REAL ANALYSIS II

Edited By
Dr. Sachin Kaushal

SYLLABUS
Real Analysis II

Objectives:

- To allows an appreciation of the many interconnections between areas of mathematics.
- To learn about the countability of sets, metric space, continuity, discontinuities, connectedness and compactness for set of real numbers.

Sr. No.	Content
1	Uniform convergence and differentiation, Equi-continuous families of functions, Arzela's Theorem and Weierstrass Approximation Theorem
2	Reimann Stieltje's integral , Definition and existence of integral, Properties of integration ,R-S integral as a limit of sum
3	Differentiation and integration, fundamental Theorem of Calculus, Mean value Theorems .
4	Lebesgue Measure ;Outer Measure , Measurable sets and Lebesgue measure, A non measurable set, Measurable functions, Littlewood's three principles
5	The Lebesgue Integral of bounded functions, Comparison of Riemann and Lebesgue Integrals,The integral of a non-negative function, General Lebesgue integral, Convergence of measure .

CONTENT

Unit 1: Equicontinuous

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Explain the equicontinuity
- Describe the properties of equicontinuous
- Discuss the equicontinuity and uniform convergence
- Define stochastic equicontinuity

Introduction

In last unit, you have studied about the uniform converges and differentiation. This unit provides you the explanation of Equicontinuity. In mathematical analysis, a family of functions is equicontinuous if all the functions are continuous and they have equal variation over a given neighbourhood, in a precise sense described herein. In particular, the concept applies to countable families, and thus sequences of functions.

1.1 Equicontinuity

The equicontinuity appears in the formulation of Ascoli's theorem, which states that a subset of C(X), the space of continuous functions on a compact Hausdorff space X, is compact if and only if it is closed, pointwise bounded and equicontinuous. A sequence in C(X) is uniformly convergent if and only if it is equicontinuous and converges pointwise to a function (not necessarily continuous a-prior). In particular, the limit of an equicontinuous pointwise convergent sequence of continuous functions f_n on either metric space or locally compact space is continuous. If, in addition, f_n are homomorphic, then the limit is also homomorphic.

The uniform boundedness principle states that a pointwise bounded family of continuous linear operators between Banach spaces is equicontinuous.

1.2 Families of Equicontinuous

Let X and Y be two metric spaces, and F a family of functions from X to Y.

The family F is equicontinuous at a point $x_0 \in X$ if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_0), f(x)) < \varepsilon$ for all $f \in F$ and all x such that $d(x_0, x) < \delta$. The family is equicontinuous if it is equicontinuous at each point of X.

The family F is uniformly equicontinuous if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_1), f(x_2)) < \varepsilon$ for all $f \in F$ and all $x_1, x_2 \in X$ such that $d(x_1, x_2) < \delta$.

For comparison, the statement all functions f in F are continuous' means that for every $\varepsilon > 0$, every $f \in F$, and every $x_0 \in X$, there exists a $\delta > 0$ such that $d(f(x_0), f(x)) < \varepsilon$ for all $x \in X$ such that $d(x_0, x) < \delta$. So, for continuity, δ may depend on ε, x_0 and f; for equicontinuity, δ must be independent of f; and for uniform equicontinuity, δ must be independent of both f and x_0.

More generally, when X is a topological space, a set F of functions from X to Y is said to be equicontinuous at x if for every $\varepsilon > 0$, x has a neighbourhood U_x such that

$$d_Y(f(y), f(x)) < \varepsilon$$

for all $y \in U_x$ and $f \in F$. This definition usually appears in the context of topological vector spaces.

When X is compact, a set is uniformly equicontinuous if and only if it is equicontinuous at every point, for essentially the same reason as that uniform continuity and continuity coincide on compact spaces.

Some basic properties follow immediately from the definition. Every finite set of continuous functions is equicontinuous. The closure of an equicontinuous set is again equicontinuous. Every member of a uniformly equicontinuous set of functions is uniformly continuous, and every finite set of uniformly continuous functions is uniformly equicontinuous.

Example:

- A set of functions with the same Lipschitz constant is (uniformly) equicontinuous. In particular, this is the case if the set consists of functions with derivatives bounded by the same constant.
- Uniform boundedness principle gives a sufficient condition for a set of continuous linear operators to be equicontinuous.
- A family of iterates of an analytic function is equicontinuous on the Fatou set.

Properties of Equicontinuous

- If a subset $\mathcal{F} \subseteq C(X, Y)$ is totally bounded under the uniform metric, and then $\mathcal{F}$ is equicontinuous.
- Suppose X is compact. If a sequence of functions $\{f_n\}$ in $C(X\ \mathbb{R}k)$ is equibounded and equicontinuous, then the sequence $\{f_n\}$ has a uniformly convergent subsequence. (Arzelà's theorem)
- Let $\{f_n\}$ be a sequence of functions in C(X, Y). If $\{f_n\}$ is equicontinuous and converges pointwise to a function $f : X \to Y$, then f is continuous and $\{f_n\}$ converges to f in the compact-open topology.

Notes

1.3 Equicontinuity and Uniform Convergence

Let X be a compact Hausdorff space, and equip C(X) with the uniform norm, thus making C(X) a Banach space, hence a metric space. Then Ascoli's theorem states that a subset of C(X) is compact if and only if it is closed, pointwise bounded and equicontinuous. This is analogous to the Heine-Borel theorem, which states that subsets of $\mathbb{R}^n$ are compact if and only if they are closed and bounded. Every bounded equicontinuous sequence in C(X) contains a subsequence that converges uniformly to a continuous function on X.

In view of Ascoli's theorem, a sequence in C(X) converges uniformly if and only if it is equicontinuous and converges pointwise. The hypothesis of the statement can be weakened a bit: a sequence in C(X) converges uniformly if it is equicontinuous and converges pointwise on a dense subset to some function on X (not assumed continuous). This weaker version is typically used to prove Ascoli's theorem for separable compact spaces. Another consequence is that the limit of an equicontinuous pointwise convergent sequence of continuous functions on a metric space, or on a locally compact space, is continuous.

In the above, the hypothesis of compactness of X cannot be relaxed. To see that, consider a compactly supported continuous function g on $\mathbb{R}$ with g(0) = 1, and consider the equicontinuous sequence of functions $\{f_n\}$ on $\mathbb{R}$ defined by $f_n(x) = g(x - n)$. Then, f_n converges pointwise to 0 but does not converge uniformly to 0.

This criterion for uniform convergence is often useful in real and complex analysis. Suppose we are given a sequence of continuous functions that converges pointwise on some open subset G of $\mathbb{R}^n$. As noted above, it actually converges uniformly on a compact subset of G if it is equicontinuous on the compact set.

In practice, showing the equicontinuity is often not so difficult. For example, if the sequence consists of differentiable functions or functions with some regularity (e.g., the functions are solutions of a differential equation), then the mean value theorem or some other kinds of estimates can be used to show the sequence is equicontinuous.

It then follows that the limit of the sequence is continuous on every compact subset of G; thus, continuous on G. A similar argument can be made when the functions are homomorphic. One can use, for instance, Cauchy's estimate to show the equicontinuity (on a compact subset) and conclude that the limit is homomorphism. Note that the equicontinuity is essential here. For example, $f_n(x)$ = arctan nx converges to a multiple of the discontinuous sign function.

1.4 Equicontinuity Families of Linear Operators

Let E, F be Banach spaces, and Γ be a family of continuous linear operators from E into F. Then Γ is equicontinuous if and only if

$$\text{Sup}\{||T|| : T \in \Gamma\} < \propto$$

that is, Γ is uniformly bounded in operator norm. Also, by linearity, Γ is uniformly equicontinuous if and only if it is equicontinuous at 0.

The uniform boundedness principle (also known as the Banach-Steinhaus theorem) states that Γ is equicontinuous if it is pointwise bounded; i.e., $\sup\{||T(x)|| : T \in \Gamma\} < \infty$ for each $x \in E$. The result can be generalized to a case when F is locally convex and E is a barreled space.

Alaoglu's theorem states that if E is a topological vector space, then every equicontinuous subset of E* is weak-* relatively compact.

1.5 Equicontinuity in Topological Spaces

The most general scenario in which equicontinuity can be defined is for topological spaces whereas *uniform* equicontinuity requires the filter of neighbourhoods of one point to be somehow comparable with the filter of neighbourhood of another point. The latter is most generally done via a uniform structure, giving a uniform space. Appropriate definitions in these cases are as follows:

A set A of functions continuous between two topological spaces X and Y is **topologically equicontinuous** at the points $x \in X$ and $y \in Y$ if for any open set O about y, there are neighbourhoods U of x and V of y such that for every $f \in A$, if the intersection of f[U] and V is non-empty, $f(U) \subseteq O$. One says A is said to be topologically equicontinuous at $x \in X$ if it is topologically equicontinuous at x and y for each $y \in Y$. Finally, A is equicontinuous if it is equicontinuous at x for all points $x \in X$.

A set A of continuous functions between two uniform spaces X and Y is **uniformly equicontinuous** if for every element W of the uniformity on Y, the set

$$\{(u, v) \in X \times X : \text{for all } f \in A.\ (f(u), f(v)) \in W\}$$

is a member of the uniformity on X

A weaker concept is that of even continuity:

A set A of continuous functions between two topological spaces X and Y is said to be **evenly continuous at** $x \in X$ and $y \in Y$ if given any open set O containing y there are neighbourhoods U of x and V of y such that $f[U] \subseteq O$ whenever $f(x) \in V$. It is **evenly continuous at** x if it is evenly continuous at x and y for every $y \in Y$, and **evenly continuous** if it is evenly continuous at x for every $x \in X$.

For metric spaces, there are standard topologies and uniform structures derived from the matrices, and then these general definitions are equivalent to the metric-space definitions.

1.6 Stochastic Equicontinuity

Stochastic equicontinuity is a version of equicontinuity used in the context of sequences of functions of random variables, and their convergence.

Let $\{H_n(\theta) : n \geq 1\}$ be a family of random functions defined from, where $\Theta \to \mathbb{R}$ where Θ is any normed metric space. Here $\{H_n(\theta)\}$ might represent a sequence of estimators applied to datasets of size n, given that the data arises from a population for which the parameter indexing the statistical model for the data is θ. The randomness of the functions arises from the data generating process under which a set of observed data is considered to be a realisation of a probabilistic or statistical model. However, in $\{H_n(\theta)\}$, θ relates to the model currently being postulated or fitted rather than to an underlying model which is supposed to represent the mechanism generating the data. Then $\{H_n\}$ is stochastically equicontinuous if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that:

$$\lim_{n\to\infty} \Pr\left(\sup_{\theta\in\Theta} \sup_{\theta'\in B(\theta,\delta)} |H_n(\theta') - H_n(\theta)| > \varepsilon \right) < \delta$$

Here $B(\theta, \delta)$ represents a ball in the parameter space, centered at θ and whose radius depends on.

Self Assessment

Fill in the blanks:

1. The ……………………..states that a pointwise bounded family of continuous linear operators between Banach spaces is equicontinuous.

2. The family F is ……………………$x_0 \in X$ if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_0), f(x)) < \varepsilon$ for all $f \in F$ and all x such that $d(x_0, x) < \delta$. The family is equicontinuous if it is equicontinuous at each point of X.
3. Suppose X is compact. If a sequence of functions $\{f_n\}$ in C(X, ℝk) is equibounded and equicontinuous, then the sequence $\{f_n\}$ has a …………………………..
4. The uniform boundedness principle is also known as …………………….. states that Γ is equicontinuous if it is pointwise bounded; i.e., $\sup\{||T(x)|| : T \in \Gamma\} < \infty$ for each $x \in E$. The result can be generalized to a case when F is locally convex and E is a barreled space.
5. …………….. is a version of equicontinuity used in the context of sequences of functions of random variables, and their convergence.

1.7 Summary

- In particular, the limit of an equicontinuous pointwise convergent sequence of continuous functions f_n on either metric space or locally compact space is continuous. If, in addition, f_n are holomorphic, then the limit is also holomorphic.
- The uniform boundedness principle states that a pointwise bounded family of continuous linear operators between Banach spaces is equicontinuous.
- The family F is **equicontinuous at a point** $x_0 \in X$ if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_0), f(x)) < \varepsilon$ for all $f \in F$ and all x such that $d(x_0, x) < \delta$. The family is **equicontinuous** if it is equicontinuous at each point of X.
- The family F is **uniformly equicontinuous** if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_1), f(x_2)) < \varepsilon$ for all $f \in F$ and all $x_1, x_2 \in X$ such that $d(x_1, x_2) < \delta$.
- If a subset $\mathcal{F} \subseteq C(X, Y)$ is totally bounded under the uniform metric, and then $\mathcal{F}$ is equicontinuous.
- Suppose X is compact. If a sequence of functions $\{f_n\}$ in C(X, ℝk) is equibounded and equicontinuous, then the sequence $\{f_n\}$ has a uniformly convergent subsequence. (Arzelà's theorem)
- Let fn be a sequence of functions in C(X, Y). If $\{f_n\}$ is equicontinuous and converges pointwise to a function $f : X \to Y$, then f is continuous and $\{f_n\}$ converges to f in the compact-open topology.
- The uniform boundedness principle (also known as the Banach-Steinhaus theorem) states that Γ is equicontinuous if it is pointwise bounded; i.e., $\sup\{||T(x)|| : T \in \Gamma\} < \infty$ for each $x \in E$. The result can be generalized to a case when F is locally convex and E is a barreled space.
- Stochastic equicontinuity is a version of equicontinuity used in the context of sequences of functions of random variables, and their convergence.

1.8 Keywords

Stochastic Equicontinuity: Stochastic equicontinuity is a version of equicontinuity used in the context of sequences of functions of random variables, and their convergence

Uniformly Equicontinuous: The family F is uniformly equicontinuous if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_1), f(x_2)) < \varepsilon$ for all $f \in F$ and all $x_1, x_2 \in X$ such that $d(x_1, x_2) < \delta$.

Equicontinuous at a Point: The family F is equicontinuous at a point $x_0 \in X$ if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d(f(x_0), f(x)) < \varepsilon$ for all $f \in F$ and all x such that $d(x_0, x) < \delta$. The family is equicontinuous if it is equicontinuous at each point of X.

Uniform Boundedness: The uniform boundedness principle states that a pointwise bounded family of continuous linear operators between Banach spaces is equicontinuous.

1.9 Review Questions

1. Explain the Equicontinuity and Families of Equicontinuous.
2. Describe the Properties of equicontinuous.
3. Discuss the Equicontinuity and uniform convergence.
4. Describe Equicontinuity families of linear operators.
5. Explain the Equicontinuity in topological spaces.
6. Define Stochastic equicontinuity.

Answers: Self Assessment

1. uniform boundedness principle
2. equicontinuous at a point
3. uniformly convergent subsequence
4. the Banach-Steinhaus theorem
5. Stochastic equicontinuity

1.10 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 2: Arzelà's Theorem and Weierstrass Approximation Theorem

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the Arzelà's Theorem
- Describe the Weierstrass Approximation Theorem

Introduction

In last unit you have studied about the uniform convergence and Equicontinuity. This unit provides you the explanation of Arzelà's Theorem and Weierstrass Approximation theorem. Our setting is a compact metric space *X* which you can, if you wish, take to be a compact subset of R*n*, or even of the complex plane (with the Euclidean metric, of course). Let *C*(*X*) denotes the space of all continuous functions on *X* with values in C (equally well, you can take the values to lie in R). In *C*(*X*) we always regard the distance between functions *f* and *g* in *C*(*X*) to be

$$\text{dist}(f, g) = \max\{|f(x) - g(x)| : x \in X\}$$

2.1 Arzelà–Ascoli Theorem

A sequence $\{f_n\}_{n \in N}$ of continuous functions on an interval I = [a, b] is uniformly bounded if there is a number M such that

$$|f_n(x)| \leq M$$

for every function f_n belonging to the sequence, and every $x \in [a, b]$. The sequence is *equicontinuous* if, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f_n(x) - f_n(y)| < \varepsilon \text{ Whenever } |x - y| < \delta$$

for every f_n belonging to the sequence. Succinctly, a sequence is equicontinuous if and only if all of its elements have the same modulus of continuity. In simplest terms, the theorem can be stated as follows:

Consider a sequence of real-valued continuous functions $(f_n)_{n\in N}$ defined on a closed and bounded interval [a, b] of the real line. If this sequence is uniformly bounded and equicontinuous, then there exists a subsequence (f_{nk}) that converges uniformly.

Notes

Proof: The proof is essentially based on a diagonalization argument. The simplest case is of real-valued functions on a closed and bounded interval:

Let $I = [a, b] \subset R$ be a closed and bounded interval. If F is an infinite set of functions $f : I \to R$ which is uniformly bounded and equicontinuous, then there is a sequence f_n of elements of F such that f_n converges uniformly on I.

Fix an enumeration $\{x_i\}_{i = 1,2,3,...}$ of rational numbers in I. Since F is uniformly bounded, the set of points $\{f(x_1)\}_{f \in F}$ is bounded, and hence by the Bolzano-Weierstrass theorem, there is a sequence $\{f_{n1}\}$ of distinct functions in F such that $\{f_{n1}(x_1)\}$ converges. Repeating the same argument for the sequence of points $\{f_{n1}(x_2)\}$, there is a subsequence $\{f_{n2}\}$ of $\{f_{n1}\}$ such that $\{f_{n2}(x_2)\}$ converges.

By mathematical induction this process can be continued, and so there is a chain of subsequences

$$\{f_{n1}\} \supset \{f_{n2}\} \supset \ldots$$

such that, for each k = 1, 2, 3,…, the subsequence $\{f_{nk}\}$ converges at $x_1,...,x_k$. Now form the diagonal subsequence {f} whose mth term f_m is the mth term in the mth subsequence $\{f_{nm}\}$. By construction, f_m converges at every rational point of I.

Therefore, given any $\varepsilon > 0$ and rational x_k in I, there is an integer $N = N(\varepsilon, x_k)$ such that

$$|f_n(x_k) - f_m(x_k)| < \varepsilon/3, \quad n, m \geq N.$$

Since the family F is equicontinuous, for this fixed å and for every x in I, there is an open interval U_x containing x such that

$$|f(s) - f(t)| < \varepsilon/3$$

for all $f \in F$ and all s, t in I such that $s, t \in U_x$.

The collection of intervals U_x, $x \in I$, forms an open cover of I. Since I is compact, this covering admits a finite subcover $U_1, ..., U_J$. There exists an integer K such that each open interval U_j, $1 \leq j \leq J$, contains a rational x_k with $1 \leq k \leq K$. Finally, for any $t \in I$, there are j and k so that t and x_k belong to the same interval U_j. For this choice of k,

$$|f_n(t) - f_m(t)| \leq |f_n(t) - f_n(x_k)| + |f_n(x_k) - f_m(x_k)| + |f_m(x_k) - f_m(t)| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3$$

for all $n, m > N = \max\{N(\varepsilon, x_1), ..., N(\varepsilon, x_K)\}$. Consequently, the sequence $\{f_n\}$ is uniformly Cauchy, and therefore converges to a continuous function, as claimed. This completes the proof.

Theorem 1: Weierstrass Approximation Theorem

Let $f: [a, b] \to \mathbb{R}$ be continuous. Then there is a sequence of polynomials $\{P_n\}_{n=1}^{\infty}$ such that $p_n \to f$ uniformly.

> *Notes* It is important that [a, b] is a closed interval. If it was open, we could take (0, 1) and $f(x) = 1/x$, which is unbounded. But every polynomial is bounded on (0, 1) and therefore no sequence of polynomials could converge to f uniformly.

It will suffice to prove Weierstrass Approximation Theorem on [0, 1] from which the general case can be easily obtained. Recall the notion of uniform continuity from Analysis 1.

