \quad n \geqslant 1,$$

Consequently, using the monotone convergence theorem, we have

$$\int_X x^{-\nu} d\lambda(x) \leqslant \sum_{n \geqslant 1} n^{-\nu} < +\infty$$

NOTES

Characteristically, a measurable function which is bounded by an integrable function is integrable. Every integrable function is measurable. If a sequence of measurable functions converges almost everywhere, then its limit is measurable. If a sequence of measurable functions converges asymptotically, then its limit is measurable. Basically, the set of measurable functions is defined as a linear space. Additionally, the intersection and union of two measurable functions are measurable.

Integration with respect to a measure is termed as Lebesgue integration. The following definition illustrates that the Lebesgue integration functions as anticipated on the simple functions represented as linear combinations of characteristic functions of disjoint sets.

Integral of a Simple Function

Definition 2: Assume that (X, δ, μ) is a measure space, $E_1, \ldots, E_n$ are disjoint sets in S, and $c_1, \ldots, c_n \in [0, \infty]$. Then,

$$\int \Big(\sum_{k=1}^{n} c_k \chi_{E_k}\Big) \, d\mu = \sum_{k=1}^{n} c_k \mu(E_k).$$

Without loss of simplification, we can assume that $E_1, \ldots, E_n$ is an δ-partition of X by replacing n by $n + 1$ and setting $E_{n+1} = X \setminus (E_1 \cup \ldots \cup E_n)$ and $c_{n+1} = 0$. If P is the δ-partition $E_1, \ldots, E_n$ of X, then,

$$\mathcal{L}\big(\textstyle\sum_{k=1}^{n} c_k \chi_{E_k}, P\big) = \textstyle\sum_{k=1}^{n} c_k \mu(E_k)$$

Consequently,

$$\int \big(\sum_{k=1}^{n} c_k \chi_{E_k}\big) \, d\mu \geq \sum_{k=1}^{n} c_k \mu(E_k).$$

Integration is Order Preserving

Definition 3: Assume that (X, δ, μ) is a measure space and $f, g : X \to [0, \infty]$ are δ-measurable functions such that $f(x) \leq g(x)$ for all $x \in X$. Then $\int f \, d\mu \leq \int g \, d\mu$.

Suppose P is an δ-partition $A_1, \ldots, A_m$ of X.

Then,

$$\inf_{A_j} f \leq \inf_{A_j} g$$

Where for each $j = 1, \ldots, m$. Accordingly, $\mathcal{L}(f, P) \leq \mathcal{L}(g, P)$.

Hence, $\int f \, d\mu \leq \int g \, d\mu$.

NOTES

5.8 THE L^p-SPACES

Characteristically, in mathematics, the L^p spaces are function spaces defined using a natural generalization of the p-norm for Finite Dimensional Vector Spaces (FDVS). They are also sometimes called Lebesgue spaces, named after Henri Lebesgue, although according to the Bourbaki group (Bourbaki 1987) they were first introduced by Frigyes Riesz (Riesz 1910). L^p spaces form an significant class of Banach spaces in functional analysis, and of topological vector spaces.

The set of L^p-functions where $p \geq 1$ generalizes L^2-space. As an alternative of square integrable, the measurable function f must be p-integrable for f to be in L^p.

On a measure space X, the L^p norm of a function f is,

$$|f|_{L^p} = \left(\int_X |f|^p\right)^{1/p}$$

The L^p-functions are the characteristic functions for which this integral converges. For $p \neq 2$, the space of L^p-functions is a Banach space which is not a Hilbert space.

Suppose, p is a positive real number. Then a measurable function f defined on [0, 1] is said to belong to the space L^p if $\int |f|^p < \infty$. Hence, L^1 precisely consists of Lebesgue integrable functions on [0, 1]. Consequently,

$$|f+g|^P \leq 2^P(|f|^P + |g|^P),$$

We have,

$$\int |f+g|^P \leq 2^P \int (|f|^P + 2^P |f|^P)$$

And therefore, if $f, g \in L^p$, then $f + g \in L^p$. Additionally, if α is a scalar and $f \in L^p$, then clearly αf belongs to L^p. Hence, $\alpha f + \beta g \in L^p$ whenever $f, g \in L^p$ and α, β are scalars.

5.9 CONVEX FUNCTIONS

Characteristically, a real valued function is called **convex** if the line segment between any two points on the graph of the function does not lie below the graph between the two points. Equivalently, a function is convex if its epigraph, i.e., the set of points on or above the graph of the function is a convex set. A twice-differentiable function of a single variable is convex if and only if its second derivative is non-negative on its entire domain. Recognised examples of convex functions of a single variable include the quadratic function x^2 and the exponential function e^x. In simple terms, a convex function refers to a function whose graph is shaped like a cup $\cup$, while a concave function's graph is shaped like a cap $\cap$.

Convex functions play a significant role in several areas of mathematics. Even in infinite dimensional spaces, under suitable additional hypotheses, the convex functions continue to satisfy such properties and as a result, they are the most significant functionals in the calculus of variations. In probability theory, a convex function applied to the expected value of a random variable is always bounded above by the expected value of the convex function of the random variable. This result, known as Jensen's inequality, can be used to deduce inequalities such as the arithmetic–geometric mean inequality and Hölder's inequality.

Definition 1: A function ϕ defined an open interval (a, b) is known as a **convex function** if for each $x, y \in (a, b)$ and λ, μ such that $\lambda, \mu \geq 0$ and $\lambda + \mu = 1$, we have,

$$\phi\,(\lambda x + \mu y) \leq \lambda\phi(x) + \mu\phi(y)$$

The end points a, b can take the values $-\infty, \infty$, respectively.

If we take $\mu = 1 - \lambda$, $\lambda \geq 0$, then $\lambda + \mu = 1$ and so ϕ will be convex if,

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y) \qquad \text{... (5.3)}$$

If we take $a < s < t < u < b$ and

$$\lambda = \frac{t-s}{u-s},\ \mu = \frac{u-t}{u-s},\ u = x,\ s = y,$$

Then,

$$\lambda + \mu = \frac{t-s+u-t}{u-s} = \frac{u-s}{u-s} = 1$$

Therefore, Equation (5.3) reduces to the form,

$$\phi\left(\frac{t-s}{u-s}u + \frac{u-t}{u-s}s\right) \leq \frac{t-s}{u-s}\varphi(u) + \frac{u-t}{u-s}\varphi(s)$$

Or,

$$\phi(\text{t}) \leq \frac{t-s}{u-s}\varphi(u) + \frac{u-t}{u-s}\varphi(u) \qquad \text{... (5.4)}$$

Thus, the segment joining $(s, \phi(s))$ and $(u, \phi(n))$ is never below the graph of ϕ. A function ϕ is sometimes said to be convex on (a, b) if for all $x, y \in (a, b)$.

$$f\left(\frac{x+y}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

Remember that, this definition is a consequence of major definition taking, $\lambda = \mu = 1/2$.

If for all positive numbers λ, μ satisfying $\lambda + \mu = 1$, then we have

$$\phi(\lambda x + \mu y) \leq \lambda\phi(x) + \mu\phi(y)$$

Then, ϕ is said to be strictly convex.

NOTES

Theorem 5.4: Let ϕ be convex on (a, b) and $a < s < t < u < b$, then

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(u)-\phi(s)}{u-s} \le \frac{\phi(u-\phi(t)}{u-t}$$

If ϕ is strictly convex, equality will not occur.

Proof: Let $a < s < t < u < b$ and suppose ϕ is convex on (a, b).

Since,

$$\frac{t-s}{u-s} + \frac{u-t}{u-s} = \frac{t-s+u-t}{u-s} = \frac{u-s}{u-s} = 1$$

Then, from the convexity of ϕ,

$$\phi\left(\frac{t-s}{u-s}u + \frac{u-t}{u-s}s\right) \le \frac{t-s}{u-s}\phi(u) + \frac{u-t}{u-s}\phi(s)$$

$$\Rightarrow \quad \phi(t) \le \frac{t-s}{u-s}\phi(u) + \frac{u-t}{u-s}\phi(s) \qquad \text{.... (5.5)}$$

$$\Rightarrow \quad (u-s)\phi(t) \le (t-s)\phi(u) + (u-t)\phi(s)$$

$$\Rightarrow \quad (u-s)(\phi(t)-\phi(s) \le (t-s)\phi(u) + u\phi(s) - t\phi(s) - u\phi(s) + s\phi(s)$$

$$\Rightarrow \quad (u-s)(\phi(t)-\phi(s)) \le (t-s)(\phi(u)-\phi(s))$$

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(u)-\phi(s)}{u-s} \qquad \text{... (5.6)}$$

Hence, the first inequality is proved. In the same way, the second inequality can be proved. If ϕ is strictly converse, equality shall not be there in Equation (5.5) and so it cannot be in Equation (5.6). This completes the proof.

Theorem 5.5: A differentiable function ϕ is convex on (a, b) iff ϕ' is a monotonically increasing function. If ϕ'exists on (a, b), then ϕ is convex iff $\phi'' \ge 0$ on (a, b) and strictly convex if $\psi'' > 0$ on (a, b).

Proof: Consider a differentiable and convex function ϕ and also consider that $a < s < t < u < v < b$. Then applying Theorem 5.4 to $a < s < t < u$, we obtain

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(u)-\phi(s)}{u-s} \le \frac{\phi(u)-\phi(t)}{u-t}$$

And applying Theorem 5.4 to $a < t < u < v$, we obtain

$$\frac{\phi(u)-\phi(t)}{u-t} \le \frac{\phi(v)-\phi(t)}{v-t} \le \frac{\phi(v)-\phi(v)}{v-u}$$

Hence,

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(v)-\phi(u)}{v-u}$$

For $t \to s$, $\dfrac{\phi(t)-\phi(s)}{t-s}$ decreases to $\phi'(s)$ and for $u \to v$, $\dfrac{\phi(v)-\phi(u)}{v-u}$

increases to $\phi'(v)$. Hence, $\phi'(v) \geq \phi'(s)$ for all $s < v$ and so ϕ' is monotonically increasing function. Further, if ϕ'' exists, due to monotonicity of ϕ', it can never be negative.

Conversely, let $\psi'' \geq 0$. We shall now prove that ψ is convex. Suppose, on the contrary that ϕ is not convex on (a, b). Therefore, there are points $a < s < t < u < b$, such that,

$$\frac{\phi(t)-\phi(s)}{t-s} > \frac{\phi(u)-\phi(t)}{u-t}$$

This means that the slope of chord over (s, t) is larger than the slope of the chord over (t, u). But slope of the chord over (s, t) is equal to $\phi'(\alpha)$, for some $\alpha \in (s, t)$ and slope of the chord over (t, u) is $\phi'(\beta)$, $\beta \in (t, u)$.

But $\phi'(\alpha) > \phi'(\beta)$ implies ϕ' is not monotone increasing and so ψ'' cannot be greater than zero which is a contradiction. Hence, ϕ is convex.

If $\phi'' > 0$, then ϕ is strictly convex, for otherwise there would exist collinear points of the graph of ϕ and we would have $\phi'(\alpha) = \phi'(\beta)$ for appropriate α and β with $\alpha < \beta$. But then $\phi'' = 0$ at some point between α and β which is a contradiction to $\phi'' \geq 0$.

This completes the proof.

Theorem 5.6: If ϕ is convex on (a, b), then ϕ is absolutely continuous on each closed subinterval of (a, b).

Proof: Assume that $[a, b] \subset (a, b)$. If $x, y \in [c, d]$, then we have $a < c \leq x \leq y \leq d < b$ and so by Theorem 5.4, we have

$$\frac{\phi(c)-\phi(a)}{c-a} \leq \frac{\phi(y)-\phi(x)}{y-x} \leq \frac{\phi(b)-\phi(d)}{b-d}$$

Thus,

$$|\phi(y)-\phi(x)| \leq M\,|x-y|,\ x, y \in [c, d]$$

and so ϕ is absolutely continuous there.

Theorem 5.7: Every convex function on an open interval is continuous.

Proof: When $a < x_1 < x < x_2 < b$, the convexity of a function ϕ implies,

$$\phi(x) \leq \frac{x_2 - x}{x_2 - x_1}\phi(x_1) + \frac{x - x_1}{x_2 - x_1}\phi(x_2) \qquad \dots (5.7)$$

For $x \to x_1$ in Equation (5.7), we obtain $\phi(x_1 + 0) \leq \phi(x_1)$ and for $x_2 \to x$ we obtain $\phi(x) \leq \phi(x + 0)$.

Hence, $\phi(x) = \phi(x + 0)$ for all values of x in (a, b).

Similarly, $\phi(x - 0) = \phi(x)$ for all values of x. Therefore, $\phi(x - 0) = \phi(x + 0) = \phi(x)$ and so ϕ is continuous.

Definition 2: Let ϕ be a convex function on (a, b) and $x_0 \in (a, b)$. The line,

$$y = m(x - x_0) + \phi(x_0) \qquad \dots (5.8)$$

NOTES

Typically, $(x_0, \phi(x_0))$ is called a supporting line at x_0 if it always lies below the graph of ϕ, i.e., when

$$\phi(x) \geq m(x - x_0) + \phi(x_0) \qquad ...(5.9)$$

The line given by Equation (5.8) is a supporting line iff its slope m lies between the left and the right hand derivative at x_0. Therefore, precisely, there is at least one supporting line at each point.

5.10 JENSEN'S INEQUALITY

In mathematical analysis, the term Jensen's inequality is named after the Danish mathematician Johan Jensen, it was proved by Jensen in 1906. The Jensen's inequality relates the value of a convex function of an integral to the integral of the convex function. Given its generality, the inequality appears in many forms depending on the context. In its simplest form the inequality states that the convex transformation of a mean is less than or equal to the mean applied after convex transformation; it is a simple corollary that the opposite is true of concave transformations.

Jensen's inequality generalizes the statement that the secant line of a convex function remains above the graph of the function, which is Jensen's inequality for two points: the secant line consists of weighted means of the convex function for $t \in [0,1]$,

$$t\,f(x_1) + (1 - t)\,f(x_2)$$

while the graph of the function is the convex function of the weighted means,

$$f(t\,x_1 + (1 - t)\,x_2)$$

Consequently, the Jensen's inequality is,

$$f(t\,x_1 + (1 - t)\,x_2) \leq t\,f(x_1) + (1 - t)\,f(x_2)$$

The classical form of Jensen's inequality includes several numbers and weights. The inequality can be stated quite commonly using either the language of measure theory or equivalently the probability.

Consider $EX = \sum_{x \in X} xp(x)$, where E denotes expectation.

Theorem 5.8. (Jensen's Inequality): If f is a convex function and X is a random variable, then $Ef(X) \geq f(EX)$. Furthermore, if f is strictly convex, then equality implies that $X = EX$ with probability 1, that is, X is constant.