Let $I \subset \mathbb{R}$ and f be a real-valued function on I. We say that f is uniformly continuous on I if

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, y \in I : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Also remind, that a continuous function on [a, b] is always uniformly continuous.

Notes

Definition 1: Define

$$p_{kn}(x) = \binom{n}{k} x^k (1 - x)^{n-k}, \ \forall\, n \in \mathbb{N} \text{ and } 0 \le k \le n.$$

Note that, $p_{kn}(x)$ becomes a probability mass function of binomial distribution with probability of successful trial equal to x if $x \in [0, 1]$.

Lemma: (a) $\forall\, n \in \mathbb{N} : \sum_{k=0}^{n} p_{kn}(x) = 1$, (b) $\forall\, n \in \mathbb{N} : \sum_{k=0}^{n} k_{pkn}(x) = nx$, (c) $\forall\, n \in \mathbb{N} : \sum_{k=0}^{n} (k - nx)^2$ $p_{kn}(x) = nx(1 - x)$.

Proof: (a) If $x \in [0, 1]$ the equality follows from normalisation of probability distribution. In general

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

therefore

$$\sum_{k=0}^{n} p_{kn}(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} = (x + 1 - x)^n = 1.$$

(b) If $x \in [0, 1]$, we can define a random variable $Y_{n,x}$ the number of heads observed on unfair x-coin tossed n-times. Then

$$\mathbb{P}(Y_{n,x} = k) = \binom{n}{k} x^k (1 - x)^{n-k} = p_{kn}(x).$$

Moreover, we find the following relation with (b)

$$\mathbb{E}[Y_{n,x}] = \sum_{k=0}^{n} k_{p_{kn}}(x) = nx.$$

In general

$$k\binom{n}{k} = k\frac{n!}{k!(n-k)!} = n\frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} = n\binom{n-1}{k-1},$$

so

$$\sum_{k=0}^{n} k_{p_{kn}}(x) = \sum_{k=0}^{n} k\binom{n}{k} x^k (1 - x)^{n-k} = \sum_{k=0}^{n} n\binom{n-1}{k-1} x^k (1 - x)^{n-k} =$$

$$= \sum_{k=1}^{n} n\binom{n-1}{k-1} x^k (1 - x)^{n-k} = nx\sum_{k=1}^{n} \binom{n-1}{k-1} x^{k-1} (1 - x)^{n-k} =$$

$$= nx\sum_{i=0}^{n-1} \binom{n-1}{i} x^i (1 - x)^{n-1-i} = nx(x + 1 - x)^{n-1} = nx.$$

(c) If $x \in [0, 1]$, we can rewrite the formula as

$$\mathrm{Var}(Y_{n,x}) = \sum_{k=0}^{n} (k - nx)^2 p_{kn}(x) = nx\,(1 - x).$$

In general

$$k(k-1)\binom{n}{k} = n(n-1)\frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} = n(n-1)\binom{n-2}{k-2},$$

$$\sum_{k=0}^{n} k(k-1)p_{kn}(x) = \sum_{k=0}^{n} k(k-1)\binom{n}{k} x^k(1-x)^{n-k}$$

$$= \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2} x^k(1-x)^{n-k}$$

$$= n(n-1)x^2 \sum_{k=2}^{n}\binom{n-2}{k-2} x^{k-2}(1-x)^{n-k}$$

$$= n(n-1)x^2 \sum_{i=0}^{n-2}\binom{n-2}{i} x^i(1-x)^{n-2-i} = n(n-1)x^2.$$

Hence

$$\sum_{k=0}^{n} (k-nx)^2 p_{kn}(x) = \sum_{k=0}^{n} (k^2 - 2knx + n^2x^2)p_{kn}(x)$$

$$= \sum_{k=0}^{n} (k(k-1) + k - 2knx + n^2x^2)p_{kn}(x)$$

$$= n(n-1)x^2 + nx - 2nxnx - n^2x$$

$$n^2x - nx^2 + nx - 2n^2x^2 + n^2x^2 = nx(1-x).$$

Definition 2: For any f: [0, 1] $\to \mathbb{R}$ define its Bernstein polynomials $B_n^f(x)$ such that

$$B_n^f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) p_{kn}(x).$$

Theorem 2: Weierstrass Approximation Theorem, special case

Let f be a real-valued function on [0, 1]. If f is continuous then $B_n^f \to f$ uniformly.

Proof: We want

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\ \ \forall x \in [0,1] : |B_n^f(x) - f(x)| < \varepsilon.$$

Let ε > 0 be given. Since f is uniformly continuous on [0, 1]

$$\exists 5 > 0\ \ \forall x, y \in [0,1] : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon/2.$$

Using this fact we can estimate

$$|B_n^f(x) - f(x)| = \left| \sum_{k=0}^{n} f\left(\frac{k}{n}\right) p_{kn}(x) - f(x) \underbrace{\sum_{k=0}^{n} p_{kn}(x)} \right|$$

$$= \left|\sum_{k=0}^{n} (f(k/n) - f(x))p_{kn}(x)\right| \leq \sum_{k=0}^{n} \left|(f(k/n) - f(x)\right| p_{kn}(x)$$

$$= \sum_{k:\left|\frac{k}{n}-x\right|<\delta}^{n} \left|(f(k/n) - f(x)\right| p_{kn}(x) + \sum_{k:\left|\frac{k}{n}-x\right|\geq\delta}^{n} \left|(f(k/n) - f(x)\right| p_{kn}(x)$$

$$< \frac{\varepsilon}{2} \sum_{k:\left|\frac{k}{n}-x\right|<\delta}^{n} p_{kn}(x) + 2\|f\|_{sup} \sum_{k:\left|\frac{k}{n}-x\right|\geq\delta}^{n} 1 \cdot p_{kn}(x).$$

We used estimate

$$|f(k/n) - f(x)| \leq |f(k/n)| + |f(x)| \leq 2\,\|f\|_{sup}.$$

Now observe that in the second sum we have the following condition on k

Notes

$$|k/n - x| \geq \delta \Rightarrow \left(\frac{k-nx}{n}\right)^2 \geq \delta^2 \Rightarrow \frac{(k-nx)^2}{n^2\delta^2} \geq 1.$$

By using this remark in the second sum and by increasing number of summants in the first sum we get

$$\frac{\varepsilon}{2}\sum_{k=0}^{n} p_{kn}(x) + 2\|f\|_{\sup} \sum_{k:\left|\frac{k}{n}-x\right|\geq\delta}^{n} 1\cdot p_{kn}(x)$$

$$\leq \frac{\varepsilon}{2}\cdot 1 + \frac{2\|f\|_{\sup}}{n^2\delta^2}\sum_{k:\left|\frac{k}{n}-x\right|\geq\delta}^{n}(k-nx)^2 p_{kn}(x)$$

$$\leq \frac{\varepsilon}{2} + \frac{2\|f\|_{\sup}}{n^2\delta^2}\sum_{k=0}^{n}(k-nx)^2 p_{kn}(x) = \frac{\varepsilon}{2} + \frac{2\|f\|_{\sup}}{n^2\delta^2} n \underbrace{x(1-x)}_{\leq 1} \leq \frac{\varepsilon}{2} + \frac{2\|f\|_{\sup}}{n\delta^2}.$$

Let N be such that $\frac{2\|f\|_{\sup}}{N\delta^2} < \frac{\varepsilon}{2}$. Then $\forall n \geq N$ and $\forall x \in [0, 1]$

$$\left|B_n^f(x) - f(x)\right| < \frac{\varepsilon}{2} + \frac{2\|f\|_{\sup}}{n\delta^2} \leq \frac{\varepsilon}{2} + \frac{2\|f\|_{\sup}}{N\delta^2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Proof: In this proof we shrink our function f to [0, 1], where it can be approximated uniformly by Bernstein polynomials and then scale these polynomials from [0, 1] to [a, b] where they will approximate the original function. Define $g : [0, 1] \to \mathbb{R}$, $g(t) = f(x(t))$. Where $x(t) = a + (b - a)t$ for $t \in [0, 1]$. We see that g is continuous since it is a composite of two continuous functions. As $B_n^g \to$ g uniformly. Define $q_n(x) = B_n^g(t(x)) = B_n^g\left(\frac{x-a}{b-a}\right)$. Then

$$\|a_n - f\|_{\sup} = \sup_{x\in[a,b]}|q_n(x) - f(x)| = \sup_{x\in[a,b]}\left|B_n^g(t(x)) - g(t(x))\right|$$

$$= \sup_{t\in[0,1]}\left|B_n^g(t) - g(t)\right| = \left\|B_n^g - g\right\|_{\sup} \to 0, \text{ since } B_n^g \to g \text{ uniformly.}$$

Figure 2.1

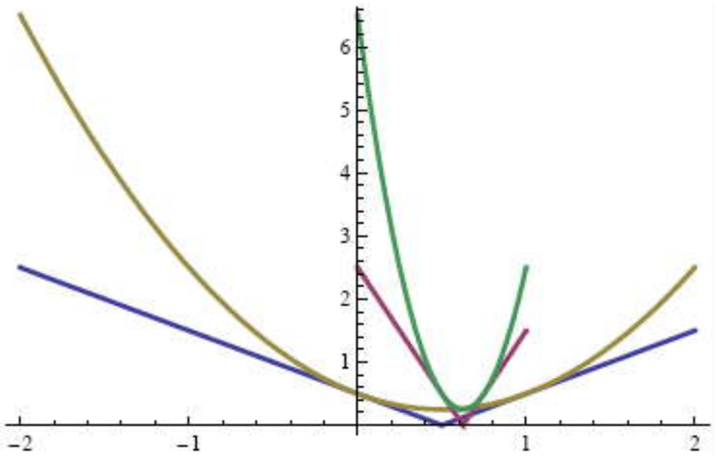
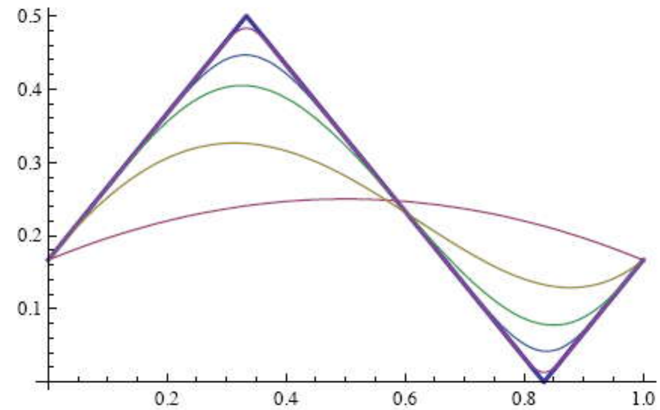


Figure 19.1, the original function f(x) = | x – 1/2 | on [–2, 2], is shrinked to [0, 1], and approximated uniformly by Bernstein polynomials. These polynomials are then scaled to [–2, 2], to approximate uniformly the original function f on [–2, 2].

Notes

Figure 2.2: Approximation of f(x) = | − |x−1/3| + 1/2| on [0,1] by B2(x),BI(x), B (5(x),B10(x) and Bi500(x)



2.2 Fourier Series

Firstly, let us look at some definitions. We denote the set of all Riemann integrable functions on [a, b] by $\mathfrak{R}[a, b]$.

Definition 3: For any f, g ∈ $\mathfrak{R}[a, b]$ define inner product of f and g ⟨.,.⟩ such that

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Recall from Linear Algebra that the inner product space is finite-dimensional vector space V equipped with mapping $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ (or $\mathbb{R}$), satisfying these three properties:

1. $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$
2. $\langle v, u\rangle = \overline{\langle v, u\rangle}$
3. $\langle u, u\rangle \in \mathbb{R}$ and $\langle u, u\rangle \geq 0$ with equality ⇔ u = 0

We see that our inner product does not satisfy all three properties since ⟨f, f⟩ = 0 ⇔ f(x) = 0 does not hold. It suffices to take $f(x) = \begin{cases} 1 & \text{for } x \in [0,1) \\ 0 & \text{for } x = 1 \end{cases}$.

Definition 4: We define the two-norm $\|\cdot\|_2$ on f ∈ $\mathfrak{R}[a, b]$ such that

$$\|f\|_2 = \sqrt{\langle f, f\rangle} = \sqrt{\int_a^b f(x)^2 dx}.$$

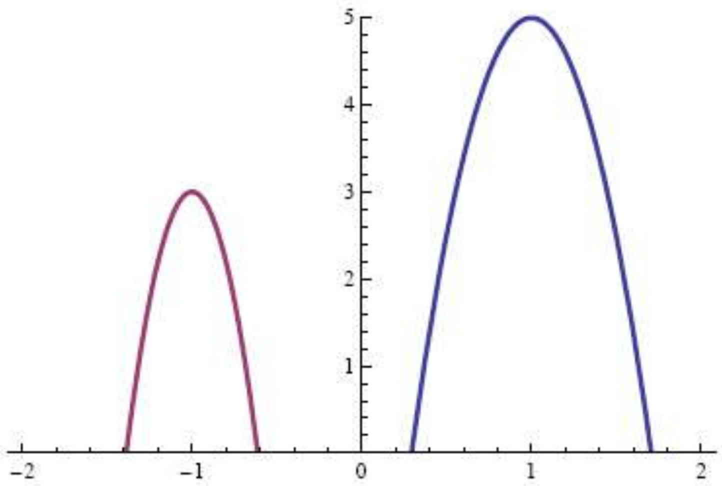
Definition 5: A collection of Riemman integrable functions $\{\phi_n\}_{n=1}^{\infty}$ on [a, b] is called an orthogonal system if

$$\langle \phi_m, \phi_n\rangle = \int_a^b \phi_m(x)\phi_n(x)dx = 0, \qquad \forall m \neq n.$$

If in addition $\forall n \in \mathbb{N} : \|\phi_n\|_2 = 1$ we call $\{\phi_n\}_{n=1}^{\infty}$ an orthonormal system.

Example: Consider two continuous functions as on Figure 17.3. We have fg = 0, hence ⟨f,g⟩ = 0. Therefore, they are orthogonal.

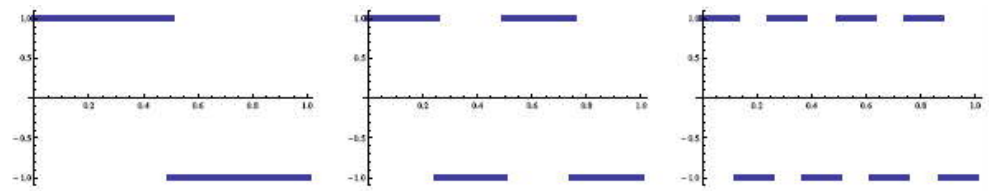
Figure 2.3: Two Orthogonal Continuous Functions on [−2,2]



Example: A collection $\Xi = \left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(nx)\right\}, n \in \mathbb{N}$, is called the trigonometrical orthonormal system on $[-\pi, \pi]$, since $\|f\|_2 = 1$ for all $f \in \Xi$ and $\langle f, g\rangle = 0$ for all $f \neq g$.

Example: For another example of orthonormal system see Figure 19.4.

Figure 19.4



Orthonormal system of functions ϕ_n: [0, 1] → {–1, 1}. Each ϕ_n divides interval [0, 1] into $1/2^n$ subintervals. $\int_0^1 \phi_n(x)^2 dx = 1$ and $\int_0^1 \phi_n(x)\phi_m(x)dx = 0$ if $n \neq m$

Definition 6: Let $\{\phi_n\}_{n=1}^{\infty}$ be an o.n.s. on [a,b] and $f \in \mathfrak{R}[a, b]$. We define Fourier coefficients of f w.r.t. $\{\phi_n\}_{n=1}^{\infty}$ as

$$a_n = \langle f, \phi_n\rangle = \int_a^b f(x)\phi(x)dx, \qquad n \in \mathbb{N}.$$

$\sum_{n=1}^{\infty} a_n\phi_n$ is called the Fourier series of f w.r.t. $\{\phi_n\}_{n=1}^{\infty}$.

Notes

1. $\sum_{n=1}^{\infty} a_n\phi_n$ does not necessarily converge.
2. f(x) is not necessarily equal to its Fourier series.

Example: Let f(x) = x on [0, 1] and $\{\phi_n\}_{n=1}^{\infty}$. We get Fourier coefficients $a_n = \langle x, \phi_n \rangle = \int_0^1 x\phi(x)dx = -\left(\frac{1}{2^n}\right)^2 \frac{2^n}{2} = -\frac{1}{2^{n+1}}$. Therefore, we can compute Fourier series for f(x) = x which is $-\sum_{n=1}^{\infty} \frac{1}{2^{n+1}} \phi_n(x)$.

Self Assessment

Fill in the blanks:

1. For any f: [0, 1] → $\mathbb{R}$ define its .... $B_n^f(x)$ such that

$$B_n^f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) p_{kn}(x).$$

2. Let f be a real-valued function on [0, 1]. If f is continuous then ....
3. Define the two—norm $\|\cdot\|_2$ on f ∈ $\Re$[a, b] such that ....
4. A collection of Riemman integrable functions $\{\phi_n\}_{n=1}^{\infty}$ on [a, b] is called an ....

   $\langle \phi_m, \phi_n \rangle = \int_a^b \phi_m(x)\phi_n(x)dx = 0, \quad \forall m \neq n.$

2.3 Summary

- Let I ⊂ $\mathbb{R}$ and f be a real-valued function on I. We say that f is uniformly continuous on I if

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in I : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

  Also remind, that a continuous function on [a, b] is always uniformly continuous.

- $\forall n \in \mathbb{N} : \sum_{k=0}^{n} p_{kn}(x) = 1$, (b) $\forall n \in \mathbb{N} : \sum_{k=0}^{n} k_{pkn}(x) = nx$, (c) $\forall n \in \mathbb{N} : \sum_{k=0}^{n} (k - nx)^2 p_{kn}(x) = nx(1 - x)$.

  If x ∈ [0, 1], we can define a random variable $Y_{n,x}$ the number of heads observed on unfair x-coin tossed n-times. Then

$$\mathbb{P}(Y_{n,x} = k) = \binom{n}{k} x^k (1 - x)^{n-k} = p_{kn}(x).$$

  Moreover, we find the following relation with (b)

$$\mathbb{E}[Y_{n,x}] = \sum_{k=0}^{n} k_{P_{kn}}(x) = nx.$$

- For any f, g ∈ $\Re$[a, b] define inner product of f and g ⟨.,.⟩ such that

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

- Recall from Linear Algebra that the inner product space is finite-dimensional vector space V equipped with mapping ⟨·,·⟩ : V × V→ $\mathbb{C}$ (or $\mathbb{R}$).

2.4 Keyword

Fourier Series: For any f, g ∈ $\Re$[a, b] define inner product of f and g ⟨.,.⟩ such that

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

Recall from Linear Algebra that the inner product space is finite-dimensional vector space V equipped with mapping $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ (or $\mathbb{R}$), satisfying these three properties.

2.5 Review Questions

1. The Arzela-Ascoli Theorem is the key to the following result: A subset F of C(X) is compact if and only if it is closed, bounded, and equicontinuous. Prove this.

2. You can think of Rn as (real-valued) C(X) where X is a set containing n points, and the metric on X is the discrete metric (the distance between any two different points is 1). The metric thus induced on Rn is equivalent to, but (unless n = 1) not the same as, the Euclidean one, and a subset of Rn is bounded in the usual Euclidean way if and only if it is bounded in this C(X). Show that every bounded subset of this C(X) is equicontinuous, thus establishing the Bolzano-Weierstrass theorem as a generalization of the Arzela-Ascoli Theorem.