Proof: To prove this Theorem 5.8, we will apply induction on the number of mass points. From the definition of convex functions, for two points we have,

$$P_1 f(x_1) + P_2 f(x_2) \geq f(P_1 x_1 + P_2 x_2)$$

Suppose the theorem holds for $k - 1$ mass points. For $1 \leq i \leq k - 1$, then we can write,

$$p_i' = p_i/(1 - p_k)$$

So, we obtain

NOTES

$$\sum_{i=1}^{k} p_i f(x_i) = p_k f(x_k) + (1-p_k)\sum_{i=1}^{k-1} p_i' f(x_i)$$

$$\geq p_k f(x_k) + (1-p_k) f\left(\sum_{i=1}^{k-1} p_i' x_i\right) \qquad \text{(From Induction Hypothesis)}$$

$$\geq f\left(p_k x_k + (1-p_k)\sum_{i=1}^{k-1} p_i' x_i\right) \qquad \text{(From the Definition of Convexity)}$$

Hence, the theorem is proved.

Check Your Progress

1. Give the definition of outer measure .
2. How is the measure on an algebra defined?
3. State the Carathéodory's extension theorem as per the measure theory.
4. Define the unique extension theorem.
5. What is complete measure?
6. How is the complete measure constructed?
7. What is measure space? Give the definition of the term measure space.
8. State the definition of δ-partition.
9. Define the term L^p spaces.
10. Give the definition of convex function.
11. State the Jensen's inequality theorem.

5.11 HÖLDER AND MINKOWSKI INEQUALITIES

The Hölder's and Minkowski's inequalities are defined below for analysis.

Hölder's Inequality

In real analysis, Hölder's inequality, named after Otto Hölder, is a fundamental inequality between integrals and an indispensable tool for the study of L^p spaces.

Theorem 5.9: Hölder's Inequality: Let (S, Σ, μ) be a measure space and let p, $q \in [1, \infty]$ with $1/p + 1/q = 1$. Then for all measurable real valued function or complex valued function f and g on S,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

If, in addition, $p, q \in (1, \infty)$ and $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then Hölder's inequality becomes an equality if and only if $|f|^p$ and $|g|^q$ are linearly dependent in $L^1(\mu)$, meaning that there exist real numbers $\alpha, \beta \geq 0$, not both of them zero, such that $\alpha|f|^p = \beta\,|g|^q$ μ-almost everywhere.

NOTES

The numbers p and q above are said to be Hölder conjugates of each other. The special case $p = q = 2$ gives a form of the Cauchy–Schwarz inequality. Hölder's inequality holds even if $\|fg\|_1$ is infinite, the right-hand side also being infinite in that case. Conversely, if f is in $L^p(\mu)$ and g is in $L^q(\mu)$, then the pointwise product fg is in $L^1(\mu)$.

Hölder's inequality is used to prove the Minkowski inequality, which is the triangle inequality in the space $L^p(\mu)$, and also to establish that $L^q(\mu)$ is the dual space of $L^p(\mu)$ for $p \in [1, \infty)$.

If S is a measurable subset of $\boldsymbol{R}^n$ with the Lebesgue measure, and f and g are measurable real or complex valued function on S, then Hölder inequality is

$$\int_S |f(x)g(x)|\,\mathrm{d}x \le \left(\int_S |f(x)|^p\,\mathrm{d}x\right)^{\frac{1}{p}} \left(\int_S |g(x)|^q\,\mathrm{d}x\right)^{\frac{1}{q}}.$$

Hölder's inequality was first found by Leonard James Rogers (Rogers (1888)), and discovered independently by Hölder (1889).

Minkowski Inequality

In real analysis, the Minkowski inequality establishes that the L^p spaces are normed vector spaces. The inequality is named after the German mathematician Hermann Minkowski.

Theorem 5.10: Minkowski Inequality: Let S be a measure space, let $1 \le p < \infty$ and let f and g be elements of $L^p(S)$. Then $f + g$ is in $L^p(S)$, and we have the triangle inequality,

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

With equality for $1 < p < \infty$ if and only if f and g are positively linearly dependent, i.e., $f = \lambda_g$ for some $\lambda \ge 0$ or $g = 0$. Here, the norm is given by:

$$\|f\|_p = \left(\int |f|^p d\mu\right)^{\frac{1}{p}}$$

If $p < \infty$, or in the case $p = \infty$ by the essential supremum,

$$\|f\|_\infty = \operatorname{ess\,sup}_{x \in S} |f(x)|.$$

The Minkowski inequality is the triangle inequality in $L^p(S)$. In fact, it is a special case of the more general fact,

$$\|f\|_p = \sup_{\|g\|_q = 1} \int |fg|\,d\mu, \qquad \tfrac{1}{p} + \tfrac{1}{q} = 1$$

It can be easily state that the right hand side satisfies the triangular inequality.

Like Hölder's inequality, the Minkowski inequality can be specialized to sequences and vectors by using the counting measure:

$$\left(\sum_{k=1}^{n}|x_k+y_k|^p\right)^{1/p} \le \left(\sum_{k=1}^{n}|x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n}|y_k|^p\right)^{1/p}$$

for all real (or complex) numbers $x_1, ..., x_n, y_1, ..., y_n$ and where n is the cardinality of S (the number of elements in S).

The Minkowski inequality can be generalized to other functions $\phi(x)$ beyond the power function x^p. The generalized inequality has the form,

$$\phi^{-1}(\sum_{i=1}^{n}\phi(x_i+y_i)) \le \phi^{-1}(\sum_{i=1}^{n}\phi(x_i)) + \phi^{-1}(\sum_{i=1}^{n}\phi(y_i))$$

Various sufficient conditions on ϕ have been found by Mulholland and others. For example, for $x \ge 0$ one set of sufficient conditions from Mulholland is,

1. $\phi(x)$ is continuous and strictly increasing with $\phi(0) = 0$.
2. $\phi(x)$ is a convex function of x.
3. log $\phi(x)$ is a convex function of log (x).

5.12 COMPLETENESS OF L^P

Characteristically, the L^p functions have certainly ambiguous pointwise values. However, we usually consider L^p functions as complete functions. A simple example of this construction, for a measure that has no sets of measure 0, consequently requires no quotient is given by,

$$\ell^p = \{\text{complex sequences } \{c_i\} \text{ with } \sum_i |c_i|^p < \infty\}$$

With standard norm,

$$|(c_1, c_2, \ldots)|_{\ell^p} = (\textstyle\sum_i |c_i|^p)^{1/p}.$$

The analogue of the following statement for ℓ^p is more elementary.

Statement: The Space L^p (X) is a Complete Metric Space

Essentially, to prove a Cauchy sequence f_i in L^p (X) has a subsequence which converges pointwise off a set of measure 0 in X.

The vector space of equivalence classes of measurable functions on (S, Σ, μ) is denoted ad $L^0(S, \Sigma, \mu)$. By definition, it contains all the L^p, and is equipped with the topology of convergence in measure. When μ is a probability measure (i.e., μ(S) = 1), this mode of convergence is named convergence in probability.

The description is easier when μ is finite. If μ is a finite measure on (S, Σ), the 0 function admits for the convergence in measure the following fundamental system of neighbourhoods,

$$V_\varepsilon = \Big\{f : \mu\big(\{x : |f(x)| > \varepsilon\}\big) < \varepsilon\Big\}, \qquad \varepsilon > 0$$

NOTES

The topology can be defined by any metric d of the form,

$$d(f,g) = \int_S \varphi\big(|f(x) - g(x)|\big)\, \mathrm{d}\mu(x)$$

Where φ is bounded continuous concave and non-decreasing on $[0, \infty)$, with $\varphi(0) = 0$ and $\varphi(t) > 0$ when $t > 0$ (for example, $\varphi(t) = \min(t, 1)$). Such a metric is called Lévy metric for L^0. Under this metric the space L^0 is complete (it is again an F-space). The space L^0 is in general not locally bounded and not locally convex.

For the infinite Lebesgue measure λ on $\boldsymbol{R}^n$, the definition of the fundamental system of neighbourhoods could be modified as follows:

$$W_\varepsilon = \left\{ f : \lambda\left(\left\{ x : |f(x)| > \varepsilon \text{ and } |x| < \frac{1}{\varepsilon} \right\}\right) < \varepsilon \right\}$$

The resulting space $L^0(\boldsymbol{R}^n, \lambda)$ coincides as topological vector space with $L^0(\boldsymbol{R}^n, g(x)\, d\lambda(x))$, for any positive λ–integrable density g.

Theorem 5.11: The space $L^p(X)$ is a complete metric space.

Proof: The triangle inequality here is Minkowski's inequality. To prove completeness, choose a subsequence f_{ni}, such that,

$$|f_{n_i} - f_{n_{i+1}}|_p \; < \; 2^{-i}$$

And, put

$$g_n(x) \;=\; \sum_{1 \le i \le n} |f_{n_{i+1}}(x) - f_{n_i}(x)|$$

And,

$$g(x) \;=\; \sum_{1 \le i < \infty} |f_{n_{i+1}}(x) - f_{n_i}(x)|$$

The infinite sum is not necessarily claimed to converge to a finite value for every x. The triangle inequality shows that $|g_n|_p \le 1$. Fatou's Lemma asserts that for $[0, \infty]$-valued measurable functions h_i

$$\int_X \left(\liminf_i h_i\right) \;\le\; \liminf_i \int_X h_i$$

Thus, $|g|_p \le 1$, so is finite. Consequently,

$$f_{n_1}(x) + \sum_{i \ge 1} \left(f_{n_{i+1}}(x) - f_{n_i}(x)\right)$$

Converges for almost all $x \in X$. Let $f(x)$ be the sum at points x where the series converges, and on the measure zero set where the series does not converge put $f(x) = 0$. Certainly,

NOTES

$$f(x) = \lim_i f_{n_i}(x) \quad \text{(for almost all } x\text{)}$$

Now, prove that this almost everywhere pointwise limit is the L^p-limit of the original sequence. For $\varepsilon > 0$ take N such that $|f_m - f_n|_p < \varepsilon$ for $m, n \geq N$. Fatou's lemma gives,

$$\int |f - f_n|^p \leq \liminf_i \int |f_{n_i} - f_n|^p \leq \varepsilon^p$$

Thus $f - f_n$ is in L^p and hence f is in L^p. And $|f - f_n|_p \to 0$.

Theorem 5.12: L^p is complete, i.e., every Cauchy sequence converge.

Proof: If, $\sum_k \|f_k\|_p < \infty$, then $\sum_k f_k$ converges in L^p norm to an element in L^p.

$$G_n(x) = \sum_{j=1}^{n} |f_j(x)|,$$

$$G(x) = \sum_{j=1}^{\infty} |f_j(x)|$$

And observe that $\|G_n\|_p \leq \sum \|f_j\|_p < \infty$ by assumption. Monotone convergence: $G \in L^p$. In particular $G(x) < \infty$, i.e., $\sum f_k$ converges at least pointwise.

We have, $|\overline{F}| \leq G$, hence, $F \in L^p$.

Moreover,

$$|F(x) - \sum_{j=1}^{n} f_j(x)|^p \leq (2G(x))^p \in L^1.$$

This specifies that $\|F - \sum^n f_j\|_p \to 0$, i.e., $\sum f_j$ converges to F in L^p.

Now, if F_n is a Cauchy sequence in L^p, consider a sequence n_k so that $\|F_n - F_m\| < 2^{-j}$ for $m, n \geq n_j$. Set $f_1 = F_{n_1}$ and $f_j = F_{n_j} - F_{n_{j-1}}$ for $j > 1$.

NOTES

Again,

$$F_{n_k} = \sum_{n=1}^{k} f_n.$$

To prove that the series converges in L^p consider that F_n is Cauchy and then explain that F_n and F_{nk} have the same limit.

Consequence: All L^p spaces are normed complete vector spaces. These are also called Banach spaces.

5.13 CONVERGENCE IN MEASURE

Convergence in measure is either of following two distinct mathematical concepts both of which generalize the concept of convergence in probability.

Let $\boldsymbol{f, f_n}$ $(\boldsymbol{n} \in \mathbb{N}) : \boldsymbol{X} \to \mathbb{R}$ be the measurable functions defined on a measure space $(\boldsymbol{X}, \Sigma, \mu)$. The sequence $\boldsymbol{f_n}$ is said to converge globally in measure to $\boldsymbol{f}$ if for every $\varepsilon > \mathbf{0}$,

$$\lim_{n\to\infty} \mu(\{x \in X : |f(x) - f_n(x)| \geq \varepsilon\}) = 0.$$

And to converge locally in measure to $\boldsymbol{f}$ if for every $\varepsilon > \mathbf{0}$ and every $\boldsymbol{F} \in \Sigma$ with $\mu\,(\boldsymbol{F}) < \infty$,

$$\lim_{n\to\infty} \mu(\{x \in F : |f(x) - f_n(x)| \geq \varepsilon\}) = 0.$$

On a finite measure space, both notions are equivalent. Otherwise, convergence in measure can refer to either global convergence in measure or local convergence in measure.

Definition 1: A sequence $<f_n>$ of measurable functions is said to converge to f in measure if for given $\varepsilon > 0$, there is an N such that for all $n \geq N$ we have,

$$m\{x \mid f(x) - f_n(x) \mid \geq \varepsilon\} < \varepsilon.$$

Theorem 4.13 (F. Riesz): Let $<f_n>$ be a sequence of measurable functions that converges in measure to f. Then there is a subsequence $<f_{nk}>$ which converges to f almost everywhere.

Proof: Since $<f_n>$ is a sequence of measurable functions which converges in measure to f, for any positive integer k there is an integer n_k such that for $n \geq n_k$, we have,

$$m\{x \mid f_n(x) - f(x) \mid \geq \frac{1}{2^k}\} < \frac{1}{2^k}$$

Let, $E_k = \{x \mid\mid f_{n_k}(x) - f(x) \mid \geq \dfrac{1}{2^k}\}$

NOTES

Then if $x \notin \bigcup_{k=i}^{\infty} E_k$, we have

$$| f_{n_k}(x) - f(x) | \frac{1}{2^k} \text{ for } k \geq i$$

And so $f_{n_k}(x) \to f(x)$.

Hence, $f_{n_k}(x) \to f(x)$ for any $x \notin A = \bigcap_{i=1}^{\infty} \bigcup_{k=i}^{\infty} E_k$

But,

$$mA \leq m\left[\bigcup_{k=i}^{\infty} E_k\right]$$

$$= \sum_{k=i}^{\infty} mE_k = \frac{1}{2^{k-1}}$$

Hence, the measure of A is zero.

Example 5.3: A sequence $<f_n>$ which converges to zero in measure on [0,1] but such that $<f_n(x)>$ does not converge for any x in [0,1] can be constructed as follows:

Let $n = k + 2^v$, $0 \leq k \leq 2^v$, and set $f_n(x) = 1$ if $x \in [k2^{-v}, (k+1)2^{-v}]$ and $f_n(x) = 0$ otherwise. Then, $m\{x \,||\, f_n(x)| > \varepsilon\} \leq \frac{2}{n}$ and so, $f_n \to 0$ in measure, although for any $x \in [0, 1]$, the sequence $<f_n(x)>$ has the value1 for arbitrarily large values of n. So it does not converge.