3. Let f(x) = x on [0, 1] and let $\{\phi_n\}_{n=1}^{\infty}$ be as in Ex. 2.3. We get Fourier coefficients $a_n = \langle x, \phi_n \rangle$ $= \int_0^1 x\phi(x)dx = -\left(\frac{1}{2^n}\right)^2 \frac{2^n}{2} = -\frac{1}{2^{n+1}}$, (computation of the integral is left as exercise). Therefore, we can compute Fourier series for f(x) = x which is $-\sum_{n=1}^{\infty} \frac{1}{2^{n+1}} \phi_n(x)$.

Answer: Self Assessment

1. Bernstein polynomials
2. $B_n^f \rightarrow f$ uniformly
3. $\|f\|_2 = \sqrt{\langle f, f \rangle} = \sqrt{\int_a^b f(x)^2 dx}$.
4. Orthogonal System

2.6 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 3: The Riemann Integration

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define the Riemann Integral of a function
- Derive the conditions of Integrability and determine the class of functions which are always integrable
- Discuss the algebra of integrable functions
- Compute the integral as a limit of a sum

Introduction

You are quite familiar with the words 'differentiation' and distinguishing 'integration'. You know that in ordinary language, differentiation refers to separating things while integration means putting things together. In Mathematics, particularly in Calculus and Analysis, differentiation and integration are considered as some kind of operations on functions. You have used these operations in our study of Calculus.

There are essentially two ways of describing the operation of integration. One way is to view it as the inverse operation of differentiation. The other way is to treat it as some sort of limit of a sum.

The first view gives rise to an integral which is the result of reversing the process of differentiation. This is the view which was generally considered during the eighteenth century.

Accordingly, the method is to obtain, from a given function, another function which has the first function as its derivative. This second function, if it be obtained, is called the indefinite integral of the first function. This is also called the 'primitive' or anti-derivative of the first function. Thus, the integral of a function f(x) is obtained by finding an anti-derivative or primitive function F(x) show that $F'(x) = f(x)$. The indefinite integral of f(x), is symbolized by the notation $\int f(x)\,dx$.

The second view is related to the limiting process. It gives rise to an integral which is the limit of all the values of a function in an interval. This is the integral of a function f(x) over an interval [a,b,]. It is called the definite integral and is denoted by

$$\int_a^b f(x)dx$$

The definite integral is a number since geometrically it corresponds an area of a region enclosed by the graph of a function.

Although both the notions of integration are closely related, yet, you will see later, the definite integral turns out to be a mare fundamental concept. In fact, it is the starting point for some important generalizations like the double integrals, triple integrals, line integrals etc., which you may study on Advanced Calculus.

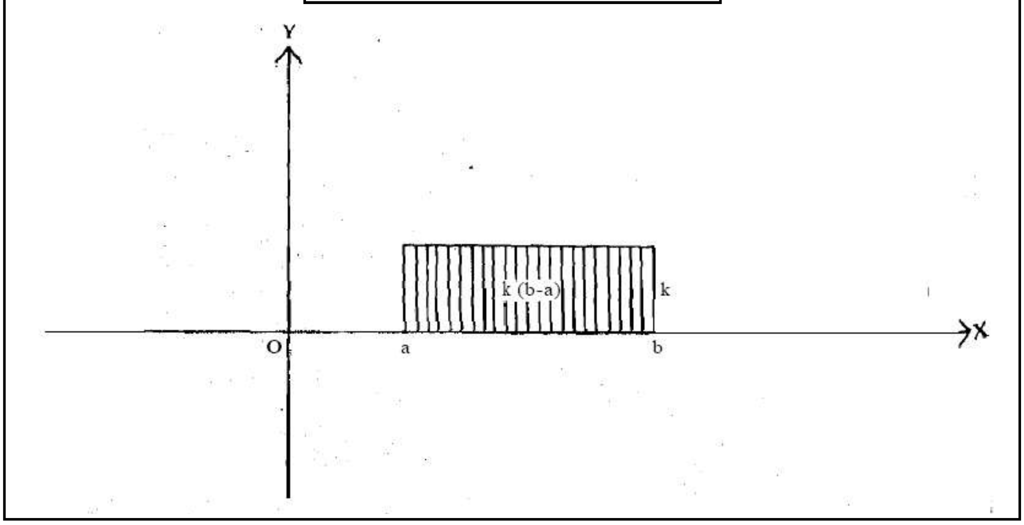
The integral in the anti-derivative sense was given by Neyrtan. This notion was found to be adequate so long as the functions to be integrated were continuous. But in the early 19th century, Fourier brought to light the need for making integration meaningful for the functions that are not continuous. He came across such functions in applied problems. Cauchy formulated rigorous definition of the integral of a function. He essentially provided a general theory of integration but only for continuous functions. Cauchy's theory of Integration for continuous functions is sufficient for piece-wise continuous functions as well as for the functions having isolated discontinuities. However, it was G.B.F. Riemann [1826-1866] a German mathematician who extended Cauchy's integral to the discontinuous functions also. Riernann answered the question "what is the meaning of $\int f(x)\,dx$?"

The concept of definite integral was given by Riemann in the middle of the nineteenth century. That is why, it is called Riemann Integral. Towards the end of 19th Century, T.J. Stieltjes [1856-1894] of Holland, introduced a broader concept of integration replacing certain linear functions used in Riemann Integral by functions of more general forms. In the beginning of this century, the notion of the measure of a set of real numbers paved the way to the foundation of modern theory of Lebesgue Integral by an eminent French Mathematician H. Lebesgue [1875-1941], a beautiful generalisation of Riemann Integral which you may study in some advanced courses of Mathematics. In this unit, the Riemann Integral will be defined without bringing in the idea of differentiation. As you have been go through the usual connection between the Integration and Differentiation. Just by applying the definition, it is not always easy to test the integrability of a function. Therefore, condition of integrability will be derived with the help of which it becomes easier to discuss the integrability of functions. Then just as in the case of continuity and derivability, we will also consider algebra of integrable functions. Finally, in this unit, second definition of integral as the limit of a sum will be given to you and you will be shown the equivalence of the two definitions.

3.1 Riemann Integration

The study of the integral began with the geometrical consideration of calculating areas of plane figures. You know that the well-known formula for computing the area of a rectangle is equal to the product of the length and breadth of the rectangle. The question that arises from this formula is that of finding the correct modification of this formula which we can apply to other plane figures. To do so, consider a function defined on a closed interval [a,b] of the real line, which assumes a constant value $K \geq 0$ throughout the interval. The graph of such a function gives rise to a rectangular region bounded by the X-axis and the ordinates $x = a$, $x = b$ as shown in the Figure 20.1.

Figure 3.1



Obviously, the area enclosed is k (b-a). NOW, suppose that]a, b[is further divided into smaller intervals by inserting points of division, say

$$a = x, \leq x_1 \leq x_2 \leq x_3 \leq x_4 = b,$$

and the function f is defined so as to take a constant value at each of the resulting sub-intervals i.e.,

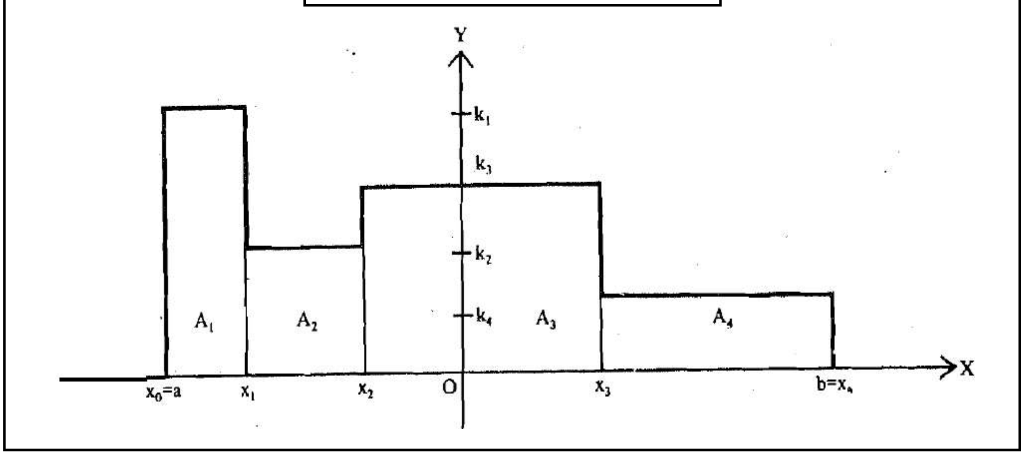
$$f(x) = \begin{cases} k_1, \text{ if } x \in [x_0, x_1[\\ k_2, \text{ if } x \in [x_1, x_2[\\ k_3, \text{ if } x \in [x_2, x_3[\\ k_4, \text{ if } x \in [x_1, x_4] \end{cases}$$

Further, suppose that d_i = length of the ith interval $]x_i, x_{i-1}[$ i.e.,

$d_1 = |x_1 - x_0|$, $d_2 = |x_2 - x_1|$, $d_3 = |x_3 - x_2|$, and $d_4 = |x_4 - x_3|$.

Then, we get four rectangular regions and the area of each region is A, = k_1d_1, A, = k_2d_2, A, = k_3,d_3, and A, = k_4,d_4, as shown in Figure 20.2.

Figure 3.2



The total area enclosed by the graph of the function, X-axis and the ordinates x=a, x=b is equal to the sum of these areas i.e.

$$\text{Area} = A_1 + A_2 + A_3 + A_4$$

$$= k_1d_1 + k_2d_2 + k_3d_3 + k_4d_4.$$

Note that in the last equation, we have generalized the notion of area. In other words, we are able to compute the area of a region which is not of rectangular shape. How did we get it? By breaking up the region into a series of non-overlapping rectangles which include the totality of the figure and summing up their respective areas. This is simply a slight obstraction of the same process which is used in Geometry.

Since the graph of the function in figure 20.2 consists of 4 different steps, such a function, is called a step function. What we have obtained is the area of a region bounded by

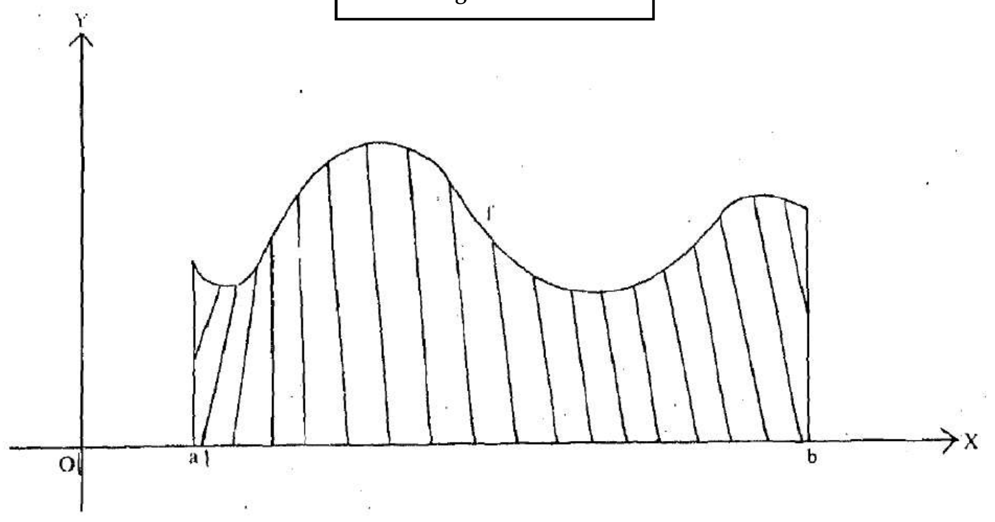
1. a non-negative step function
2. the vertical lines defined by x=a and x=b
3. the X-axis.

This area is just the sum of the areas of a finite number of f non-overlapping rectangles resulting from the graph of the given function. The area is nothing but a real number.

Now suppose that the graph of a given function is as shown in the Figure 20.3.

Does it make any sense to obtain the area of the region under the graph off? If so, how can we compute its value? To answer this question, we introduce the notion of the integral of a function as given by Riemann.

Figure 3.3



To introduce the notion of an integral of a function, we will require such a real number which results for applying the function and which represents the area of the region bounded by the graph off, the vertical lines x=a, x=b and the X-axis. This can be achieved by approximating the given function by suitable step functions. The area of the region will, then, be approximated by the areas enclosed by these step functions, which in turn are obtained as sum of the areas of non-overlapping rectangles as we have computed for the Figure 20.2. This is precisely the idea behind the formal treatment of the integral which we discuss here. First, we introduce some terminology and basic notions which will be used throughout the discussion.

Let f be a real function defined and bounded on a closed interval [a,b].

Recall that a real function f is said to be bounded if the range of f is a bounded subset of R, that is, if there exist numbers m and M such that $m \le f(x) \le M$ for each $x \in [a,b]$. M is an upper bound and m is a lower bound of f in [a,b]. You also know that when f is bounded, its supremum and infimum exist. We introduce the concept of a partition of [a,b] and other related definitions:

Definition 1: Partition

Let [a,b] be a given interval. By a partition P of [a,b] we mean a finite set of points $\{x_0, x_l, ..., x,\}$, where

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

We write $\Delta x_i = x_i - x_{i-1}$, (i=l, 2, ..., n). So Δx_i is the length of the ith sub-interval given by the partition P.

Definition 2: Norm of a Partition

Norm of a partition P, denoted by $|P|$, is defined by $|P| = \max_{t \le i \le n} Ax,$. Namely, the norm of P is the length of largest sub-interval of [a, b] induced by P. Norm of P is also denoted by $\mu(P)$.

There is a one-to-one correspondence between the partitions of [a,b] and finite subsets of]a, b[. This induces a partial ordering on the set of partitions of [a,b]. So, we have the following definition.

Definltion 3: Refinement of a Partition

Let P, and P, be two partitions of [a,b]. We say that P, is finer than P, or P_2, refines P, or P_2 is a refinement of P_1 if $P_1 \subset P_2$, that is, every point of P_1 is a point of P,.

You may note that, if P, and P_2, are any two partitions of [a,b], then P, $\cup$ P_2 is a common refinement of P, and P_2. For example, if $P_1 = \left\{0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1\right\}$ and $P_2 = \left\{0, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1\right\}$ are partitions of [0, 1], then P_2 is a refinement of P_1 and $P_1 \cup P_2 = \left\{0, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1\right\}$ is their common refinement.

We now introduce the notions of upper sums and lower sums of a bounded function f on an interval [a, b], as given by Darboux. These are sometimes referred to as Darboux Sums.

Definition 4: Upper and Lower Sums

Let f: $[a,b] \to R$ be a bounded function, and let P = $(x_0, X_1 ... x,)$ be a partition of [a,b]. For i = 1, 2 ....., n, let M_i and m_i be defined by

$$M_i = \text{lub } (f(x) : x_{i-1} \le x \le x_i)$$

$$m_i = \text{glb } (f(x) : x_{i-1} \le x \le x_i)$$

i.e. M_i and m_i be the supremum and infimum of f in the sub-interval $[x_{i-1}, x_i]$.

Then, the upper (Riemann) sum of f corresponding to the partition P, denoted by U (P,f), is defined by

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

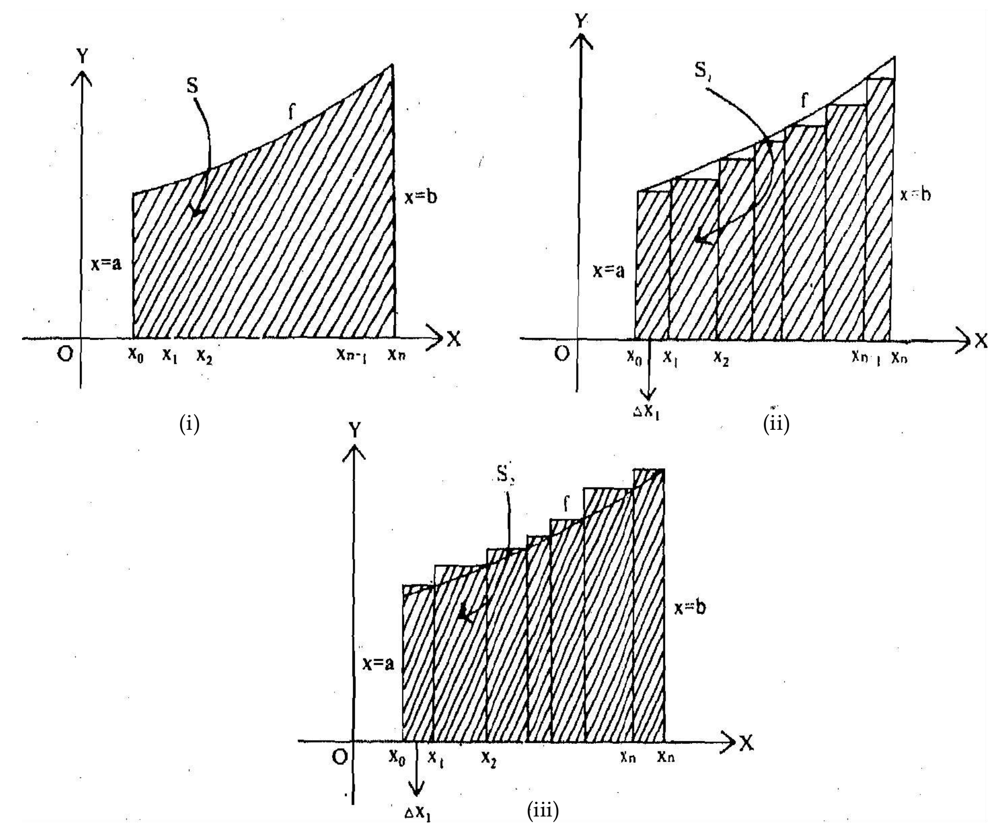
Notes

The lower (Riemann) sum off corresponding to the partition P, denoted by L(P, f), is defined by

$$U(P,f) = \sum_{i=2}^{n} M_i \, \Delta x_i.$$

Before we pass on to the definition of upper and lower integrals, it is good for you to have the geometrical meaning of the upper and lower sums and to visualize the above definitions pictorially. You would, then, have a feeling for what is going on, and why such definitions are made. Refer to Figure 20.4.

Figure 3.4



In figure 20.4(i) the graph off: [a,b] $\rightarrow$ R is drawn. The partition P = $\{x_0, x_1, ... x_n\}$ divides the interval [a,b] into sub-intervals $[x_0, x_1]$, $[x_1, x_2]$, ... $[x_{n-1}, x_n]$. Consider the area S under the graph off. In the first sub-interval $[x_0, x_1]$, m_1 is the g.l.b. of the set of values f(x) for x in $[x_0, x_1]$. Thus $m_1 \Delta x_1$ is the area of the small rectangle with sides m_1 and Δx_1 as shown in the figure 20.4(ii). Similarly $m_2 \Delta x_2$ ... m_n A x_n are areas of such small rectangles and $\sum_{i=1}^{n} m,$ Ax, i.e. lower sum L (P,f) is the area S_2 which is the sum of areas of such small rectangles. The area S_1 is less than the area S under the graph of f.

In the same way $M_I \Delta X_1$ I is the area of the Large rectangle with sides M_I and ΔX_1 and $\sum_{i=1}^{n} M_i - 1x$, i.e., the upper sum U(P, f) is the area S_2 which is the sum of areas or such large rectangles as shown in Figure 20.4(iii). The area S_2 is more than the area S under the graph off. It is intuitively clear that if the points in the partition P are increased, the areas S_1 and S_2 approach the area S.

We claim that the sets of upper and lower sums corresponding to different partitions of [a,b] are bounded. Indeed, let m and M be the infimum and supremum of f in [a.b].