Definition 2: A sequence $\{f_n\}$ of almost everywhere finite valued measurable functions is said to be fundamental in measure, if for every $\varepsilon > 0$,

$$m(\{x : | f_n(x) - f_m(x) | \geq \varepsilon\}) \to 0 \text{ as } n \text{ and } m \to \infty.$$

Definition 3: A sequence $\{f_n\}$ of real valued functions is called fundamental almost everywhere if there exists a set E_0 of measure zero such that, if $x \notin E_0$ and $\varepsilon > 0$, then an integer $n_0 = n_0 = (x, \varepsilon)$ has the property that,

$$| f_n(x) - f_m(x) | < \varepsilon, \text{ whenever } n \geq n_0 \text{ and } m \geq n_0.$$

Definition 4: A sequence $\{f_n\}$ of almost everywhere finite valued measurable functions is said to converge to the measurable function f almost uniformly if, for every $\varepsilon > 0$, there exists a measurable set F such that $m(F) < \varepsilon$ and also the sequence $\{f_n\}$ converges to f uniformly on F^c.

Theorem 5.14: If $\{f_n\}$ is a sequence of measurable functions which converges to f almost uniformly, then $\{f_n\}$ converges to f almost everywhere.

Proof: Let F_n be a measurable set such that $m(F_n) < 1/n$ and such that the sequence $\{f_n\}$ converges to f uniformly on $F_n^{\,c}$, $n = 1, 2, \ldots$. If $F = \bigcap_{n=1}^{\infty} F_n$, then $m(F) \leq \mu(F_n) < \frac{1}{n}$ so that $m(F) = 0$, and it is clear that, for $x \in F^c$, $\{f_n(x)\}$ converges to $f(x)$.

NOTES

5.14 ALMOST UNIFORM CONVERGENCE

In the mathematical analysis, the term uniform convergence is a specific mode of convergence of functions which is stronger than pointwise convergence.

If the domain of the functions is a measure space E then the related notion of '**almost uniform convergence'** can be defined. We define that a sequence of functions (f_n) **converges almost uniformly** on E if for every $\delta > 0$ there exists a measurable set E_δ with measure less than δ such that the sequence of functions (f_n) converges uniformly on $E \setminus E_\delta$. In other words, almost uniform convergence means there are sets of arbitrarily small measure for which the sequence of functions converges uniformly on their complement.

The almost uniform convergence of a sequence does not mean that the sequence converges uniformly almost everywhere as might be inferred from the name. Almost uniform convergence implies almost everywhere convergence and convergence in measure.

Theorem 5.15: Almost uniform convergence implies convergence in measure.

Proof: If $\{f_n\}$ converges to f almost uniformly, then for any two positive numbers ε and δ there exists a measurable set F such that $m(F) < \delta$ such that $|f_n(x) - f(x)| < \varepsilon$, whenever x belongs to F^c and n is sufficiently large.

Theorem 5.16: If $\{f_n\}$ converges in measure to f, then $\{f_n\}$ is fundamental in measure. Also, if $\{f_n\}$ converges in measure to g, then $f = g$ almost everywhere.

Proof: The first claim of the Theorem 5.16 follows from the following relation,

$$\{x : |f_n(x) - f_m(x)| \geq \varepsilon\} \subset \{x : |f_n(x) - f(x)| \geq \frac{\varepsilon}{2}\} \cup \{x : |f_m(x) - f(x)| \geq \frac{\varepsilon}{2}\}$$

For proving the second claim, we have,

$$\{x : |f(x) - g(x)| \geq \varepsilon\} \subset \{x : f_n(x) - f(x)| \geq \frac{\varepsilon}{2}\} \cup \{x : |f_n(x) - g(x)| \geq \frac{\varepsilon}{2}\}$$

Since by appropriate selection of n, the measure of both sets on the right can be made arbitrarily small, we have

$$m(\{x : |f(x) - g(x)| \geq \varepsilon\}) = 0$$

for every $\varepsilon > 0$ which implies that $f = g$ almost everywhere.

Theorem 5.17: If $\{f_n\}$ is a sequence of measurable functions which is fundamental in measure, then some subsequence $\{f_{n_k}\}$ is **almost uniformly** fundamental.

Proof: For any positive integer k we can find an integer $\bar{n}(k)$ such that if $n \geq \bar{n}(k)$ and $m \geq \bar{n}(k)$, then

$$m(\{x : | f_n(x) - f_m(x) | \geq \frac{1}{2^k}\}) < \frac{1}{2k}$$

We write,

NOTES

$n_1 = \overline{n}(1), n_2 = (n_1 + 1) \cup \overline{n}(2),\ n_3 = (n_2 + 1) \cup \overline{n}(3)$, ...; then $n_1 < n_2 < n_3 < ..$,

So that the sequence $\{f_{n_k}\}$ is certainly a subsequence of $\{k_n\}$. If,

$$E_k = \{x : | f_{n_k}(x) - f_{n_k+1}(x) | \geq \frac{1}{2^k}\}$$

And $k \leq i \leq j$, then for every x which does not belong to $E_k \cup E_{k+1} \cup E_{k+2} \cup .....,$ we have,

$$| f_{n_i}(x) - f_{n_j}(x) | \leq \sum_{m=i}^{\infty} | f_{n_m}(x) - f_{n_{m+1}(x)} | < \sum_{m=i}^{\infty} \frac{1}{2^m} = \frac{1}{2^{i-1}}$$

So that, in other words, the sequence $\{f_{n_i}\}$ is uniformly fundamental on $E \setminus (E_k \cup E_{k+1} \cup ....)$, since

$$m(E_k \cup E_{k+1} \cup ...) \leq \sum_{m=k}^{\infty} m(E_m) < \frac{1}{2^{k-1}}$$

This completes the proof of the Theorem 5.17.

Theorem 5.18: If $\{f_n\}$ is a sequence of measurable functions which is fundamental in measure then there exists a measurable function *f* such that $\{f_n\}$ converges in measure to *f*.

Proof: By Theorem 5.18 we can find a subsequence $\left\{f_{n_k}\right\}$ which is almost uniformly fundamental and therefore fundamental almost everywhere. We write $f(x) = \lim_{k \to \infty} f_{n_k}(x)$ for every x for which the limits exists and observe that, for every $\varepsilon > 0$,

$$\{x : | f_n(x) - f(x) | \geq \varepsilon] \subset \{x : | f_n(x) - f_{n_k}(x) | \geq \frac{\varepsilon}{2}\} \cup \{x : | f_{n_k(x) - f(x)} | \geq \frac{\varepsilon}{2}\}$$

Note here that, the measure of the first term on the right hand side is by hypothesis arbitrarily small if n and n_k are sufficiently large. Also, the measure of the second term also approaches 0 (as $k \to \infty$), since almost uniform convergence implies convergence in measure. Hence, the theorem follows.

Note: Convergence in measure does not essentially imply pointwise convergence at any point.

NOTES

Check Your Progress

12. What is Hölder's inequality? State the Hölder's inequality theorem.
13. What is Minkowski inequality? State the Minkowski inequality theorem.
14. Give an example to show that the L^p functions are complete functions.
15. What are the two distinct mathematical concepts of convergence as per the measure?
16. Prove that if $\{f_n\}$ is a sequence of measurable functions which converges to f almost uniformly, then $\{f_n\}$ converges to f almost everywhere.
17. Define the term almost uniform convergence of a sequence.
18. Prove that almost uniform convergence implies convergence in measure.