Then $m \le m_i \le M_i \le M$ and so

$$m \,\Delta x_1 \le m_i \,\Delta x_i \le M_i \,\Delta x_i \le M \,\Delta x_i$$

Putting i = 1, 2, ... n and adding, we get

$$m \sum_{i=1}^{n} \Delta x, \le L(P,f) \le U(P,f) \le M \sum_{i=1}^{n} A\ x,.$$

$$\sum_{i=1}^{n} \Delta x_i = \sum_{i=1}^{n} (x_i - x_{i-1}) = x_n - x_0 = b - a$$

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

For every partition P, there is a lower sum and there is an upper sum. The above inequalities show that the set of lower sums and the set of upper sums are bounded, so that their supremum and infimum exist. In particular, the set of upper sums have an infimum and the set of lower sums have a supremum. This leads us to concepts of upper and lower in tegrals as given by Riemann and popularly known as Upper and Lower Riemann Integrals.

Definition 5: Upper and Lower Riemann Integral

Let f: [a,b] $\to$ R be a bounded function. The infimum or the greatest lower bound of, the set of all upper sums is called the upper (Riemann) integral o f f on [a,b] and is denoted by,

$$\int_a^{\bar{h}} f(x)dx.$$

i.e.

$$\int_a^{\bar{b}} f(x)dx. = \text{g.l.b. } \{U(P,f): P \text{ is a partition of } [a,b]\}.$$

The supremum or the least upper bound of the set of all lower sums is called the lower (Riemann) integral of f on [a,b] and is denoted by

$$\int_a^{b} f(x)dx$$

i.e.

$$\int_a^{b} f(x)dx = \text{l.u.b } \{L(P,f): P \text{ is a partition of } [a,b]\}.$$

Now we consider some examples where we calculate upper and lower integrals.

Example: Calculate the upper and lower integrals of the function f defined in [a, b] as follows:

$$f(x) = \begin{vmatrix} 1 \text{ when x is rational} \\ 0 \text{ when x is irrational} \end{vmatrix}$$

Solution: Let P = {x_0, x_1 … x_n} be any partition of [a,b]. Let M_i and m_1 be respectively the sup. f and inf. f in [x_{i-1}, x,]. You know that every interval contains infinitely many rational as well as irrational numbers. Therefore, m_i = 0 and M, = I for i = 1, 2 ... n. Let us find U(P,f) and L(P,f).

$$U(P,f) = \sum_{i=1}^{n} M_i \;\Delta\, x_i = \sum_{i=1}^{n} \Delta\, x_i = b - a$$

$$U(P,f) = \sum_{i=1}^{n} m_i \;\Delta\, x_i = 0$$

Therefore U(P,f) = b – a and L(P,f) = 0 for every, partition P of [a,b]. Hence

$\int_a^b f(x)\,dx$ = g.l.b. {U(P,f): P is a partition of [a,b]]

= g.l.b. {b – a] = b – a.

$\int_a^b f(x)dx$ = l.u.b. {L(P,f): P is a partition of [a,b]]

= l.u.b. {0] = 0.

Example: Let f be a constant function defined in [a,b]. Let f(x) = k ∀ x ∈ [a,b]. Find the upper and lower integrals of f.

Solution: With the same notation as in example 1, M_i = k and m_i = k ∀ i.

$$U(P,f) = \sum_{i=1}^{n} M_i \; A\, x_i = \sum_{i=1}^{n} A\, x_i = k(b-a)$$

$$\text{and } L(P,f) = \sum_{i=1}^{n} m_i \; A\, x_i = \sum_{i=1}^{n} A\, xi = k(b-a)$$

Therefore U(P,f) = k(b – a) and L(P,f) = k (b – a) for every partition P of [a,b].

Consequently $\int_a^{\bar{b}} f(x)dx = k(b-a)$ and $\int_a^b f(x)dx = k(b-a)$

Now try the following exercise.

Exercise

Find the upper and lower Riemann integrals of the function f defined in [a,b] as follows:

$$f(x) = \begin{cases} \;\;1 \text{ when x is ratinal} \\ -1 \text{ when x is irrational} \end{cases}$$

You have seen that sometimes the upper and lower integrals are equal (as in Example) and sometimes they are not equal (as in Example). Whenever they are equal, the function is said to be integrable. So integrability is defined as follows:

Definition 6: Riemann Integral

Let f: [a,b] – R be a bounded function. The function f is said to be Riemann integrable or simply integrable or R-integrable over [a,b] if $\int_a^b f(x)\,dx = \int_{-}^{b} f(x)\,dx$ and iff is Riemann integrable, we denote the common value by $\int_a^b f(x)\,dx.$ This is called the Riemann integral a r simply the integral off on [a, b].

Notes

Example: Show at the function f considered in example is not Riemann integrable.

Solution: As shown in above example, $\int_a^b f(x)\,dx = b-a$ and $\int_a^b f(x)\,dx = 0$ and so $\int_a^{\bar{b}} f(x)\,dx \neq \int^b f(x)\,dx$ and consequently f is not Riemann integrable.

Example: Show that a constant function is Riemann integrable over [a,b] and find $\int_1^h f(x)\,dx$.

Solution: As proved in above example, $\int_a^{\bar{b}} f(x)\,dx = k(b-a) = \int_1^b f(x)\,dx$

Therefore, f is Riemann integrable on [a,b] and $\int_1^h f(x)\,dx = k(b-a)$.

Theorem 1: If the partition P_2 is a refinement of the partition P, of [a,b], then $L(P_1,f) \leq L(P_2,f)$ and $U(P_2,f) \leq U(P_1,f)$.

Proof: Suppose P_2 contains one point more than P,. Let this extra point be c. Let $P_1 = \{x_0, x_,, \ldots, x_n\}$ and $x_{i-1} < c < x_i$. Let M_i and m_i be respectively the sup. f and inf. f in $[x_{i-1}, x_i]$. Suppose sup. f and inf. f in $[x_{i-1}, c]$ are p, and q_1 and those in $[c, x_i]$ are p_2 and q_2, respectively. Then,

$$L(P_2,f) - L(P_1,f) = q_1(c-x_{i-1}) + q_2(x_i - c) - m_i\Delta x_i$$

$$= (q_1 - m_i)(c - x_{i-1}) + (q_2 - m_i)(x_i - c)$$

$\left(\text{since A } x_i = (x_i - c) + (c - x_{i-1})\right)$

Similarly $U(P_2,f) - U(P_1,f) = (p_1 - M_i)(c - x_{i-1}) + (p_2 - M_i)(x_i - c)$

Now $m_i \leq q_1 \leq p_1 \leq M_i$

$m_i \leq q_2 \leq p_2 \leq M_i$

Therefore

$$L(P_2,f) - L(P_1,f) \geq 0 \text{ and } U(P_2,f) - U(P_1,f) \leq 0$$

Therefore

$$L(P_1,f) - L(P_2,f) \geq 0 \text{ and } U(P_2,f) - U(P_1,I).$$

Is P_2 contains p points more than P_1, then adding these extra points one by one to P_1 and using the above results, the theorem is proved. We can also write the theorem as

$$L(P_1,f) \leq L(P_2,f) \leq U(P_2,f) \leq U(P_1,f)$$

from which it follows that $U(P_2,f) - L(P_2,f) \leq U(P_1,f) - L(P_1,f)$. As an illustration of theorem 1, we consider the following example.

Example: Verify Theorem 1 for the function f(x) = x + 1 defined over [0, 1] and the partition $P_1\left\{0, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{3}{4}, 1\right\}$ and $P_2 = \left\{0, \frac{1}{6}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, 1\right\}$.

Solution: For partition $P_1, n = 5, x_0 = 0, x_1 = \frac{1}{4}, x_2 = \frac{1}{3}, x_1 = \frac{1}{2}, x_4 = \frac{3}{4}, x_4 = 1$ and so $\Delta x_1 = \frac{1}{4}$,

$\Delta x_2 = \frac{1}{12}, \Delta x_3 = \frac{1}{6}, \Delta x_4 = \frac{1}{4}, \Delta x_5 = \frac{1}{4}$.

Further $M_i = f(x_i)$ & $m_i = f(x_{i-1})$ for i = 1, 2, 3, 4, 5 and therefore $M_1 = \frac{5}{4}, M_2 = \frac{4}{3}, M_3 = \frac{3}{2}, M_4 = \frac{7}{4}$,

$M_5 = 2, m_1 = 1, m_2 = \frac{5}{4}, m_3 = \frac{4}{3}, m_4 = \frac{3}{2}, m_5 = \frac{7}{4}$. We have $L(P_1, f) = \sum_{i=1}^{5} m_i \Delta x_i = \frac{25}{18}$ and

$U(P_1, f) = \sum_{i=1}^{5} M_1 A x, = \frac{29}{18}$. Similarly, $L(P_2, f) = \frac{17}{12}$, and $U(P_2, f) = \frac{19}{12}$. Hence $L(P_1, f) \le L(P_2, f)$ and $U(P_2, f) \le U(P_1, f)$.

Theorem 2: $\int_a^b f(x)\,dx \le \int_a^b f(x)\,dx$.

Proof: If P_1 & P_2, are two partitions of [a,b] and $P = P_1 \cup P$, is their common refinement, then using Theorem 1, we have $L(P_1, f) \le L(P, f) \le U(P, f) \le U(P_1, f)$ and

$L(P_2, f) \le L(P, f) \le U(P, f) \le U(P_2, f)$.

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

Keeping P_2 fixed and taking l.u.b. over all P_1, we get

$$\int_a^b f(x)\,dx \le U(P_2, f)$$

Now taking g.l.b. over all P_2, we obtain

$$\int_a^b f(x)\,dx \le \int_a^b f(x)\,dx$$

This proves the result.

In Theorem 1, we have compared the lower and upper sums for a partition P_1 with those for a finer partition P_2. Next theorem, which we state without proof, gives the estimate of the difference of these sums.

Theorem 3: If a refinement P, of P_1 contains p more points and $|f(x)| \le k$, for all $x \in [a, b]$, then

$$L(P_1, f) \le L(P_2, f) \le L(P_1, f) + 2pk\delta,$$

and $U(P_1, f) \ge U(P_2, f) \ge U(P_1, f) - 2p\,k\,\delta$, where δ is the norm of P_1.

This theorem helps us in proving Darboux's theorem which will enable us to derive conditions of integrability. Firstly, we give Darboux's Theorem.

Theorem 4: Darboux's Theorem

If f: [a,b] → R is a bounded function, then to every ∈ > 0, there corresponds δ > 0 such that

(i) $U(P, f) < \int_a^{\bar{b}} f(x)\,dx + \in$

Notes

(ii) $L(P,f) < \int_a^b f(x)\,dx - \in$

for every partition P of [a,b] with $|P| < \delta$.

Proof: We consider (i). As f is bounded, there exists a positive number k such that $|f(x)| \leq k \ \forall \ x \in [a,b]$. As $\overline{\int_a^b} f(x)\,dx$ is the infimum of the set of upper sums, therefore to each $\in > 0$, there is a partition P_1 of [a,b] such that

$$U(P_1, f) < \overline{\int_a^b} f(x)\,dx + \frac{\in}{2}, \qquad (1)$$

Let $P_1 = \{x_0, x_1, ..., x_p\}$ and δ be a positive number such that $2\,k\,(p - 1)\,\delta = \in/2$. Let P be a partition of [a,b] with $|P| < \delta$. Consider the common refinement $P_2 = P \cup P_1$ of P and P_1.

Each partition has the same end points 'a' and 'b'. So P_2 is a refinement of P having at the most (p – 1) more points than P. Consequently, by Theorem 3,

$$\begin{aligned} U(P,f) - 2(p-1)\,k\,\delta &\leq U(P_2,f) \\ &\leq U(P_2,f) \\ &< \overline{\int_a^b} f(x)\,dx + \in/2. \qquad \text{(using (1))} \end{aligned}$$

Thus

$$\begin{aligned} U(P,f) &< \overline{\int_a^b} f(x)\,dx + \frac{\in}{2} + 2(p-1)\,k\,\delta \\ &= \int_a^6 f(x)\,dx + \in, \text{ with } |P| < \delta. \end{aligned}$$

Task Write down the proof of part (ii) of Darboux's Theorem.

As mentioned earlier, Darboux's Theorem immediately leads us to the conditions of integrability. We discuss this in the form of the following theorem:

Theorem 5: Condition of Integrability

First Form: The necessary and sufficient condition for a bounded function f to be integrable over [a,b] is that to every number $\in > 0$ there corresponds $\delta > 0$ such that

$$U(P,f) - L(P,f) < \in, \ \forall \ P \text{ with } |P| < \delta.$$

Proof: We firstly prove the necessity of the condition.

Since the bounded function f is integrable on [a, b], we have

$$\overline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx = \int_a^b f(x)\,dx.$$

Let $\in > 0$ be any number. By Darboux Theorem, there is a number $\delta > 0$ such that

Notes

$$U(P,f) < \int_a^b f(x)\,dx + \in/2 \quad = \int_a^b f(x)dx + \in/2 \;\forall\; P \text{ with } |P| < \delta \tag{2}$$

Also,

$$L(P,f) > \int_a^b f(x)\,dx - \in/2 \quad = \int_a^b f(x)\,dx - \in/2 \tag{3}$$

i.e. $-L(P,f) < \int_{-a}^{b} f(x)\,dx + \in/2 \;\forall\; P$ with $|P| < \delta$

Adding (2) and (3), we get

$U(P,f) - L(P,f) < \in \forall\; P$ with $|P| < \delta.$

Next, we prove that condition is sufficient.

It is given that, for each number $\in > 0$, there is a number $\delta > 0$ such that

$$U(P,f) - L(P,f) < \in, \forall\; P \text{ with } |P| < \delta.$$

Let P be a fixed partition with $|P| < \delta.$ Then

$$L(P,f) \le \int_a^b f(x)dx \le \int_a^{\bar{b}} f(x)dx \le U(P,f).$$

Therefore, $\int_a^b f(x)\,dx - \int_a^b f(x)\,dx \le U(P,f) - L(P,f) < \in.$

Since $\in$ is arbitrary, therefore the non-negative number

$$\int_a^{\bar{b}} f(x)\,dx - \int_a^b f(x)\,dx$$

is less than every positive number. Hence it must be equal to zero that is $\int_a^{\bar{b}} f(x)\,dx = \int_a^b f(x)\,dx$ and consequently f is integrable over [a,b].

Second Form: The necessary and sufficient condition for a bounded function f to be integrable over [a,b] is that to every number $\in > 0$, there corresponds a partition P of [a,b] such that

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

3.2 Riemann Integrable Functions

As we derived the necessary and sufficient conditions for the integrability of a function, we can now decide whether a function is Riemann integrable without finding the upper and lower integrals of the function. By using the sufficient part of the conditions, we test the integrability of the functions. Here we discuss functions which are always integrable. We will show that a continuous function is always Riemann integrable. The integrability is not affected even when there are finites number of points of discontinuity or the set of points of discontinuity of the function has a finite number of limit points. It will also be shown that a monotonic function is also always Riemann integrable.

We shall denote by R(a,b), the family of all Riemann integrable functions on [a,b]. First we discuss results pertaining to continuous functions in the form of the following theorems.

Theorem 6: If f: [a, bJ $\rightarrow$ R is a continuous function, then f is integrable over [a, b], that is $f \in R(a,b)$.

Proof: If f is a continuous function on [a,b] then f is bounded and is also uniformly continuous.

To show that $f \in R$ [a,b] you have to show that to each number $\in > 0$, there is a partition P for which

$$U(P,f) - L(P,f) < \in$$

Let $\in > 0$ be given. Since f is uniformly continuous on [a,b], there is a number $\delta > 0$ such that $|f(x) - f(y)| < \frac{E}{b-a}$ whenever $|x - y| < \delta$. Let P be any partition of [a,b] with $|P| < \delta$.

We show that, for such a partition P, U (P, f) – L(P, f) < $\in$.

Now, $$U(P,f) - L(P,f) = \sum_{i=1}^{n} M_i \, \Delta x_i - \sum_{i=1}^{n} m_i \, \Delta x_i$$

$$= \sum_{i=1}^{n} (M_i - m_1)\Delta x_i, \quad (4)$$

where $\Delta x_i = x_i - x_{i-1}$, and $M_i = \sup \{f(x) | x_{i-1} \le x \le x_1\} = f(\xi_1)$ (say), for same $\xi_1 \in [x_{i-1}, x_1]$. Such a ξ_i exists because a continuous function f attains its bounds on $[x_{i-1} - x_1]$.

Similarly, $m_i = \inf \{f(x) | x_{i-1} \le x \le x_i\} = f(\eta_i)$ (say), for some $\eta_i \in [x_{i-1}, x_i]$. Hence

$M_i - m_i = f(\xi_i) - f(\eta_i) \le |f(\xi_i) - f(\eta_i)| < \in / b - a$, for all i,

since $|\xi_i - \eta_i| \le A \, x_i < \delta$. Substituting in (4) we obtain

$$U(P,f) - L(P,f) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i$$

$$< \frac{\in}{b-a}\left(\sum_{i=1}^{n} \Delta x_i\right)$$

$$\frac{E}{b-a}(b-a) = \in.$$

Thus, every continuous function is Riemann integrable,

But as remarked earlier, even when there are discontinuous of the function, it is integrable. This is given in the next two concepts which we state without proof.

Theorem 7: Let the bounded function f: [a, b] $\rightarrow$ R have a finite number of discontinuities. Then $f \in R$ (a,b).

Theorem 8: Let the sec of points of discontinuity of a, bounded function f: [a, b] $\rightarrow$ R has a finite number of limit points, then $f \in R$ (a, b).

We illustrate these theorems with the help of examples.

Example: Show that the function f where $f(x) = x^2$ is integrable in every interval [a,b].

Solution: You know that the function $f(x) = x^2$ is continuous. Therefore it is integrable in every interval [a,b].

Example: Show that the function f where f(x) = [x] is integrable in [0,2] where [x] denotes the greatest integer not greater than x.

Solution:
$$[x] = \begin{cases} 0 \text{ if } 0 \le x < 1 \\ 1 \text{ if } 1 \le x < 2 \\ 2 \text{ if } x = 2 \end{cases}$$

The points of discontinuity of f in [0,2] are 1 and 2 which are finite in number and so it is integrable in [0,2].

Example: Show that the function F defined on the interval [0,1] by

$$F(x) = \begin{cases} 2rx, \text{ when } \frac{1}{r+1} < x \le \frac{1}{r}, \text{ where r is a positive integer} \\ 0, \text{ elsewhere,} \end{cases}$$

is Riemann integrable.

Solution: The function F is discontinuous at the points $0, 1, \frac{1}{2}, \frac{1}{3}, \ldots$. The set of points of discontinuity has 0 as the only limit point. So, the limit points are finite in number and hence the function F is integrable in [0,1], by Theorem 8.

There is one more class of integrable functions and this class is that monotonic functions. This we prove in the following theorem.

Theorem 9: Every monotonic function is integrable.

Proof: We shall prove the theorem for the case where I: [a,b] $\to$ R is a monotonically increasing function. The function is bounded. f(a) and f(b) being g.l.b. and l.u.b. Let $\in$ > 0 be given number, Let n be a positive integer such that

$$n > \frac{(b-a)[f(b)-f(a)]}{\in}$$

Divide the interval [a,b] into n equal sub-intervals, by the partition P = {x_0, x_1 …, x_0} of [a, b]. Then

$$U(P,f) - L(P,f) = \sum_{i=1}^{n} (M_i - m_i)(\Delta x_i)$$

$$= \frac{b-a}{n} \sum_{i}^{n} [f(xi) - f(xi-1)]$$

$$= \frac{(b-a)}{n} [f(b) - f(a)] < \in.$$

This proves that f is integrable. Discuss the case of monotonically decreasing function as an exercise. Do it by yourself.

Exercise: Show that a monotonically decreasing function is integrable.

Now we give example to illustrate the theorem.