5.15 ANSWERS TO 'CHECK YOUR PROGRESS'

1. An outer measure $\mu*$ is an extended real valued set function defined on all subsets of a space X having the following properties:

   (a) $\mu*\phi = 0$

   (b) $A \subset B \Rightarrow \mu*A \leq \mu*B$ (Monotonicity)

   (c) $E \subset \sum_{i=1}^{\infty} E_i \Rightarrow \mu * E \leq \sum_{i=1}^{\infty} \mu * E_i$ (Subadditivity)

   The outer measure $\mu*$ is said to be finite if $\mu*X < \infty$.
2. A measure on an algebra is defined as a non-negative extended real valued set function μ which is typically defined on an algebra A of sets such that,

   (a) $\mu\phi = 0$

   (b) If $<A_i>$ is a disjoint sequence of sets in A whose union is also in A, then

   $$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu A_i$$

   Therefore, a measure on an algebra A is a measure iff A is a σ-algebra.
3. In the measure theory of real analysis, the Carathéodory's extension theorem states that, "Any premeasure defined on a given ring R of subsets of a given set Ω can be extended to a measure on the σ-algebra generated by R, and this extension is unique if the premeasure is σ-finite". The Carathéodory's extension theorem is named after the Greek mathematician Constantin Carathéodory.
4. Any set function P defined on a field $\mathcal{F}_0$ of sets and satisfying the properties of a probability measure on $\mathcal{F}_0$ extends uniquely to a probability measure on the σ-field $\mathcal{F}_0$ generated by $\mathcal{F}_0$.

5. The term complete measure or more specifically a complete measure space is defined as a specific measure space wherein every single subset of every single null set is measurable, i.e., having measure zero. More appropriately, we can state that a measure space (X, Σ, μ) is termed as complete if and only if,

   $S \subseteq N \in \Sigma$ and $\mu(N) = 0 \Rightarrow S \in \Sigma$.

6. Consider that a possibly incomplete measure space (X, Σ, μ) is given, then of this measure space there is an extension (X, Σ_0, μ_0), which is complete. The smallest of the extension, i.e., the smallest σ-algebra Σ_0 is termed as the completion of the measure space.

   Using the following assumptions or statement the completion can be constructed:

   - Let Z be the set of all the subsets of the zero-μ-measure subsets of X, instinctively those elements of Z which are already not in Σ are specifically the ones which prevent completeness from holding true.
   - Let Σ_0 be the σ-algebra created or produced by Σ and Z, i.e., the smallest σ-algebra that contains every element of Σ and of Z.
   - Let μ has an extension μ_0 to Σ_0, which is unique if μ is σ-finite, then it is called the outer measure of μ, given by the infimum.

     $$\mu_0(C) := \inf \{\mu(D) \mid C \subseteq D \in \Sigma\}$$

   Then (X, Σ_0, μ_0) is referred as a complete measure space and is termed as the completion of (X, Σ, μ).

7. In mathematics, the term measure space is defined as a fundamental and essential object of measure theory which analyses the universal generalized and simplified notions of volumes. Characteristically, it comprises of an underlying set, which is referred as the subsets of this set that are feasible and sufficient for measuring the 'σ-Algebra' and the method that is used for measuring the 'Measure'. One significant example of a measure space can be given as a probability space.

   A measure space is defined as a triple $(X, \mathcal{A}, \mu)$.

   Where, X is a Set.

   $\mathcal{A}$ is a σ-Algebra on the Set X

   μ is a Measure on $(X, \mathcal{A})$.

8. Assume that $\mathcal{S}$ is a σ-algebra on a set X. Characteristically, an $\mathcal{S}$-partition of X is defined as a finite collection $A_1, \ldots, A_m$ of disjoint sets in S such that $A_1 \cup \cdots \cup A_m = X$.

   Implementing the convention that $0 \cdot \infty$ and $\infty \cdot 0$ should both be interpreted to be 0.

9. Characteristically, in mathematics, the L^p spaces are function spaces defined using a natural generalization of the p-norm for Finite Dimensional Vector Spaces (FDVS). They are also sometimes called Lebesgue spaces, named after Henri Lebesgue.

NOTES

The set of L^p-functions where $p \geq 1$ generalizes L^2-space. As an alternative of square integrable, the measurable function f must be p-integrable for f to be in L^p.

10. A function ϕ defined an open interval (a, b) is known as a **convex function** if for each $x, y \in (a, b)$ and λ, μ such that $\lambda, \mu \geq 0$ and $\lambda + \mu = 1$, we have,

    $\phi(\lambda x + \mu y) \leq \lambda\phi(x) + \mu\phi(y)$

    The end points a, b can take the values $-\infty$, ∞, respectively.

11. If f is a convex function and X is a random variable, then $Ef(X)$ ³ $f(EX)$. Furthermore, if f is strictly convex, then equality implies that $X = EX$ with probability 1, that is, X is constant.

12. In real analysis, Hölder's inequality, named after Otto Hölder, is a fundamental inequality between integrals and an indispensable tool for the study of L^p spaces.

    Let (S, Σ, μ) be a measure space and let $p, q \in [1, \infty]$ with $1/p + 1/q = 1$. Then for all measurable real valued function or complex valued function f and g on S,

    $$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

13. In real analysis, the Minkowski inequality establishes that the L^p spaces are normed vector spaces. The inequality is named after the German mathematician Hermann Minkowski.

    Let S be a measure space, let $1 \leq p < \infty$ and let f and g be elements of $L^p(S)$. Then $f + g$ is in $L^p(S)$, and we have the triangle inequality,

    $$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

    With equality for $1 < p < \infty$ if and only if f and g are positively linearly dependent, i.e., $f = \lambda_g$ for some $\lambda \geq 0$ or $g = 0$.

14. Characteristically, the L^p functions have certainly ambiguous pointwise values. However, we usually consider L^p functions as complete functions. A simple example of this construction, for a measure that has no sets of measure 0, consequently requires no quotient is given by,

    $$\ell^p = \{\text{complex sequences } \{c_i\} \text{ with } \sum_i |c_i|^p < \infty\}$$

    With standard norm,

    $$|(c_1, c_2, \ldots)|_{\ell^p} = (\textstyle\sum_i |c_i|^p)^{1/p}.$$

15. Convergence in measure is either of following two distinct mathematical concepts both of which generalize the concept of convergence in probability.

    Let f, f_n $(n \in \mathbb{N}) : X \rightarrow \mathbb{R}$ be the measurable functions defined on a measure space (X, Σ, μ). The sequence f_n is said to converge globally in measure to f if for every $\varepsilon > 0$,

NOTES

$$\lim_{n\to\infty} \mu(\{x \in X : |f(x) - f_n(x)| \geq \varepsilon\}) = 0.$$

And to converge locally in measure to f if for every $\varepsilon > 0$ and every $F \in \Sigma$ with $\mu(F) < \infty$,

$$\lim_{n\to\infty} \mu(\{x \in F : |f(x) - f_n(x)| \geq \varepsilon\}) = 0.$$

On a finite measure space, both notions are equivalent. Otherwise, convergence in measure can refer to either global convergence in measure or local convergence in measure.

16. Let F_n be a measurable set such that $m(F_n) < 1/n$ and such that the sequence $\{f_n\}$ converges to f uniformly on F_n^c, $n = 1, 2, \ldots$. If $F = \bigcap_{n=1}^{\infty} F_n$, then $m(F) \leq \mu(F_n) < \frac{1}{n}$ so that $m(F) = 0$, and it is clear that, for $x \in F^c$, $\{f_n(x)\}$ converges to $f(x)$.
17. If the domain of the functions is a measure space E then the related notion of 'almost uniform convergence' can be defined. We define that a sequence of functions (f_n) converges almost uniformly on E if for every $\delta > 0$ there exists a measurable set E_δ with measure less than δ such that the sequence of functions (f_n) converges uniformly on $E \setminus E_\delta$. In other words, almost uniform convergence means there are sets of arbitrarily small measure for which the sequence of functions converges uniformly on their complement.
18. If $\{f_n\}$ converges to f almost uniformly, then for any two positive numbers ε and δ there exists a measurable set F such that $m(F) < \delta$ such that $|f_n(x) - f(x)| < \varepsilon$, whenever x belongs to F^c and n is sufficiently large.