Example: Show that the function f defined by the condition $f(x) = \frac{1}{2^n}$

when $\frac{1}{2^{n+1}} < x \le \frac{1}{2^n}, n = 0,1,2 \ldots$

is integrable in [0,1]

Solution: Here we have f (0) = 0,

$$f(x) = 1 \text{ when } \frac{1}{2} < x \le 1$$

$$f(x) = \frac{1}{2} \text{ when } \left(\frac{1}{2}\right)^2 < x \le \frac{1}{2}$$

Clearly f is monotonically increasing in [0, 1]. Hence it is integrable.

3.3 Algebra of Integrable Functions

As we discussed the algebra of the derivable functions. Likewise, we shall now study the algebra of the integrable functions. In the previous class, you have seen that there are integrable as well as non-integrable functions. In this section you will see that the set of all integrable functions on [a,b] is closed under addition and multiplication by real numbers, and that the integral of a sum equals the sum of the integrals. You will also see that difference, product and quotient of two integrable functions is also integrable.

All these results are given in the following theorems.

Theorem 10: If f ∈ R (a , b), and λ is any real number, then λ f ∈ R (a,b) and

$$\int_a^b \lambda\, f(x)\, dx = \lambda \int_a^b f(x)\, dx.$$

Proof: Let P = {x,, x_1,...., x_n) be a partition of [a,b]. Let M_i and m_i be the respective l.u.b. and g.l.b. of the function f in [x_{i-1}, x_i]. Then λ M_i and λ m_i are the respective l.u.b. and g.l.b. of the function λ f in [x_{i-1}, x_i], if A ≥ 0, and λ m_i and λ M_i are the respective l.u.b. and g.l.b. of h f in [x_{i-1}, x_i], if h < 0.

When $\lambda \ge 0$, then $U(P,\lambda f) = \sum_{i=1}^{n} A\, M_i\, \Delta x_i = \lambda \sum_{i=1}^{n} M_i\, Ax, = \lambda\, U(P,f).$

$$\Rightarrow \qquad \int^6 \lambda f(x)\, dx = \lambda \int_a^{\bar{b}} f(x)\, dx.$$

Similarly L(P, λ f) = A L(P,f).

$$\Rightarrow \qquad \int_{\underline{a}}^b \lambda f(x)\, dx = \lambda \int_{\underline{a}}^b f(x)\, dx$$

If $\lambda < 0, U(P,\lambda f) = \sum_{i=1}^{n} \lambda\, m_i\, \Delta x_i = \lambda\, L(P,f).$

$\Rightarrow \qquad \overline{\int_a^b} \lambda\, f(x)\, dx = \lambda \underline{\int_a^b} f(x)\, dx$

Similarly L(P, λf) = λ U L(P,f).

$\Rightarrow \qquad \int_a^h \lambda\, f(x)\, dx = A \int_a^b f(x)\, dx$

Since f is integrable in [a,b], therefore

$$\int_a^6 f(x)\, dx = \int_a^b f(x)\, dx = \int_a^b f(x)\, dx.$$

Hence $\int^6 \lambda\, f(x)\, dx = \int^6 \lambda\, f(x)\, dx = \lambda \int_a^b f(x)\, dx,$

whether $\lambda \geq 0$ or $\lambda < 0$.

Hence $\lambda\, f \in R\,[a,b]$ and $\int_a^b \lambda\, f(x)\, dx = \lambda \int_a^b f(x)\, dx.$

Now suppose that λ = –1. In this case the theorem says that if f ∈ R [a,b], then $(-f) \in R[a,b]$

$\int_a^b [-f(x)] dx = \int_{-a}^b f(x)\, dx.$

Theorem 11: If f ∈ R [a,b], g ∈ R [a,b],. then f + g ∈ R [a,b] and

$\int^b (f+g)(x) dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$

Proof: We first show that f+g ∈ R [a,b]. Let ∈ > 0 be a given number. Since f ∈ R [a,b], g ∈ R [a,b], there exist partitions P and Q of [a,b] such that U(P,f) – L(Pλf) < ∈/2 and U (Q,g) – L (Q,g) < ∈/2

If T is a partition of [a,b] which refines both P and Q, then

U(T,f) – L(T,f) < ∈/2 [U(T,f) – L(T,f) ≤ U(P,f) – L(P,f)].

Similarly,

U(T,g) – L(T,g) < ∈/ 2 (5)

Also note that, if $M_i = \sup \{f(x) : x_{i-1} \leq x \leq x_i\}$

and

$N_i = \sup \{g(x): x_{i-1}\ 6\ x \leq x_i\}$

then,

$\sup \{f(x) + g(x): x_{i-1} \leq x \leq x_i\} \leq M_1 + N_i$.

Using this, it readily follows that

U(T, f+g) ≤ U(T,f) + U(T,g)

for every partition T of [a,b]. Similarly

L(T,f+g) ≥ L(T,f) + L(T,g)

for every partition T of [a,b].

Thus U (T,f+g) – L (T,f+g) ≤ [U(T,f) + U(T,g) – L [(T,f) + L(T,g)]

$= [U(T,f) - L(T,f)] + [U(T,g) - L(T,g)] < \frac{\in}{2} + \frac{\in}{2} = \in$ for T occurring in (5). This shows that f + g ∈ R(a,b)

It remains to show that $\int_a^b [f(x) + g(x)]\, dx = \int_a^b [f(x)\, dx + \int_a^b g(x)$

Now

$$\int^b (f+g)(x)dx = \overline{\int_a^b} (f+g)\,(x)\,dx \le U(P, f+g) \le U(P, f) + U(P, g) \tag{6}$$

for any partition P of [a,b]. Given any $\in > 0$ we can find a partition P of [a,b] such that

$$U(P, f) < \int_a^b f(x)\,(x)\,dx + \in/2$$

$$U(P, g) < \int_a^b g(x)\,dx + \in/2 \tag{7}$$

Substituting (7) in (6), we obtain

$$\int_a^b (f+g)\,(x)\,dx < \int_a^b f(x)\,dx + \int_a^b g(x)\,dx + \in \tag{8}$$

Since (8) holds for arbitrary $\in > 0$, we obtain

$$\int_a^b (f+g)\,(x)\,dx \le \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \tag{9}$$

Replacing f and g by –f and –g in (9) we obtain

$$\int^b (-f-g)\,(x)\,dx \le \int^b \{-f\,(x)\}\,dx + \int^b \{-g(x)\}\,dx$$

or

$$\int_a^b (f+g)\,(x)\,dx \le -\int_a^b f(x)\,dx - \int_a^b g(x)\,dx$$

This is equivalent to

$$\int_a^b (f+g)\,(x)\,dx \ge \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

Combining (9) and (10), we get

$$\int_a^b (f+g)\,(x)\,dx = \int_a^b f(x)\,dx \int_a^b g(x)\,dx$$

Which proves the theorem.

Theorem 12: If f ∈ R(a,b) and g ∈ R (a,b), then f – g ∈ R(a, b) and

$$\int_a^b (f-g)\,(x)\,dx = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx.$$

Proof: Since g E R [a,b], therefore -g ∈ R [a,b] and

$$\int^b -[g(x)]dx = -\int^b g(x)\,dx$$

Now f ∈ R [a,b] and –g ∈ R [a,b] implies that f + (–g) ∈ R [a,b] and therefore,

$$\int_a^b [f+(-g)](x)\,dx = \int_a^b f(x)\,dx + \int_a^b [-g(x)]\,dx$$

that is (f – g) ∈ R [a,b] and

$$\int_a^b (f-g)\,(x)\,dx = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx.$$

For the product and quotient of two functions, we state the theorems without proof.

Theorem 13: If f ∈ R(a,b) and g ∈ R(a,b), then f g ∈ R(a,b).

Theorem 14: If f ∈ R(a,b), g ∈ R(a,b) and there exists a number t > 0 such that $|g(x)| \geq t, \forall\ x \in [a,b]$, then f/g ∈ R(a,b).

Now we give some examples.

Example: Show that the function f, where f(x) = x + [x] is integrable is [0, 2].

Solution: The function F(x) = x, being continuous is integrable in [0, 2] and the function G(x) = [x] is integrable as it has only two points namely, 1 and 2 as points of discontinuity. So their sum is, f(x) is integrable in [0, 2].

Example: Give an example of function f and g such that f + g is integrable but f and g are not integrable in [a, b].

Solution: Let f and g be defined in [a, b] such that

$$f(x) = \begin{cases} 0, \text{ when x is rational} \\ 1, \text{ when x is irrational,} \end{cases}$$

$$g(x) = \begin{cases} 1, \text{ when x is rational} \\ 0, \text{ when x is irrational} \end{cases}$$

f and g are not integrable but $(f+g) = 1\ \forall\ x \in [a,b],$ being a constant function, is integrable.

3.4 Computing an Integral

So far, we have discussed several theorems for testing whether a given function is integrable on a closed interval [a,b]. For example, we can see that a function $f(x) = x^2\ \forall\ x \in [0,2]$ is continuous as well as monotonic on the given interval and hence it is integrable over [0,2]. But this information does not give us a method for finding the value of the integral of this function. In practice, this is not so easy as we might think of. The reason is that there are some functions which are integrable by conditions of integrability but it is difficult to find the values of their integrals. For example, suppose a function is given by $f(x) = e^{x^2}$ This is continuous over every closed interval and hence it is integrable. But we cannot find its integral by our usual method of anti derivative since there is no function for which $f(x) = e^{x^2}$ is the derivative. If possible, try to find the anti derivative for this function.

Notes

In such situations, to find the integral of a given function, we use the basic definition of the integral to evaluate its integral. Indeed, the definition of integral as a limit of sum helps us in such situations.

In this section, we demonstrate this method by means of certain examples. We have found the integral $\int_a^b f(x)\,dx$ via the sums U(P,f) and L(P,f). The numbers M_i and m_i which appear in these sums are not necessarily the values of f(x), i f f is not continuous. In fact, we shall now show that $\int f(x)\,dx$ can be considered as limit of sums in which M_i and m, are replaced by values of f. This approach gives us a lot of latitude in evaluating $\int_a^b f(x)\,dx,$ as we shall see in several examples.

Let f: [a,b] $\to$ R be a bounded function. Let

la = $x_0 < X_1 <$ ...... x_n = b]

be a partition P of [a,b]. Let us choose points t_1, ... t_n, such that

$x_{i-1} \le t_i \le x_i$ (i = 1, ... n). Consider the sum

$$S(P,f) = \sum_{i=1}^{n} f(t_i)\,\Delta x_i = \sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1}).$$

Notice that, instead of M_i in U(P, f) and m_i in L(P,f), we have $f(t_i)$ in S(P, f). Since t_i 's are arbitrary points in $[x_{i-1}, x_i]$, S(P, f) is not quite well-defined. However, this will not cause any trouble in case of integrable functions.

S(P,f) is called Riemann Sum corresponding to the partition P.

We say that lim S(P,f) = A

$$|P| - 0$$

or $S(P,f) \to A$ as $|P| \to 0$ if for every number $\in < 0\ \exists\ \delta > 0$ such that

$$|S(P,f) - A| < \in \text{for P with } |P| < 6.$$

We give a theorem which expresses the integral as the limit of S(P,f).

Theorem 15: If $\lim\limits_{|P|\to 0}$ S(P,f) exists, then f $\in$ R (a,b) and $\lim\limits_{|P|\to 0} S(P,f) = \int_a^b f(x)dx.$

Proof: Let $\lim\limits_{|P|\to 0}$ S(P,f) = A. Then, given a number $\in$ > 0, there exists a number δ > 0 such that

$$|S(P,f) - A| < \in/4, \text{ for P with } |P| < \delta.$$

i.e., $\quad A - \in/4 < S(P,f) < A + \in/4,$ for P with $|P| < \delta.$ (11)

Let P = {x_0, x,, ..., x_n}. Suppose the points t, ..., t_n vary in the intervals [x_0, x,], ..., [x_{n-1}, x_n], respectively. Then, the l.u.b. of the numbers S(P,f) are given by

$$\text{l.u.b. } S(P,f) = \text{l.u.b.}\left(\sum_{i=1}^{n} f(t_i)\,\Delta\,x_i\right) = \sum_{i=1}^{n} M_i\,\Delta\,x_i = U(P,f).$$

Similarly, g.1.b. S(P,f) = L(P,f). Then, from (11), we get

$$A - \in/4 \le L(P,f) \le U(P,f)\ 5\ A + \in/4 \quad (12)$$

Therefore,

Notes

$$U(P,f) - L(P,f) \leq (A + \in/14) - (A - \in/4)$$

$$= \in/2 < \in.$$

In other words, $f \in R(a,b)$. Thus

$$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx = \int_a^b f(x)\,dx.$$

Since $L(P,f) \leq \int_a^b f(x)\,dx \leq \int^h f(x)\,dx \leq U(P,f)$, therefore

$$L(P,f) \leq \int_a^b f(x)\,dx \leq U(P,f). \quad (13)$$

From (12) and (13), we get

$$A - \in/4 \leq \int_a^b f(x)\,dx \leq A + \in/4.$$

That is,

$$\left|\int_a^b f(x)\,dx - A\right| \leq \in/4 < e.$$

Since $\in$ is arbitrary, therefore $\int_a^b f(x)\,dx - A = 0$, that is, $\int_a^b f(x)\,dx = A = \lim_{|P|\to 0} S(P,f)$. This completes the proof of the theorem.

To illustrate this theorem, we give two examples.

Example: Show that $\int_a^b dx = \int_a^b 1\,dx = b - a$.

Solution: Here, the function f: [a,b] → R is the constant function f(x) = 1.

Clearly, for any partition $P = (x_0, x_,, ..., x_n)$ of [a,b], we have

$$S(P,f) = (x_1 - x_0)f(t_1) + (x_2 - x_1)\,f(t_2) + ... + (x_n - x_{n-1})f(t_n)$$

$$= (x_1 - x_0)1 + (x_2 - x_1)1 + ... + (x_n - x_{n-1})1 = b - a.$$

Since S(P,f) = b – a, for all partitions, $\int_a^b 1\,dx = \lim_{|P|\to 0} S(P,f) = b - a.$

Example: Show that $\int_a^b x\,dx = \dfrac{b^2 - a^2}{2}$.

Solution: The function f:[a,b] → R in this example is the identity function f(x) = x.

Let $P = (a = x_,, x_,, \ldots, x_a = b)$ be any partition of [a,b]. Then

$S(P, f) = (x_1 - x_0)\,f(t_1) + (x_2 - x_1)\,f(t_2) + ... + (x_n - x_n)\,f(t_n)$, where $t_1 \in [x_0, x_1,]$, $t_2 \in [x_1, x_2]$, …

$t_n \in [x_{n-1}, x_n]$ are arbitrary. Let us choose

$$t_1 = \frac{x_0 + x_1}{2}, t_2 = \frac{x_1 + x_2}{2}, \ldots, t_n = \frac{x_{n-1} + x_n}{2}.$$

Notes

Then, $S(P,f) = (x_1 - x_0)\frac{x_1 + x_0}{2} + (x_2 - x_1)\frac{x_2 + x_1}{2} + \ldots + (x_n - x_{n-1})\frac{x_n + x_{n-1}}{2}$

$$= \frac{1}{2}\left[\left(x_1^2 - x_2^0\right) + \left(x_2^2 - x_2^1\right) + \ldots + \left(x_n^2 - x_{n-1}^2\right)\right]$$

$$= \frac{1}{2}\left(x_n^2 - x_0^2\right) = \frac{1}{2}\left(b^2 - b^2\right).$$

Here again, $S(P,f) = \frac{1}{2}\left(b^2 - a^2\right)$, no matter what the partition P we may take, Hence $\int_a^b f(x)\,dx = \int_a^b x\,dx = \lim_{|P|\to 0} S(P,f) = \frac{1}{2}(b^2 - a^2)$.

The converse of Theorem 15 is also true which we state without proof as the next theorem.

Theorem 16: If a function f is Riemann integrable on a closed interval [a,b], then $\lim_{|P|\to 0} S(P,f)$ exists and $\lim_{|P|\to 0} S(P,f) = \int_a^b f(x)\,dx$.

One of the important application of Theorem 16 is in computing the sum of certain power series. For, let us consider a partition P of [a,b] having n sub-intervals, each of length h so that nh = b – a. Then P can be written as P = (a, a + h, a + 2h, ..., a + nh = b).

Let t, = a + ih, i = 1,2 ,...., n. Then

$$S(P,f) \sum_{i=1}^{n} f(t_i)\,\Delta x_i = h[f(a+h) + f(a+2h) + \ldots + f(a+nh)].$$

When $\lim_{|P|\to 0} S(P,f)$ exists, then

$$\lim_{\substack{n\to\infty \\ h\to 0}} h[f(a+h) + f(a+2h) + \ldots + f(a+nh)] = \int_a^h f(x)\,dx.$$

In the above formulae, we can change the limits of integration from a, b to 0, a, where a $\in$ N. For, by changing h to $\frac{b-a}{an}$, it is easy to deduce from above formula that

$$\frac{(b-a)}{\alpha}\lim_{n\to\infty}\frac{1}{n}\sum_{=}^{n} f\left[a + \frac{(b-a)}{\alpha}\frac{r}{h}\right] = \int_a^b f(x)\,dx. \quad (14)$$

But, $\int_a^b f(x)\,dx = \frac{(b-a)}{\alpha}\int_0^{\alpha} f\left[a + \frac{(b-a)}{\alpha}x\right]dx.$

Therefore, from (14), we get

$$\lim_{n\to\infty}\frac{1}{n}\sum_{r=1}^{n} f\left[a + \frac{(b-a)}{\alpha}\frac{r}{n}\right] = \int_0^{\alpha} f\left[a + \frac{(b-a)}{\alpha}x\right]dx. \quad (15)$$

In (15), put a = 0, b = α. We get the following result:

If f is integrable in [0, α], then

$$\lim_{n\to\infty}\sum_{r=1}^{n}\frac{1}{n} f\left(\frac{r}{n}\right) = \int_0^{\alpha} f(x)\,dx.$$

Notes

This gives us the following method for finding the limit of sum of n terms of a series:

1. Write the general rth term of the series.

2. Express it as $\frac{1}{n}f\left(\frac{r}{n}\right)$, the product of $\frac{1}{n}$ and a function of $\frac{r}{n}$.

3. Change $\frac{r}{n}$ to x and $\frac{1}{n}$ to dx and integrate between the limits 0 and a. The n value of the resulting integral gives the limit of the sum of n terms of the series.

Since each term of a convergent series tends to 0, the addition or deletion of a finite number of terms of the series does not affect the value of the limit. Similarly, you can verify that

$$\lim_{n\to\infty}\sum_{r=1}^{2n}\left[\frac{1}{n}\phi\left(\frac{r}{n}\right)\right] = \int_0^2 \phi(x)\,dx,$$

$$\lim_{n\to\infty}\sum_{r=1}^{3n}\left[\frac{1}{n}\phi\left(\frac{r}{n}\right)\right] = \int_0^3 \phi(x)\,dx, \text{ and so on.}$$

As an illustration of these results, consider the following examples.

Example: Find the limit, when n tends to infinity, of the series

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

Solution: General (rth) term of the series is $\sum_{r=1}^{n}\frac{1}{n+r} = \sum_{r=1}^{n}\frac{1}{n}\left(\frac{1}{1+\frac{r}{n}}\right)$.