5.16 SUMMARY

- In the measure theory, the concept of a measure is a generalization of common notions, such as mass, distance/length, area, volume, etc.
- An outer measure or exterior measure is a function defined on all subsets of a given set with values in the extended real numbers satisfying some additional technical conditions. The theory of outer measures was first introduced by Constantin Carathéodory to provide an abstract basis for the theory of measurable sets and countably additive measures.
- The class β of μ^*-measurable sets are σ-algebra. If $\bar{\mu}$ is restricted to β, then $\bar{\mu}$ is a complete measure on β.
- The union of any sequence of sets in an algebra which can be replaced by a disjoint union of sets in an algebra, it follows that β is a σ-algebra.

NOTES

- If $A \in \boldsymbol{A}$ and if $<A_i>$ is any sequence of sets in $\boldsymbol{A}$, such that $A \subset \bigcup_{i=1}^{\infty} A_i$, then show that $\mu A \leq \sum_{i=1}^{\infty} \mu A_i$.
- The set function $\mu*$ is an outer measure.
- If $A \in \boldsymbol{A}$, then A is measurable with respect to $\mu*$.
- The outer measure $\mu*$ which we have defined above is known as the outer measure induced by μ.
- Algebra A of sets we use A_σ to denote those sets which are countable unions of sets of A and use $A_{\sigma\delta}$ to denote those sets which are countable intersection of sets in A_σ.
- Let μ be a measure on an algebra A, $\mu*$ be the outer measure induced by μ and E be any set. Then for $\varepsilon > 0$, there exists a set $A \in A_\sigma$ with $E \subset A$ and $\mu * A \leq \mu * E + \varepsilon$.
- In the measure theory of real analysis, the Carathéodory's extension theorem states that, "Any premeasure defined on a given ring R of subsets of a given set Ω can be extended to a measure on the σ-algebra generated by R, and this extension is unique if the premeasure is σ-finite". The Carathéodory's extension theorem is named after the Greek mathematician Constantin Carathéodory.
- The Carathéodory's extension theorem is also occasionally termed as the Carathéodory-Fréchet extension theorem, the Carathéodory–Hopf extension theorem, the Hopf extension theorem and the Hahn–Kolmogorov extension theorem.
- The Carathéodory's extension theorem considered very significant as it helps in constructing a measure by defining it on a small algebra of sets, so that its sigma additivity can be verified. Additionally, this theorem also ensures its extension to a σ-algebra.
- Any set function P defined on a field $\mathcal{F}_0$ of sets and satisfying the properties of a probability measure on $\mathcal{F}_0$ extends uniquely to a probability measure on the σ-field $\mathcal{F}_0$ generated by $\mathcal{F}_0$.
- Consequently, we can define that $\mathcal{F}$ be the collection of sets $A \subset \Omega$ having the same inner measure and outer measure, and then subsequently we can define that $P(A)$.
- The term complete measure or more specifically a complete measure space is defined as a specific measure space wherein every single subset of every single null set is measurable, i.e., having measure zero. More appropriately, we can state that a measure space (X, Σ, μ) is termed as complete if and only if,

  $$S \subseteq N \in \Sigma \quad \text{and} \quad \mu(N) = 0 \Rightarrow S \in \Sigma$$

- The term completeness can be essentially illustrated by considering the typical product space problems.

- While this approach does define a measure space, it has a flaw. Since every singleton set has one-dimensional Lebesgue measure zero,

  $\lambda^2 (\{0\} \times A) \leq \lambda(\{0\}) = 0.$

- The given larger set also have λ^2-measure zero. Consequently, as defined above this 'Two-Dimensional Lebesgue Measure' is not complete, hence some completion procedure is essential.
- Consider that a possibly incomplete measure space $(X, \sum, \mu)$ is given, then of this measure space there is an extension $(X, \sum_0, \mu_0)$, which is complete. The smallest of the extension, i.e., the smallest σ-algebra $\sum_0$ is termed as the completion of the measure space.
- The Borel measure when defined on the Borel σ-algebra specifically created or produced by the open intervals of the real line is not complete, and therefore the above defined completion procedure has to be used for defining the complete Lebesgue measure.
- A measure space is characteristically defined as a measurable space that possesses a non-negative measure. The typical examples of measure spaces include n-dimensional Euclidean space with Lebesgue measure and the unit interval with Lebesgue measure, i.e., probability.
- The Lebesgue integral depends ultimately on the idea of measure. In particular, the mathematical framework requires a set, a σ-algebra of subsets alongwith a set function that assigns a non-negative number (called its measure) to each set in the σ-algebra.
- In mathematics, specifically in the real analysis, the integral of a non-negative function of a single variable can be simply interpreted as the area between the graph of that function and the X-axis. The Lebesgue integral extends the integral to a larger class of functions. It also extends the domains on which these functions can be defined.
- The integral of a positive function f between limits a and b can be interpreted as the area under the graph of f. However, Riemann integration does not interact accurately by taking limits of sequences of functions, because producing such limiting processes are difficult for analyses.
- Lebesgue's theory defines integrals for a class of functions called measurable functions. A real valued function f on E is measurable if the preimage of every interval of the form (t, ∞), i.e., any Borel set is in X:

  $\{x \mid f(x) > t\} \in X \quad \forall t \in \mathbb{R}.$

- Characteristically, a measurable function which is bounded by an integrable function is integrable. Every integrable function is measurable. If a sequence of measurable functions converges almost everywhere, then its limit is measurable. If a sequence of measurable functions converges asymptotically, then its limit is measurable. Basically, the set of measurable functions is defined as a linear space. Additionally, the intersection and union of two measurable functions are measurable.

NOTES

- Integration with respect to a measure is termed as Lebesgue integration. The following definition illustrates that the Lebesgue integration functions as anticipated on the simple functions represented as linear combinations of characteristic functions of disjoint sets.
- The L^p-functions are the characteristic functions for which this integral converges. For $p \neq 2$, the space of L^p-functions is a Banach space which is not a Hilbert space.
- Characteristically, a real valued function is called convex if the line segment between any two points on the graph of the function does not lie below the graph between the two points. Equivalently, a function is convex if its epigraph, i.e., the set of points on or above the graph of the function is a convex set.
- In simple terms, a convex function refers to a function whose graph is shaped like a cup $\cup$, while a concave function's graph is shaped like a cap $\cap$.
- Convex functions play a significant role in several areas of mathematics. Even in infinite dimensional spaces, under suitable additional hypotheses, the convex functions continue to satisfy such properties and as a result, they are the most significant functionals in the calculus of variations.
- In probability theory, a convex function applied to the expected value of a random variable is always bounded above by the expected value of the convex function of the random variable.
- Let ϕ be convex on (a, b) and $a < s < t < u < b$, then

  $$\frac{\phi(t)-\phi(s)}{t-s} \leq \frac{\phi(u)-\phi(s)}{u-s} \leq \frac{\phi(u-\phi(t)}{u-t}$$

  If ϕ is strictly convex, equality will not occur.
- A differentiable function ϕ is convex on (a, b) iff ϕ' is a monotonically increasing function. If ϕ'exists on (a, b), then ϕ is convex iff $\phi'' \geq 0$ on (a, b) and strictly convex if $\psi'' > 0$ on (a, b).
- If ϕ is convex on (a, b), then ϕ is absolutely continuous on each closed subinterval of (a, b).
- Every convex function on an open interval is continuous.
- In mathematical analysis, the term Jensen's inequality is named after the Danish mathematician Johan Jensen, it was proved by Jensen in 1906. The Jensen's inequality relates the value of a convex function of an integral to the integral of the convex function.
- The classical form of Jensen's inequality includes several numbers and weights. The inequality can be stated quite commonly using either the language of measure theory or equivalently the probability.
- If f is a convex function and X is a random variable, then $Ef(X) \geq f(EX)$. Furthermore, if f is strictly convex, then equality implies that $X = EX$ with probability 1, that is, X is constant.