Hence, $\lim_{n\to\infty}\sum_{r=1}^{n}\frac{1}{n}\left(\frac{1}{1+\frac{r}{n}}\right) = \int_0^1\frac{1}{1+x} = dx = \log 2.$

Example: Find the limit, when n tends to infinity, of the series

$$\frac{1}{\sqrt{n^2}}+\frac{1}{\sqrt{n^2-1}}+\frac{1}{\sqrt{n^2-2^2}}+\ldots+\frac{1}{\sqrt{n^2-(n-1)^2}}.$$

Solution: Here the rth term $= \sum_{r=1}^{n}\frac{1}{\sqrt{n^2-(r-1)^2}}$

Since it contains (r – 1), we consider its (r + l)th term i.e.,

the term $\sum_{r=0}^{n}\frac{1}{\sqrt{n^2-r^2}} = \sum_{r=0}^{n}\frac{1}{n}\frac{1}{1+\left(\frac{r}{n}\right)^2}$

Notes

Therefore, $\lim_{n\to\infty} \sum_{r=1}^{n} \frac{1}{\sqrt[n]{1+\left(\frac{r}{n}\right)^2}} = \int_0^1 - \frac{1}{\sqrt{1-x^2}} dx.$ because $\lim_{n\to\infty} \frac{1}{n} = 0.$

The value of this integral, on the r.h.s. of last equality, is $\frac{\pi}{2}$,

Example: Find $\lim_{n\to\infty} \sum_{r=1}^{3^n} \frac{n^2}{(3^n + r)^3}$.

Solution: We have

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

Since the number of terms in the summation is 3^n, the resulting definite integral will have the limits from 0 to 3.

Therefore, $\lim_{n\to\infty} \sum_{r=1}^{3^n} \frac{n^2}{\left(\overline{3^n} + \bar{r}\right)^3} = \lim_{n\to\infty} \sum_{r=1}^{3^n} \frac{1}{n} \frac{1}{(3+\underline{r})^3} = \int_0^3 \frac{dx}{(3+x)^3}$

This integral you can evaluate easily.

Self Assessment

Fill in the blanks:

1. Let P, and P, be two partitions of [a,b]. We say that P, is finer than P, or P_2, refines P, or P_2 is a refinement of P_1 if ...., that is, every point of P_1 is a point of P,.
2. Let f: [a,b] $\rightarrow$ R be a bounded function. The infimum or the greatest lower hound of, the set of ail upper sums is called the upper (Riemann) integral of f on [a, b] and is denoted by,..
3. If the partition P_2 is a refinement of the partition P, of [a,b], then $L(P_1,f) \le L(P_2,f)$ and ....
4. The integrability is not affected even when there are finites number of points of .... or the set of points of discontinuity of the function has a finite number of limit points.
5. If f : [a, bJ $\rightarrow$ R is a ...., then f is integrable over [a, b], that is f $\in$ R(a,b).

3.5 Summary

- In this unit, you have been introduced to the concept of integration without bringing in the idea of differentiation. As upper and lower sums and integrals of a bounded function f over closed interval [a,b] have been defined. You have seen that upper and lower Riemann integrals of a bounded function always exist. Only when the upper and lower Riemann integrals are equal, the function f is said to be Riemann integrable or simply integrable over [a,b] and we write it as f $\in$ R [a,b] and the value of the integral of f over [a,b] is denoted by $\int_a^b f(x)\,dx$. Also in this section, it has been shown that in passing from a partition PI to a finer partition P2, the upper sum does not increase and the lower sum does not decrease. Further, you have seen that the lower integrable of a function is less than or equal to the upper integral. Further condition of integrability has been derived with the help of which the integrability of a function can be decided without finding the upper and lower integrals. Using the condition of integrability, it has been shown that a function f is integrable on [a,b] if it is continuous or it has a finite number of points of discontinuities or the set of points of discontinuities have finite number of limit points. Also you have seen that a monotonic function is integrable. As in the case of continuous and derivable functions, the sum, difference, product and quotient of integrable functions is integrable. Riemann sum S(P,f) of a function f for a partition P has been defined and you have been shown that $\lim_{|P|\to 0} S(P,f)$ exists if and only if $f \in R$ [a,b] and $\int_a^b f(x)\,dx = \lim_{|P|\to 0} S(P,f)$. Using this idea a number of problems can be solved.

3.6 Keywords

Partition: Let [a,b] be a given interval. By a partition P of [a,b] we mean a finite set of points $\{x_0, x_1, ..., x,\}$, where

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

We write $\Delta x_i = x_i - x_{i-1}$, (i=l, 2, ..., n). So Δx_i is the length of the ith sub-interval given by the partition P.

Norm of a Partition: Norm of a partition P, denoted by $|P|$, is defined by $|P| = \max_{t \le i \le n} Ax,$. Namely, the norm of P is the length of largest sub-interval of [a, b] induced by P. Norm of P is also denoted by $\mu(P)$.

Darboux's Theorem: If f: [a,b] $\to$ R is a bounded function, then to every $\in > 0$, there corresponds $\delta > 0$ such that

(i) $U(P,f) < \overline{\int_a^b} f(x)\,dx + \in$

(ii) $L(P,f) < \int_a^b f(x)\,dx - \in$

3.7 Review Questions

1. Find the upper and lower Riemariri integrals of the function f defined in [a, b] as follows

$$f(x) = \begin{cases} 1 \text{ when x is ratinal} \\ -1 \text{ when x is irrational} \end{cases}$$

2. Show that the function f where f(x) = x[x] is integrable in [0, 2].

3. Show that the function f defined in LO, 21 such that f(x) = 0, when $x = \frac{n}{n+1}$ or $\frac{n+1}{n}$ $(n = 1,2,3,\ldots)$, and f(x) = 1, elsewhere, is integrable.

Notes

4. Prove that the function f defined in [0, 1] by the condition that if r is a positive integer, $f(x) = (-1)^{r-1}$ when $\frac{1}{r+1} < x \leq \frac{1}{r}$, and f(x) = 0, elsewhere, is integrable.

5. Show that the function f defined in [0, 1], for integer a > 2, by $f(x) = \frac{1}{a^{r-1}}$, when $\frac{1}{a^r} < x < \frac{1}{a^{r-1}}$ $(r = \frac{1}{a^{r-1}} 1, 2, 3)$, and f(0) = 0, is integrable.

6. Give example of functions f and g such that f – g, fg, f/g are integrable but f and g may not be integrable over [a, b].

7. Find the limit, when n tends 10 infinity, of the series

$$\frac{\sqrt{n}}{\sqrt{n^3}} + \frac{\sqrt{n}}{\sqrt{(n+4)^3}} + \frac{\sqrt{n}}{\sqrt{(n+8)^3}} + \ldots + \frac{\sqrt{n}}{\sqrt{[n+4(n-1)]^3}}$$

8. Find the limit, when 11 tends to infinity, of the series

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

Answers: Self Assessment

1. $P_1 \subset P_2$
2. $\int_a^{\bar{h}} f(x)dx.$
3. $U(P_2,f) \leq U(P_1,f)$
4. discontinuity
5. continuous function

3.8 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 4: Properties of Integrals

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Identify the properties of the integral and
- Use them to find the Riemann Stieltjes integral of functions

Introduction

In last unit you have studied about Riemann integral. In this unit, we are going to see the properties of Riemann Stieltjes integral.

4.1 Properties of Riemann Integral

As you were introduced to some methods which enabled you to associate with each integrable function f defined on [a,b], a unique real number called the integral $\int_a^b f(x)\,dx$ in the sense of Riemann. A method of computing this integral as a limit of a sum was explained. All this leads us to consider some nice properties which are presented as follows:

Property 1: If f and g are integrable on [a, b] and if

$$f(x) \leq g(x)\ \forall\ x \in [a,b],$$

then

$$\int^b f(x)\,dx \leq \int^b g(x)\,dx$$

Proof: Define a function h: [a,b] → R as

$$h = g - f\,.$$

Since f and g are integrable on [a, b], therefore, the difference h is also integrable on [a, b].

Since

$$f(x) \leq g(x) \geq g(x) - f(x) \geq 0,$$

therefore h(x) ≥ 0 for all x E [a,b].

Notes

Consequently, if P = {x_0, x_1, … x_n} be any partition of [a,b]

and m, be the inf. of h in [x, 1, x,], then

$$m_i \geq 0 \ \forall \ i = 1, 2, \ldots n$$

$$\Rightarrow \sum_{i=1}^{n} m_i \Delta \, x_i \geq 0$$

$$\Rightarrow L(P,h) \geq 0$$

Thus for every partition P, the lower sum L(P,h) ≥ 0.

In other words, Sup. (1 (P,h): P is a partition of [a,b]) ≥ 0

or

$$\int_a^h f(x)\, dx$$

Since h is integrable in [a,b], therefore

$$\int_a^b h(x)\, dx = \int_a^b h(x)\, dx = \int_a^b h(x)\, dx.$$

Thus

$$\int_a^b h(x)\, dx \geq 0$$

or

$$\int_a^b (g - f)\,(x)\, dx \geq 0$$

$$\Rightarrow \qquad \int_a^b g(x)\, dx \geq \int_a^b f(x)\, dx$$

which proves the property

Property 2: If f, is integrable on [a, b] then $|f|$ is also integrable on

$[a, b]$ and $\left|\int_a^b f(x)\, dx\right| \leq \int_a^b |f(x)|\, dx$

Proof: The inequality follows at once from Property 1 provided it is known that $|f|$ is integrable on [a,b]. Indeed, you know that $-|f| \leq f \leq |f|$.

Therefore,

$$\int_{-a}^b |f(x)|\, dx \leq \int^b f(x)\, dx \leq \int_a^b |f(x)|\, dx$$

which proves the required result. Thus, it remains to show that $|f|$ is integrable.

Let ∊ > 0 be any number. There exists a partition P of [a, b] such that

$$U(P,f) - L(P,f) \leq \in$$

Let $P = \{x_0, x_1, x_2, \ldots, x_n\}$.

Let M_i' and m_i' denote the supremum and infimum of $|f|$ and M_i and m_i denote the supremum and infimum of f in $[x_{i-1}, x_i]$.

You can easily check that

$$M_i - m_i \geq M_i' - m_i'.$$

This implies that $\sum_{i=1}^{n}(M_i - m_i)\, A\, x_i \geq \sum_{i=1}^{n}(M_1' - m_1')\,\Delta x_i$,

i.e., $U(P,|f|) - L(P,|f|) \leq U(P,f) - L(P,f) < \in$,

This shows that $|f|$ is integrable on [a,b].

Note that the inequality established in Property 2 may be thought of as a Integrability and differentiability generalization of the well-known triangle inequality

$$|a+b| \leq |a| + |b|$$

In other words, the absolute value of the limit of a sum never exceeds the limit of the sum of the absolute values.

You know that in the integral $\int_a^b f(x)\,dx$, the lower limit a is less than the upper limit b. It is not always necessary. In fact the next property deals with the integral in which the lower limit a may be less than or equal to or greater than the upper limit b.

For that, we have the following definition:

Definition 1: Let f be integrable on [a,b], that is, $\int_a^b f(x)\,dx$ exists when b > a. Then

$$\int_a^b f(x)\,dx = 0, \text{ if } a = b$$

$$= -\int_b^a f(x)\,dx, \text{ if } a > b.$$

Now have the following property.

Property 3: If a function f is integrable in [a,b] and $|f(x)| \leq k \ \forall\, x \in [a,b]$, then $\left|\int_a^b f(x)\,dx\right| \leq k|b-a|$.

Proof: There are only three possibilities namely either a < b or a > b or a = b. We discuss the cases as follows:

Case (i): a < b

Since $|f(x)| \leq k \quad \forall\, x \in [a,b]$, therefore

$-k \leq f(x) \leq k \quad \forall\, x \in [a,b]$

$\Rightarrow \quad \int_{a}^{h} -k dx \le \int_{a}^{b} f(x)\, dx \le \int_{a}^{b} k\, dx$ (why?)

$\Rightarrow \quad -k(b-a) \le \int_{a}^{b} f(x)\, dx \le k(b-a)$

$\therefore \quad \left|\int_{a}^{b} f(x) dx\right| \le k(b-a) = k|b-a|$

which completes the proof of the theorem.

Case (ii): a > b

In this case, interchanging a and b in the Case (i), you will get

$$\left|\int_{b}^{a} f(x)\, dx\right| \le k(a-b)$$

i.e. $\quad \left|-\int_{a}^{b} f(x)\, dx\right| \le k(a-b)$

i.e. $\quad \left|\int_{a}^{b} f(x)\, dx\right| \le k(a-b) = k|b-a|.$

Case (iii): a = b

In this case also, the result holds,

since $\int_{a}^{b} f(x)\, dx = 0$ for a = b and $k|b-a| = 0$ for a = b.

Let [a,b] be a fixed interval. Let R [a,b] denote the set of all Riemann integrable functions on this interval. We have shown that if f,g C R [a,b], then f + g f.g and λf for A ∈ R belong to R [a,b]. Combining these with Property, we can say that the set R [a,b] of Riemann integrable functions is closed under addition, multiplication, scalar multiplication and the formatian of the absolute value.

If we consider the integral as a function Int: R[a,b] → R defined by

$$\text{Int (f)} = \int_{a}^{b} f(x)\, dx$$

with domain R [a,b] and range contained in R, then this function has the following properties:

$$\text{lnt (f+g) = Int (f) + Int (g), lnt } (\lambda f) = \lambda \text{ Int (f)}$$

In other words, the function lnt preserves 'Vector sums' and the scalar products. In the language of Linear Algebra, the function lnt acts as a linear transformation. This function also has an additional interesting property such as

$$\text{lnt (f)} \le \text{lnt(g)}$$

whenever

$$f \le g .$$

We state yet another interesting property (without proof) which shows that the Riemann Integral is additive on an interval.

Notes

Property 4: If f is integrable on [a,b] and c ∈ [a,b], then f is integrable on [a,c] and [c,b] and conversely. Further in either case

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

According to this property, if we split the interval over which we are integrating into two parts, the value of the integral over the whole will be the sum of the two integrals over the subintervals. This amounts to dividing the region whose area must be found into two separate parts while the total area is the sum of the areas of the separate portions.

We now state a few more properties of the definite integral $\int_a^b f(x)\,dx$, these are:

(i) $\int_n f(x)\,dx = \int_a f(a-x)\,dx.$

(ii) $\int_0^{2a} f(x)\,dx = \int_0^a f(x)\,dx + \int_0^a f(2a-x)dx.$

(iii) $\int_{-a}^{a} f(x)\,dx = \begin{cases} 2\int_0^a f(x)\,dx \text{ if f is an even function} \\ 0 \text{ iff is an old function.} \end{cases}$

(iv) $\int_0^{na} f(x)\,dx = n\int_0^a f(x)\,dx$ if f is periodic with period 'a' and n is a positive integer provided the integrals exist.

Self Assessment

Fill in the blanks:

1. If f, is .... on [a, b] then $|f|$ is also integrable on $[a,b]$ and $\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.$

2. The inequality follows at once from Property 1 provided it is known that $|f|$ is .... on [a, b]. Indeed, you know that $-|f| \le f \le |f|$.

3. If a function f is integrable in [a, b] and ...., then $\left|\int_a^b f(x)\,dx\right| \le k|b-a|$.

4. If f is integrable on [a, b] and ...., then f is integrable on [a, c] and [c, b] and conversely.

4.2 Summary

- Sum of two Riemann Stieltjes integrable functions is also Riemann Stieltjes integrable.
- Scalar product of a Riemann Stieltjes integrable function is also Riemann Stieltjes integrable.
- Modulus of a Riemann Stieltjes integrable function is also Riemann Stieltjes integrable.
- Square of a Riemann Stieltjes integrable function is also Riemann Stieltjes integrable.
- If a function is Riemann Stieltjes integrable on an interval, then it is also Riemann Stieltjes integrable on any of its subinterval.

4.3 Keyword

Riemann Stieltjes Integrable: Sum of two Riemann Stieltjes integrable functions is also Riemann Stieltjes integrable.

4.4 Review Questions

1. Calculate if a < b, $\int_b^a f d\alpha$
2. Suppose f is a bounded valued function on [a, b] and f2 ∈ R on [a, b]. Does it follow that f ∈ R on [a, b]?
3. Show that 0∫1 x2dx = 3/5 where α(n) = x3
4. Show that 0∫2 [x]dx = 3/5 where α(x) = x2 = 3.

Answers: Self Assessment

1. integrable	2. integrable
3. $\lvert f(x) \rvert \le k \ \forall \ x \ E[a,b]$	4. c ∈ [a, b]

4.5 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Notes

Unit 5: Introduction to Riemann-Stieltjes Integration, using Riemann Sums

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss Riemann–Stieltjes sums
- Know the Cauchy Criterion for Riemann–Stieltjes Integrability

Introduction

We will approach Riemann-Stieltjes integrals using Riemann-Stieltjes sums instead of the upper and lower sums. The main reasons are to study Riemann-Stieltjes integrals with "integrators" $\alpha(x)$ that are not monotone, but are "of bounded variation," and (most important) here you are able to define Riemann-Stieltjes integrals when the values of my functions belong to an infinite dimensional vector space, where upper and lower sums don't make sense. This makes little difference in the case of real-valued functions, since functions of bounded variation can always be expressed as the difference of two monotone functions. At first, we don't need "bounded variation," so that concept's development will wait until it is needed.

Throughout this note, our functions f(x) will be "finite-valued." They may be real, complex, or vector-valued. Their values will thus lie in a vector space. They can thus be added pointwise, and multiplied by scalars, and their values always have finite "distance from zero," denoted $|f(x)|$, which can denote absolute value or *norm*, such as the length of a vector, or the "L^p norm" and the "L^q norm". In case f(x) is actually a function of t for each x... We always assume that the "absolute value" is *complete*; Cauchy sequences converge.

Notes

5.1 Riemann-Stieltjes Sums

A Riemann-Stieltjes sum for a function f(x) defined on an interval [a, b] is formed with the help of

1. A partition π of [a, b], namely an ordered, finite set of points x_i, with $a = x_0 < x_1 < \cdots < x_n = b$ (where n is a positive integer that can be any positive integer, and one that we will often write as $n = n_\pi$),

2. A selection vector $\xi = (\xi_1, ..., \xi_n)$ that has n_π components that must satisfy $x_{i-1} \leq \xi_i \leq x_i$, for i = 1, 2, ..., n.

   and

3. An integrator $\alpha(x)$, which is a function defined on [a, b] that plays the role of the x in dx ...

   A Riemann-Stieltjes sum for f over [a, b] with respect to the partition π, using the selection vector ξ, and integrator α, may be denoted (in greatest detail!) as follows, and it is given by the value of the sum following it:

4. $$RS\ (f, \alpha, [a, b], \pi, \xi) := \sum_{i=1}^{n_\pi} f(\xi_i)(\alpha(x_i) - \alpha(x_{i-1})).$$

5.2 More Notation: The Mesh (Size) of a Partition

In this definition, as in the Riemann-sums definition, we can write $\Delta x_i := x_i - x_{i-1}$ or $\Delta\alpha_i := \alpha(x_i) - \alpha(x_{i-1})$. These are convenient because they are short and suggest the dx or $d\alpha$ in an integral. But they can cause confusion because they leave out the dependence they have on x_{i-1}. The Δx_i is used in the Riemann-Stieltjes context.

A partition π can be thought of as "dividing" the interval [a, b] into subintervals. We may write $\pi\,|\,[a, b]$ and read this as "π divides [a, b]," or "partitions [a, b]." We will denote the intervals of π by $I_i := [x_{i-1}, x_i]$. When we wish to work with 2 partitions at the same time we will have to distinguish between them somehow, for example we can use y_j to denote the other's points and J_j to denote its intervals, etc.

We measure the fineness of a partition using the length of the longest interval in the partition. This number is written

$$\text{mesh}(\pi) := \max_{1 \leq i \leq n_\pi} (x_i - x_{i-1}) = \max_{1 \leq i \leq n_\pi} \Delta x_i.$$

This definition of mesh size is used and not $\max_{1 \leq i \leq n_\pi} (\alpha(x_i) - \alpha(x_{i-1}))$ even in the Riemann-Stieltjes context.

5.3 The Riemann-Stieltjes – sum Definition of the Riemann-Stieltjes Integral

Definition: A real-valued function f(x) defined on the bounded and closed interval [a, b] is Riemann-Stieltjes integrable on [a, b] with respect to α if there exists a number RSI such that for all $\varepsilon > 0$ there exists $\delta > 0$ such that for every partition π of [a, b],

$$\text{mesh}(\pi) < \delta \Rightarrow |RS(f, \pi) - RSI| < \varepsilon.$$

We write

$$\int_a^b f\, d\alpha = \int_a^b f(x)\, d\alpha(x) := RSI$$

and we call this the Riemann-Stieltjes integral of f over [a, b] with respect to α.

If f and α are real-valued and we imagine the set of all numbers RS(f, α, π) that can be formed (using all possible appropriate selection vectors and all possible partitions whose mesh sizes are less than δ), the definition demands that they all lie in the open interval (RSI – ε, RSI + ε). When we had $\alpha(x) = x$ this led to a Theorem.

Theorem: If f is Riemann integrable on [a, b] then f is bounded on [a, b].

This Theorem has to be modified in the Riemann-Stieltjes context! A simple example: suppose that [a, b] is [0, 1] and that $\alpha(x) = 0$ if $0 \le x \le c$, where $0 < c < 1$, and $\alpha(x) = 1$ if $c < x \le 1$. Then every function f(x) that is continuous at c is Riemann-Stieltjes integrable on [0, 1] with respect to this α. In particular the function that is 1/x except at zero, where we define it to be zero, is Riemann-Stieltjes integrable on [0, 1] with respect to this α, but f is not bounded. The difference is that when $\alpha(x)$ was just x, we had $\Delta x_i > 0$ for every i. In our example, $\Delta\alpha_i = 0$ unless I_i contains c and some d with c < d. What we need is that on the set where the function α "really" varies, f must be bounded. To make a definition, we will extend the definitions of f and α beyond the interval [a, b] by setting them equal to their values at the endpoints. Thus we think of f(x) = f(a) if x < a and f(x) = f(b) if x > b, with the same idea used to extend α. We now define the oscillation of f on an interval U by

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

We allow the interval to be open or half-open now!

As before, we will let $\omega_i = \omega_i(f) = \omega(f, I_i)$ when I_i is an interval (closed!) of a partition π. But now we need to use oscillations of α as well.

Definition: If $\alpha(x)$ is defined for $x \in [a, b]$, we denote by $\Omega = \Omega(\alpha, [a, b])$ the set of all $c \in [a, b]$ such that every open interval U that contains c contains $x_1 < c < x_2$ with $|\alpha(x_1) - \alpha(x_2)| > 0$.

Notes Here c can be a or b because of our extension beyond [a, b]! For instance, if for all $\delta > 0$ there exists x_2 such that $a < x_2 < a + \delta$ and $|\alpha(a) - \alpha(x_2)| > 0$, then $a \in \Omega(\alpha, [a, b])$ because for every $x_1 < a$ we have $|\alpha(x_1) - \alpha(x_2)| = |\alpha(a) - \alpha(x_2)| > 0$.

Task Prove that $\Omega(\alpha, [a, b])$ is closed.

Theorem: If f is Riemann-Stieltjes integrable on [a, b] with respect to α then f is bounded on $\Omega(\alpha, [a, b])$.

Proof: There exists a sequence $\{x_n\}$ in $\Omega := \Omega(\alpha, [a, b])$ such that $|f(x_n)| > n$. Since f(x) is finite at every point x in Ω, there are infinitely many distinct x_n, and so some subsequence (that we will still denote $\{x_n\}$) converges to a point x^* in Ω. We now choose $\varepsilon = 1$ in the definition of Riemann-Stieltjes integrability, and obtain a corresponding $\delta > 0$. We can then construct a partition π_o with mesh size less than δ in such a way that x^* is contained in the interior of some interval I_{i_o} of π_o (unless x^* is an endpoint of [a, b]; in that case, we can, by the Note, still use the following argument, with $I_{i_o} = I_1$ or $I_{i_o} = I_{n_\pi}$). We know that every neighbourhood of x^* contains infinitely many of the x_n. Now we will refine π_o. We know that Int(I_{i_o}) contains points $\hat{x}_1 < x^* < \hat{x}_2$ with $|\alpha(\hat{x}_1) - \alpha(\hat{x}_2)| > 0$. We add these points to π_o, giving us a new partition π, and mesh(π) < δ. We will now call $[\hat{x}_1, \hat{x}_2]$, which is an interval of π, $\hat{I}$. Next we pick the components ξ_i of a selection vector ξ in an arbitrary way when $I_i \ne \hat{I}$, and we let $\hat{\xi}$ be some $x_N \in \hat{I}$. Then $|RS(f, \pi, \xi) - RSI| < 1$. We next modify ξ by changing only $\hat{\xi} = x_N$ to $\hat{\xi}' := x_M$, where $x_M \in \hat{I}$, and we call the new selection vector ξ'. Then $|RS(f, \pi,\xi') - RSI| < 1$, $RS(f,\pi,\xi') - RS(f,\pi,\xi) = (f(x_M) - f(x_N))(\alpha(x_i) - \alpha(x_2))$ and

$$RS(f, p, \xi') - RSI = RS(f, \pi, \xi) - RSI + (f(x_M) - f(x_N))(\alpha(\hat{x}_1) - \alpha(\hat{x}_2)).$$

By choosing M very large compared to N we can arrange that $|f(x_M) - f(x_N)|\,|\alpha(\hat{x}_1) - \alpha(\hat{x}_2)| > 2$. Then

$$1 > |RS(f, \pi, \xi') - RSI| \geq |RS(f, \pi, \xi') - RS(f, \pi, \xi)| - |RS(f, \pi, \xi) - RSI| > 2 - 1 = 1.$$

The definition of Riemann-Stieltjes integrability is contradicted. Hence f is bounded on$\Omega(\alpha, [a, b])$ if f is Riemann-Stieltjes integrable with respect to α.

Notes From now on, we will usually say "f is Riemann-Stieltjes integrable" instead of "f is Riemann-Stieltjes integrable with respect to α."

5.4 A difficulty with the Definition; The Cauchy criterion for Riemann-Stieltjes integrability

In order to tell whether f is Riemann-Stieltjes integrable we have to know $\int_a^b f(x)\, d\alpha(x)$. The idea of a Cauchy sequence leads to the following Theorem, which gives an equivalent definition.

Theorem: Cauchy criterion for Riemann-Stieltjes Integrability

A function defined on [a, b] is Riemann-Stieltjes integrable over [a, b] with respect to α, defined on [a, b], if and only if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all partitions π and π' of [a, b], and for all selection vectors ξ and ξ' associated with π and π', respectively,

$$\text{mesh}(\pi) < \delta \text{ and } \text{mesh}(\pi') < \delta \Rightarrow |RS(f, \alpha, \pi, \xi) - RS(f, \alpha, \pi', \xi')| < \varepsilon.$$

Proof: First we suppose that f is Riemann-Stieltjes integrable over [a, b] with respect to α. Then, using $\varepsilon/2$ in the definition of Riemann-Stieltjes integrability, we obtain $\delta > 0$ and RSI such that for all partitions π of [a, b],

$$\text{mesh}(\pi) < \delta \Rightarrow |RS(\pi) - RSI| < \varepsilon/2$$

Now we suppose that π and π' are partitions of [a, b] and that

$$\text{mesh}(\pi) < \delta \text{ and } \text{mesh}(\pi') < \delta.$$

Then for all selection vectors ξ and ξ' associated with π and π', respectively,

$$|RS(f, \alpha, \pi, \xi) - RS(f, \alpha, \pi', \xi')| \leq |RS(\pi, \xi) - RSI| + |RSI - RS(\pi', \xi')| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

This completes half the proof.

Next we suppose that the Cauchy condition, given in the Theorem, is satisfied. We have to find a candidate for $\int_a^b f(x)\, d\alpha(x)$. We first construct a sequence of partitions of [a, b]. We let π_n denote the partition that divides [a, b] into n equal parts $\left(\pi_n \text{ has points } x_{ni} := a + i\dfrac{b-a}{n}\right)$. Finally we define selection vectors ξ_n by

$$\xi_{ni} := a + i\,\frac{b-a}{n},\ i = 1, ..., n \text{ and define } \sigma_n := \sum_{i=1}^{n} f(\xi_{ni})(\alpha(x_{ni}) - \alpha(x_{n,i-1})),$$

a Riemann-Stieltjes sum ($\sigma_n = RS(f, \alpha, \pi_n, \xi_n)$). Now, given $\varepsilon > 0$, we use $\varepsilon/2$ in the Cauchy criterion, and obtain $\delta > 0$ such that

$$\text{mesh}(\pi) < \delta \text{ and } \text{mesh}(\pi') < \delta \Rightarrow |RS(f, \alpha, \pi, \xi) - RS(f, \alpha, \pi', \xi')| < \varepsilon/2.$$

Then, if n and n′ are so large that $(b - a)/n < \delta$ and $(b - a)/n' < \delta$, we have

$$\text{mesh}(\pi_n) < \delta \text{ and } \text{mesh}(\pi_{n'}) < \delta \Rightarrow |\sigma_n - \sigma_{n'}| < \varepsilon/2.$$

This means (since ε was arbitrary) that $\{\sigma_n\}$ is a Cauchy sequence in our space. Thus we define

$$\text{RSI} := \lim_{n\to\infty} \sigma_n$$

and it remains to show that if $\pi | [a, b]$ then

$$\text{mesh}(\pi) < \delta \Rightarrow |\text{RS}(\pi) - \text{RSI}| < \varepsilon.$$

This is essentially done. We choose the first n such that $\text{mesh}(\pi_n) < \delta$, and we suppose that $\text{mesh}(\pi) < \delta$. Then

$$|\text{RS}(\pi) - \text{RI}| \leq |\text{RS}(\pi) - \sigma_n| + |\sigma_n - \text{RSI}| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

since $\text{RS}(\pi) - \sigma_n = \text{RS}(\pi) - \text{RS}(f, \alpha, \pi_n, \xi_n)$. The proof is complete.

> *Notes* Nothing is said at first about the functions f and α, beside the demand that the integrability definition hold.

If f and α have a discontinuity at the same point, then the Riemann-Stieltjes integral does not exist.

They also show that if the Riemann-Stieltjes integral exists, then this integration-by-parts formula holds:

$$\int_a^b f\, d\alpha = -\int_a^b \alpha\, df + f(b)\alpha(b) - f(a)\alpha(a)$$

(in the applications have far, far back in my mind, the integral on the right would have to be $\int_a^b df\, \alpha$, in order to keep the order of "multiplication" the same). The proof amounts to rearranging the Riemann-Stieltjes sums, adding and subtracting terms in such a way that the ξ_i become partition points and the x_i become selection-vector components when $1 < i < n_\pi$. There are some leftovers, and these turn out to be the "boundary" term $f(x)\alpha(x)|_a^b$.

Wheeden and Zygmund state several properties, routine to prove, about Riemann-Stieltjes integrals:

$$\int_a^b f\, d\alpha \text{ is linear in both f and } \alpha$$

as long as all the integrals involved exist, and if $\int_a^b f\, d\alpha$ exists and $a < c < b$, then both of $\int_a^c f\, d\alpha$ and $\int_c^b f\, d\alpha$ exist, and $\int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha$.

What has been covered applies to all Riemann-Stieltjes integrals. That continuity plays a role has already been mentioned.

5.5 Functions of Bounded Variation: Definition and Properties

In what we do from now on, at least one of f and α will be a function of bounded variation, unless otherwise stated. We will begin by discussing real-valued functions of bounded variation. This material can also be found in Measure and Integral, by Wheeden and Zygmund.

Definition: A function $f: [a, b] \to \mathbb{R}$ is a function of bounded variation on [a, b] if

$$V(f, [a, b]) := \sup_{\pi|[a,b]} \sum_1^{n_\pi} |f(x_i) - f(x_{i-1})| < \infty \text{ and we say that } f \in BV[a, b].$$

To go farther it will be useful to have some more notation. If π is a partition of [a, b] we will write

$$V_\pi = V_\pi(f, [a, b]) := \sum_1^{n_\pi} |f(x_i) - f(x_{i-1})|,$$

so that $V = \sup_{\pi|[a,b]} V_\pi$ (here, f and [a, b] are "assumed").

We can call V_π the "π-variation" of f over [a, b]. Since f has a finite value for each $\pi|[a, b]$, V_π is always finite. However, V can be infinite. This is so, for example, if f is the Dirichlet function.

Each V_π pays attention only to the absolute value of the difference between the values at the opposite ends of an interval of the partition π. We will need to take the signs of those differences into account, and they will lead to two new "variations."

For a real number x we define its positive part to be $x^+ := \max\{0, x\}$ and we define its negative part to be $x^- := \max\{0, -x\}$. Both "parts" are non-negative, and we have $x^+ + x^- = |x|$ and $x^+ - x^- = x$.

Example: Prove that for all real numbers x and y, $(x + y)^+ \le x^+ + y^+$ and $(x + y)^- \le x^- + y^-$. These are "triangle inequalities!" What can be said about $(xy)^+$ and $(xy)^-$?

We now define the "positive" and "negative" "π-variations" of f over [a, b]:

$$P_\pi = P_\pi(f, [a, b]) := \sum_1^{n_\pi} (f(x_i) - f(x_{i-1}))^+ \text{ and } N_\pi = N_\pi(f, [a, b]) := \sum_1^{n_\pi} (f(x_i) - f(x_{i-1}))^-.$$

Definition: The positive variation, P = P(f, [a, b]) and the negative variation N = N(f, [a, b]) of f over [a, b] are given by $P = \sup_{\pi|[a,b]} P_\pi$ and $N = \sup_{\pi|[a,b]} N_\pi$ respectively.

For example, if f increases on [a, b], $P_\pi = V_\pi = f(b) - f(a)$ and $N_\pi = 0$. If we look at $f(x) := |x|$ on [–1, 1] we will always have $0 \le P_\pi \le 1$ and $0 \le N_\pi \le 1$, and $0 \le V_\pi \le 2$.

Because of how x^+ and x^- were defined, we always have (for any function)

$$P_\pi + N_\pi = V_\pi \text{ and } P_\pi - N_\pi = f(b) - f(a)$$

If τ is a refinement of π, we always have $O_\pi \le O_\tau$, where O stands for any of the letters N, P or V. This follows from several applications of the triangle inequality.

5.6 Some Properties of Functions of Bounded Variation

If $f \in BV[a, b]$ then f is bounded on [a, b].

Proof: Suppose $a \le x \le b$. Then, if we let $\pi := \{a, x, b\}$,

$$|f(x)| = |f(x) - f(a) + f(a)| \le |f(a)| + |f(x) - f(a)| + |f(b) - f(x)| = |f(a)| + V_\pi \le |f(a)| + V.$$

The space BV[a, b] is a vector space. For all $c \in \mathbb{R}$ and all $f \in BV[a, b]$, $V(cf, [a, b]) = |c|V(f, [a, b])$. For all $f \in BV[a, b]$ and $g \in BV[a, b]$, $V(f, [a, b]) \le V(f, [a, b]) + V(g, [a, b])$; $V(f, [a, b]) = 0$ if and only if f is constant.

Proof: The second assertion follows from these facts: for all $\pi|[a, b]$, $V_\pi(cf, [a, b]) = |c|V_\pi(f, [a, b])$; $\sup\{|c|x: x \in E\} = |c| \sup\{x: x \in E\} = |c| \sup E$. The first assertion and the first part of the third one follow from the second one and the triangle inequality. Finally, suppose that $V(f, [a, b]) = 0$ and that $a \le x \le b$. Then, with $\pi := \{a, x, b\}$, $|f(x) - f(a)| \le |f(x) - f(a)| + |f(b) - f(x)| = V_\pi = 0$. Therefore $f(x) \equiv f(a)$.

If $f \in BV[a, b]$ and $a < c < b$ then $f \in BV[a, c]$ and $f \in BV[c, b]$, and conversely. Moreover, $V = V(f, [a, b]) = V(f, [a, c]) + V(f, [c, b])$.

Proof: If $f \in BV[a, b]$ and $a < c < b$, let partitions $\sigma | [a, c]$ and $\tau | [c, b]$ be given. Then $\pi := \sigma \cup \tau$ is a partition of $[a, b]$ so $V_\sigma + V_\tau = V_\pi \leq V$, hence $V_\sigma \leq V$ and $V_\tau \leq V$. Thus $f \in BV[a, c]$ and $f \in BV[c, b]$. Conversely, suppose that $a < c < b$ and that $f \in BV[a, c]$ and $f \in BV[c, b]$. Let $\pi | [a, b]$. Then $\pi_c := \pi \cup \{c\}$ is a refinement of π. Therefore $V_\pi \leq V_{\pi_c} = V_\sigma + V_\tau$, where $\sigma := \pi_c \cap [a, c]$ and τ is defined similarly. By hypothesis, $V_\pi \leq V_{\pi_c} = V_\sigma + V_\tau \leq V(f, [a, c]) + V(f, [c, b])$. Thus $V(f, [a, b]) \leq V(f, [a, c]) + V(f, [c, b]) < \infty$. This proves part of the asserted equality. To show the other inequality, now that we know $V < \infty$ let partitions $\sigma | [a, c]$ and $\tau | [c, b]$ be given. We recall that earlier we had $V_\sigma + V_\tau = V_{\pi_c} \leq V$, so $V_\sigma + V_\tau \leq V$ whenever $\sigma | [a, c]$ and $\tau | [c, b]$ were arbitrary partitions of $[a, c]$ and $[c, b]$, respectively.

Thus $\sup\limits_{\sigma|[a,c]} (V_\sigma + V_\tau) = V(f, [a, c]) + V_\tau \leq V$, and so $\sup\limits_{\tau|[c,b]} (V(f, [a, c]) + V_\tau) = V(f, [a, c]) + V(f, [c, b]) \leq V$.

Note The first inequality holds for an arbitrary $\tau | [c, b]$, making the second one valid.

Example: Prove that the equality in (25) holds for every function $f: [a, b] \to \mathbb{R}$, whether f is a function of bounded variation or not.

Motivated by (25), when $f: [a, b] \to \mathbb{R}$ and $a \leq x \leq b$ we can define the three functions

$$V(x) := V(f, [a, x]),\ P(x) := P(f, [a, x]) \text{ and } N(x) := N(f, [a, x]).$$

Each of these is an increasing function of x. Jordan's Theorem asserts that if $f \in BV[a, b]$ we can represent f in terms of P(x) and N(x).

Theorem: Jordan

A function $f \in BV[a, b]$ if and only if there exist functions g and h, both increasing on [a, b], such that $f(x) = g(x) - h(x)$ for $a \leq x \leq b$. If this is the case, then $P(x) \leq g(x) - g(a)$, $N(x) \leq h(x) - h(a)$ and $f(x) = f(a) + P(x) - N(x)$ for $a \leq x \leq b$.

Proof: Suppose first that $f(t) = g(t) - h(t)$, $t \in [a, b]$, where the functions g and h are both increasing on [a, b]. Let $\pi | [a, b]$ (later, we will apply this when $x \in [a, b]$ and $\pi | [a, x]$). Then

$$\Delta f_i = f(x_i) - f(x_{i-1}) = \Delta g_i - \Delta h_i \begin{cases} \leq \Delta g_i \\ \geq -\Delta h_i \end{cases}.$$

Thus $-\Delta h_i \leq \Delta f_i \leq \Delta g_i$, so $|\Delta f_i| \leq \max\{\Delta g_i, \Delta h_i\} \leq \Delta g_i + \Delta h_i$ for $1 \leq i \leq n_\pi$. Hence $V_\pi(f) \leq V_\pi(g) + V_\pi(h) = g(b) - g(a) + h(b) - h(a) < \infty$, so $f \in BV[a, b]$.