NOTES

- In real analysis, Hölder's inequality, named after Otto Hölder, is a fundamental inequality between integrals and an indispensable tool for the study of L^p spaces.
- Let (S, Σ, μ) be a measure space and let $p, q \in [1, \infty]$ with $1/p + 1/q = 1$. Then for all measurable real valued function or complex valued function f and g on S,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

- If, in addition, $p, q \in (1, \infty)$ and $f \in L^p(\mu)$ and $g \in L^q(\mu)$, then Hölder's inequality becomes an equality if and only if $|f|^p$ and $|g|^q$ are linearly dependent in $L^1(\mu)$, meaning that there exist real numbers $\alpha, \beta \geq 0$, not both of them zero, such that $\alpha|f|^p = \beta\,|g|^q$ μ-almost everywhere.
- Hölder's inequality is used to prove the Minkowski inequality, which is the triangle inequality in the space $L^p(\mu)$, and also to establish that $L^q(\mu)$ is the dual space of $L^p(\mu)$ for $p \in [1, \infty)$.
- In real analysis, the Minkowski inequality establishes that the L^p spaces are normed vector spaces. The inequality is named after the German mathematician Hermann Minkowski.
- Let S be a measure space, let $1 \leq p < \infty$ and let f and g be elements of $L^p(S)$. Then $f + g$ is in $L^p(S)$, and we have the triangle inequality
- Characteristically, the L^p functions have certainly ambiguous pointwise values. However, we usually consider L^p functions as complete functions.
- Essentially, to prove a Cauchy sequence f_i in $L^p(X)$ has a subsequence which converges pointwise off a set of measure 0 in X.
- The vector space of equivalence classes of measurable functions on (S, Σ, μ) is denoted ad $L^0(S, \Sigma, \mu)$. By definition, it contains all the L^p, and is equipped with the topology of convergence in measure. When μ is a probability measure (i.e., $\mu(S) = 1$), this mode of convergence is named convergence in probability.
- The space $L^p(X)$ is a complete metric space.
- The infinite sum is not necessarily claimed to converge to a finite value for every x.
- Converges for almost all $x \in X$. Let $f(x)$ be the sum at points x where the series converges, and on the measure zero set where the series does not converge put $f(x) = 0$.
- L^p is complete, i.e., every Cauchy sequence converge.
- Convergence in measure is either of following two distinct mathematical

t é nconcepts both of which generalize the concept of convergence in probabiao- t

- If $\{f_n\}$ is a sequence of measurable functions which is fundamental in measure

NOTES

NOTES

- **The L^p spaces:** The L^p spaces are function spaces defined using a natural generalization of the p-norm for Finite Dimensional Vector Spaces (FDVS). They are also sometimes called Lebesgue spaces, named after Henri Lebesgue.
- **Convex function:** A function ϕ defined an open interval (a, b) is known as a convex function if for each $x, y \in (a, b)$ and λ, μ such that $\lambda, \mu \geq 0$ and $\lambda + \mu = 1$, we have, $\phi(\lambda x + \mu y) \leq \lambda\phi(x) + \mu\phi(y)$. The end points a, b can take the values $-\infty, \infty$, respectively.
- **Jensen's inequality:** In mathematical analysis, the term Jensen's inequality is named after the Danish mathematician Johan Jensen, it was proved by Jensen in 1906. The Jensen's inequality relates the value of a convex function of an integral to the integral of the convex function.
- **Hölder's inequality:** In real analysis, Hölder's inequality, named after Otto Hölder, is a fundamental inequality between integrals and an indispensable tool for the study of L^p spaces.
- **Minkowski inequality:** In real analysis, the Minkowski inequality establishes that the L^p spaces are normed vector spaces. The inequality is named after the German mathematician Hermann Minkowski.
- **Almost uniform convergence:** If the domain of the functions is a measure space E then the related notion of 'almost uniform convergence' can be defined by means of a sequence of functions (f_n) that converges almost uniformly on E if for every $\delta > 0$ there exists a measurable set E_δ with measure less than δ such that the sequence of functions (f_n) converges uniformly on $E \setminus E_\delta$.

5.18 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What are the outer measures?
2. Define the extension of a measure.
3. State about the uniqueness of extension.
4. What is the completeness of measure?
5. What do you understand by the measure space?
6. Define the integration with respect to a measure.
7. State the L^p spaces.
8. What are the convex functions.
9. State the Jensen's inequality.
10. Define the Hölder's inequality.
11. State the Minkowski's inequality.
12. What is the completeness of L^p?
13. Why the convergence in measure is used?

NOTES

Long-Answer Questions

1. Briefly discuss the measures and outer measures giving theorems, proofs and appropriate examples.
2. Explain the concept of extension of a measure and uniqueness of extension with the help of theorems, proofs, and examples.
3. What is completion of a measure? Explain with the help of examples.
4. Briefly explain the measures spaces and show that the measure μ on $\mathcal{A}$ is a set function having domain $\mathcal{A}$ satisfying the following:
   (a) $\mu(A) \geq 0$ for all $A \in \mathcal{A}$.
   (b) $\mu(\phi) = 0$.
5. Discuss the concept of integration with respect to a measure giving relevant examples.
6. Show that a measure space (E, X, μ) where E is a set, X is a σ-algebra of subsets of E, and μ is a (non-negative) measure on E can be defined on the sets of X.
7. What are L^p spaces? Explain the concept with the help of appropriate examples.
8. Prove that the L^p spaces are function spaces defined using a natural generalization of the p-norm for Finite-Dimensional Vector Spaces (FDVS).
9. Describe the convex functions giving definitions, theorems, proofs, and examples.
10. Elaborate on the Jensen's inequality with the help of theorems and examples.
11. Explain Hölder's inequalities and Minkowski's inequalities giving theorems and proofs.
12. Differentiate between the Holder's inequalities and Minkowski's inequalities.
13. Briefly discuss about the completeness of L^p giving theorems and proofs.
14. How the convergence in measure is done? Explain giving appropriate theorems and examples.
15. Discuss about the almost uniform convergence with the help of theorems, proofs and examples.
16. Show that if $\{f_n\}$ is a sequence of measurable functions which converges to f almost uniformly, then $\{f_n\}$ converges to f almost everywhere.
17. Prove that if $\{f_n\}$ converges in measure to f, then $\{f_n\}$ is fundamental in measure. Also, if $\{f_n\}$ converges in measure to g, then $f = g$ almost everywhere.

5.19 FURTHER READING

Rudin, Walter. 2017. *Real and Complex Analysis*, Third Edition. Noida: McGraw-Hill Education.

NOTES

Gupta, S. L. and Nisha Rani. 2004. *Fundamental Real Analysis*, Fourth Edition. New Delhi: Vikas Publishing House Pvt. Ltd.

Carothers, N. L. 2000. *Real Analysis*, First Edition. Cambridge (U.K.): Cambridge University Press.

Bartle, Robert G. and Donald R. Sherbert. 2014. *Introduction to Real Analysis*, Fourth Edition. New York: Wiley.

Trench, William F. 2002. *Introduction to Real Analysis*, London: Pearson.

Loeb, Peter A. 2016. *Real Analysis*, Basel (Switzerland): Birkhäuser.

Royden, Halsey. 2015. *Real Analysis*, Fourth Edition. Noida: Pearson Education India.