Next, we show that $f(x) = f(a) + P(x) - N(x)$ for $a \leq x \leq b$. But we will do this just by showing it for $x = b$. Then we can use (25) and let each $x \in [a, b]$ play the role of b. This will show the existence of the functions $g(x)(= f(a) + P(x))$ and $h(x)(= N(x))$. After that is done, we'll prove the P–g and N–h inequalities.

By the definitions of P, N and V we know there exist sequences $\{\pi_k\}$, $\{\rho_k\}$ and $\{\sigma_k\}$ such that $P_{\pi_k} \to P$, $N_{\pi_k} \to N$ and $V_{\sigma_k} \to V$. Let us define $\tau_k := \pi_k \cup \rho_k \cup \sigma_k$. As $P_{\pi_k} \leq P_{\tau_k} \leq P$. By the Squeeze Principle $P_{\tau_k} \to P$. Similarly, $N_{\tau_k} \to N$ and $V_{\tau_k} \to V$. By Limit Theorems

$P + N = V$ and $P - N = f(b) - f(a)$ and the second is the same as $f(x) = f(a) + P(x) - N(x)$

when $x = b$. By above we can use any $x \in [a, b]$ in place of b by restricting our attention to f on [a, x].

Now suppose that f(x) is defined as the difference of two increasing functions on [a, b]: $f(t) = g(t) - h(t)$. We have the following observation: $t_1 \to t^+$ is increasing and $t_1 \to t^-$ is decreasing.

Therefore, with the help of (28), applied to partitions of [a, x], $(\Delta f_i)^+ \leq (\Delta g_i)^+ = \Delta g_i$ and $\Delta h_i = (-\Delta h_i)^- \geq (\Delta f_i)^-$.

Hence $P_\pi(f, [a, x]) \leq P_\pi(g, [a, x]) = g(x) - g(a)$. Similarly, $h(x) - h(a) = N_\pi(-h, [a, x]) \geq N_\pi(f, [a, x])$. When, in each case, we take the supremum over all $\pi \mid [a, x]$, we get $P(x) \leq g(x) - g(a)$ and $N(x) \leq h(x) - h(a)$. These "say" that there is no "wasted cancellation" in the formula $f(x) = f(a) + P(x) - N(x)$.

Task Prove that if $f(x) \in BV[a, b]$ and $f(x)$ is continuous at $x_o \in [a, b]$ then so are P(x), N(x) and V(x).

Self Assessment

Fill in the blanks:

1. A .... f(x) defined on the bounded and closed interval [a, b] is Riemann-Stieltjes integrable on [a, b] with respect to α if there exists a number RSI such that for all ε > 0 there exists δ > 0 such that for every partition π of [a, b],

$$\text{mesh}(\pi) < \delta \Rightarrow |RS(f, \pi) - RSI| < \varepsilon.$$

2. If f is Riemann integrable on [a, b] then f is ....
3. If α(x) is defined for x ∈ [a, b], we denote by Ω = Ω(α, [a, b]) the set of all c ∈ [a, b] such that every open interval U that contains c contains $x_1 < c < x_2$ with ....
4. If f is .... on [a, b] with respect to α then f is bounded on Ω(α, [a, b]).
5. If f and α have a .... at the same point, then the Riemann-Stieltjes integral does not exist.

5.7 Summary

- A Riemann-Stieltjes sum for a function f(x) defined on an interval [a, b] is formed with the help of
  (a) A partition π of [a, b], namely an ordered, finite set of points x_i, with $a = x_0 < x_1 < \cdots < x_n = b$ (where n is a positive integer that can be any positive integer, and one that we will often write as $n = n_\pi$),
  (b) A selection vector $\xi = (\xi_1, ..., \xi_n)$ that has n_π components that must satisfy $x_{i-1} \leq \xi_i \leq x_i$, for i = 1, 2, ..., n. and
  (c) An integrator α(x), which is a function defined on [a, b] that plays the role of the x in dx ...

  A Riemann-Stieltjes sum for f over [a, b] with respect to the partition π, using the selection vector ξ, and integrator α, may be denoted (in greatest detail!) as follows, and it is given by the value of the sum following it:

  (d) $RS(f, \alpha, [a, b], \pi, \xi) := \sum_{i=1}^{n_\pi} f(\xi_i)(\alpha(x_i) - \alpha(x_{i-1}))$.

  We try to allow context to let us drop some of the items inside the RS(...).
- In this definition, as in the Riemann-sums definition, we can write $\Delta x_i := x_i - x_{i-1}$ or $\Delta\alpha_i := \alpha(x_i) - \alpha(x_{i-1})$. These are convenient because they are short and suggest the dx or dα in an integral. But they can cause confusion because they leave out the dependence they have on x_{i-1}. The Δx_i is used in the Riemann-Stieltjes context.

Notes

- A partition π can be thought of as "dividing" the interval [a, b] into subintervals. We may write $\pi | [a, b]$ and read this as "π divides [a, b]," or "partitions [a, b]." We will denote the intervals of π by I_i: = $[x_{i-1}, x_i]$. When we wish to work with 2 partitions at the same time we will have to distinguish between them somehow, for example we can use y_j to denote the other's points and J_j to denote its intervals, etc.
- A real-valued function f(x) defined on the bounded and closed interval [a, b] is Riemann-Stieltjes integrable on [a, b] with respect to α if there exists a number RSI such that for all $\varepsilon > 0$ there exists $\delta > 0$ such that for every partition π of [a, b],

$$\text{mesh}(\pi) < \delta \Rightarrow |RS(f, \pi) - RSI| < \varepsilon.$$

5.8 Keywords

Cauchy Criterion for Riemann-Stieltjes Integrability: A function defined on [a, b] is Riemann-Stieltjes integrable over [a, b] with respect to α, defined on [a, b], if and only if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all partitions π and π' of [a, b], and for all selection vectors ξ and ξ' associated with π and π', respectively,

$$\text{mesh}(\pi) < \delta \text{ and mesh}(\pi') < \delta \Rightarrow |RS(f, \alpha, \pi, \xi) - RS(f, \alpha, \pi', \xi')| < \varepsilon.$$

Jordan: A function $f \in BV[a, b]$ if and only if there exist functions g and h, both increasing on [a, b], such that $f(x) = g(x) - h(x)$ for $a \le x \le b$. If this is the case, then $P(x) \le g(x) - g(a)$, $N(x) \le h(x) - h(a)$ and $f(x) = f(a) + P(x) - N(x)$ for $a \le x \le b$.

5.9 Review Questions

1. Identify the properties of the integral.
2. Use them to find the Riemann stieltjes integral of functions.

Answers: Self Assessment

1. real-valued function
2. bounded on [a, b]
3. $|\alpha(x_1) - \alpha(x_2)| > 0$
4. Riemann-Stieltjes integrable
5. discontinuity

5.10 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Notes

Unit 6: Differentiation of Integrals

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define Differentiation of Integrals
- Discuss the Theorems on the Differentiation of Integrals

Introduction

In this unit, we are going to study about differentiation of integrals. Suppose $\vee$ is a function of two variables which can be integrated with respect to one variable and which can be differentiated with respect to another variable. We are going to see under what conditions the result will be the same if these two limit process are carried out in the opposite order.

6.1 Differentiation of Integrals

In mathematics, the problem of differentiation of integrals is that of determining under what circumstances the mean value integral of a suitable function on a small neighbourhood of a point approximates the value of the function at that point. More formally, given a space X with a measure μ and a metric d, one asks for what functions $f : X \to R$ does

$$\lim_{r \to u} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y)\,d\mu(y) = f(x)$$

for all (or at least μ-almost all) $x \in X$? (Here, as in the rest of the article, $B_r(x)$ denotes the open ball in X with d-radius r and centre x.) This is a natural question to ask, especially in view of the heuristic construction of the Riemann integral, in which it is almost implicit that f(x) is a "good representative" for the values of f near x.

6.2 Theorems on the Differentiation of Integrals

Lebesgue Measure

One result on the differentiation of integrals is the Lebesgue differentiation theorem, as proved by Henri Lebesgue in 1910. Consider n-dimensional Lebesgue measure λ^n on n-dimensional Euclidean space R^n. Then, for any locally integrable function $f : R^n \to R$, one has

$$\lim_{r \to 0} \frac{1}{\lambda^n(B_r(x))} \int_{B_r(x)} f(y)\, d\lambda^n(y) = f(x)$$

for λ^n-almost all points $x \in R^n$. It is important to note, however, that the measure zero set of "bad" points depends on the function f.

Borel Measures on R^n

The result for Lebesgue measure turns out to be a special case of the following result, which is based on the Besicovitch covering theorem: if μ is any locally finite Borel measure on R^n and $f : R^n \to R$ is locally integrable with respect to μ, then

$$\lim_{r \to 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y)\, d\mu(y) = f(x)$$

for μ-almost all points $x \in R^n$.

Gaussian Measures

The problem of the differentiation of integrals is much harder in an infinite-dimensional setting. Consider a separable Hilbert space $(H, \langle , \rangle)$ equipped with a Gaussian measure γ. As stated in the article on the Vitali covering theorem, the Vitali covering theorem fails for Gaussian measures on infinite-dimensional Hilbert spaces. Two results of David Preiss (1981 and 1983) show the kind of difficulties that one can expect to encounter in this setting:

- There is a Gaussian measure γ on a separable Hilbert space H and a Borel set $M \subseteq H$ so that, for γ-almost all $x \in H$,

$$\lim_{r \to 0} \frac{\lambda(M \cap B_r(x))}{\gamma(B_r(x))} = 1$$

- There is a Gaussian measure γ on a separable Hilbert space H and a function $f \in L^1(H, \gamma; R)$ such that

$$\lim_{r \to 0} \inf \left\{ \frac{1}{\gamma(B_s(x))} \int_{Ds(x)} f(y)\, d\gamma(y) \middle| x \in II, 0 < s < r \right\} - +\infty$$

However, there is some hope if one has good control over the covariance of γ. Let the covariance operator of γ be $S : H \to H$ given by

$$\langle Sx, y \rangle = \int_H \langle x, z \rangle \langle y, z \rangle\, d\gamma(z)$$

or, for some countable orthonormal basis $(e_i)_{i \in N}$ of H,

$$Sx = \sum_{i \in N} \sigma_1^2 \langle x, e_i \rangle e_i.$$

In 1981, Preiss and Jaroslav Tišer showed that if there exists a constant $0 < q < 1$ such that

$$\sigma_{i+1}^2 \le q\sigma_i^2,$$

then, for all $f \in L^1(H, \gamma; R)$,

$$\frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y)\, d\mu(y) \xrightarrow[r \to 0]{\gamma} f(x)$$

where the convergence is convergence in measure with respect to γ. In 1988, Tišer showed that if

$$\sigma_{i+1}^2 \leq \frac{\sigma_i^2}{i^\alpha}$$

for some α > 5/2, then

$$\frac{1}{\mu(B_r(x))}\int_{B_r(x)} f(y)\,d\mu(y) \xrightarrow[r\to 0]{} f(x)$$

for γ-almost all x and all f ∈ $L^p(H, \gamma; R)$, p > 1.

As of 2007, it is still an open question whether there exists an infinite-dimensional Gaussian measure γ on a separable Hilbert space H so that, for all f ∈ $L^1(H, \gamma; R)$,

$$\lim_{r\to 0} \frac{1}{\gamma(B_r(x))}\int_{B_r(x)} f(y)\,d\gamma(y) = f(x)$$

for γ-almost all x ∈ H. However, it is conjectured that no such measure exists, since the σ_i would have to decay very rapidly.

Example: If $\alpha \neq 0, \phi(\alpha) = \arctan\left(\frac{1}{\alpha}\right)$

The function $\frac{\alpha}{x^2+\alpha^2}$ is not continuous at the point (x, α) = (0, 0) and the function ϕ(α) has a discontinuity α = 0, because ϕ(α) approaches $+\frac{\pi}{2}$ as $\alpha \to 0^+$ and approaches $-\frac{\pi}{2}$ as $\alpha \to 0^-$.

If we now differentiate ϕ(α) = $\int_0^1 \frac{\alpha}{x^2+\alpha^2}dx$ with respect to α under the integral sign, we get $\frac{d}{d\alpha}\phi(\alpha) = \int_0^1 \frac{x^2-\alpha^2}{x^2+\alpha^2}dx = -\frac{x}{x^2+a^2}\Big|_0^1 = -\frac{1}{1+\alpha^2}$ which is, of course, true for all values of α except α = 0.

Example: The principle of differentiating under the integral sign may sometimes be used to evaluate a definite integral.

Consider integrating $\phi(\alpha) = \int_0^\pi \ln(1-2\alpha\cos(x)+\alpha^2)dx \ (\text{for} |\alpha| > 1)$

Now,

$$\frac{d}{d\alpha}\phi(\alpha) = \int_0^\pi \frac{-2\cos(x)+2\alpha}{-2\alpha\cos(x)+\alpha^2}dx$$

$$= \frac{1}{\alpha}\int_0^\pi \left(1 - \frac{(1-\alpha)^2}{1-2\alpha\cos(x)+\alpha^2}\right)dx$$

$$= \frac{\pi}{\alpha} - \frac{2}{\alpha}\left\{\arctan\left(\frac{1+\alpha}{1-\alpha}\cdot\tan\left(\frac{x}{2}\right)\right)\right\}\Bigg|_0^\pi$$

Notes

As x varies from 0 to π, $\left(\frac{1+\alpha}{1-\alpha}\cdot\tan\left(\frac{x}{2}\right)\right)$ varies through positive values from 0 to ∞ when –1 < α < 1 and $\left(\frac{1+\alpha}{1-\alpha}\cdot\tan\left(\frac{x}{2}\right)\right)$ and varies through negative values from 0 to –∞ when α < –1 or α > 1.

Hence,

$$\arctan\left(\frac{1+\alpha}{1-\alpha}\cdot\tan\left(\frac{x}{2}\right)\right)\Bigg|_0^\pi = -\frac{\pi}{2} \text{ when } -1<\alpha<1$$

and

$$\arctan\left(\frac{1+\alpha}{1-\alpha}\cdot\tan\left(\frac{x}{2}\right)\right)\Bigg|_0^\pi = -\frac{\pi}{2} \text{ when } \alpha<-1 \text{ or } \alpha>1.$$

Therefore,

$$\frac{d}{d\alpha}\phi(\alpha) = 0 \text{ when } -1<\alpha<1$$

$$\frac{d}{d\alpha}\phi(\alpha) = \frac{2\pi}{\alpha} \text{ when } \alpha<-1 \text{ or } \alpha>1\,.$$

Upon integrating both sides with respect to α, we get $\phi(\alpha) = C_1$ when –1 < α < 1 and $\phi(\alpha) = 2\pi$ In $|\alpha| + C_2$ when α < –1 or α > 1.

C_1 may be determined by setting α = 0 in

$$\phi(\alpha) = \int_0^\pi \ln(1-2\alpha\cos(x)+a^2)dx$$

$$\phi(0) = \int_0^\pi \ln(1)dx$$

$$= \int_0^\pi 0\,dx$$

$$= 0$$

Thus, $C_1 = 0$. Hence, $\phi(\alpha) = 0$ when –1 < α < 1.

To determine C_2 in the same manner, we should need to substitute in $\phi(\alpha) = \int_0^\pi \ln(1-2\alpha\cos(x)+\alpha^2)dx$ a value of α greater numerically than 1. This is somewhat inconvenient. Instead, we substitute, $\alpha = \frac{1}{\beta}$, where $-1<\beta<1$. Then,

$$\phi(\alpha) = \int_0^\pi \ln\left(1-2\beta\cos(x)+\beta^2\right) - 2\ln|\beta|\Big)dx$$

$$= 0 - 2\pi \ln |\beta|$$

$$= 2\pi \ln |\alpha|$$

Therefore, $C_2 = 0$ and $\phi(\alpha) = 2\pi \ln |\alpha|$ when α < –1 or α > 1.)

The definition of $\phi(\alpha)$ is now complete:

$$\phi(\alpha) = 0 \text{ when } -1<\alpha<1 \text{ and}$$

$$\phi(\alpha) = 2\pi\ln|\alpha| \text{ when } \alpha<-1 \text{ or } \alpha>1$$

Notes

The foregoing discussion, of course, does not apply when $\alpha = \pm 1$ since the conditions for differentiability are not met.

Example: Here, we consider the integration of

$$I = \int_0^{\frac{\pi}{2}} \frac{1}{\left(a\cos^2 x + b\sin^2 x\right)} dx$$

where both a, b > 0, by differentiating under the integral sign.

Let us first find $J = \int_0^{\frac{\pi}{2}} \frac{1}{a\cos^2 x + b\sin^2 x} dx$

Dividing both the numerator and the denominator by $\cos^2 x$ yields

$$J = \int_0^{\frac{\pi}{2}} \frac{\sec^2 x}{a + b\tan^2 x} dx$$

$$= \frac{1}{b}\int_0^{\frac{\pi}{2}} \frac{1}{\left(\sqrt{\frac{a}{b}}\right)^2 + \tan^2 x} d(\tan x)$$

$$= \frac{1}{\sqrt{a, b}}\left(\tan^{-1}\left(\sqrt{\frac{b}{a}}\tan x\right)\right)\Bigg|_0^{\frac{\pi}{2}} = \frac{\pi}{2\sqrt{a, b}}.$$

The limits of integration being independent of a, $J = \int_0^{\frac{\pi}{2}} \frac{1}{a\cos^2 x + b\sin^2 x} dx$ gives us

$$\frac{\partial J}{\partial a} = -\int_0^{\frac{\pi}{2}} \frac{\cos^2 x dx}{\left(a\cos^2 x + b\sin^2 x\right)^2}$$

Whereas $J = \frac{\pi}{2\sqrt{ab}}$ gives us

$$\frac{\partial J}{\partial a} = -\frac{\pi}{4\sqrt{a^3 b}}.$$

Equating these two relations then yields

$$\int_0^{\frac{\pi}{2}} \frac{\cos^2 x\, dx}{\left(a\cos^2 x + b\sin^2 x\right)^2} = \frac{\pi}{4\sqrt{a^3 b}}$$

In a similar fashion, $\frac{\partial J}{\partial b}$ pursuing yields

$$\int_0^{\frac{\pi}{2}} \frac{\sin^2 x\, dx}{\left(a\cos^2 x + b\sin^2 x\right)^2} = \frac{\pi}{4\sqrt{ab^3}}$$

Adding the two results then produces

$$I = \int_0^{\frac{\pi}{2}} \frac{1}{\left(a\cos^2 x + b\sin^2 x\right)^2} dx = \frac{\pi}{4\sqrt{ab}}\left(\frac{1}{a} + \frac{1}{b}\right)$$

Which is the value of the integral I.

Note that if we define

Notes

$$I_n = \int_0^{\frac{\pi}{2}} \frac{1}{\left(a\cos^2 x + b\sin^2 x\right)^n} dx$$

it can easily be shown that

$$\frac{\partial I_{n-1}}{\partial a} + \frac{\partial I_{n-1}}{\partial b} + (n-1)\cdot I_n = 0.$$

Given I_1 this *partial-derivative-based* recursive relation (i.e., integral reduction formula) can then be utilized to compute all of the values of I_n for $n > 1$ (I_1, I_2, I_3, I_4, etc.).

Example: Here, we consider the integral

$$I(\alpha) = \int_0^{\frac{\pi}{2}} \frac{\ln(1+\cos\alpha\,\cos x)}{\cos x} dx\,.$$

for $0 < \alpha < \pi$.

Differentiating under the integral with respect to α, we have

$$\frac{d}{d\alpha} I(\alpha) = \int_0^{\frac{\pi}{2}} \frac{\partial}{\partial\alpha}\left(\frac{\ln(1+\cos\alpha\,\cos x)}{\cos x}\right) dx$$

$$= -\int_0^{\frac{\pi}{2}} \frac{\sin\alpha}{1+\cos\alpha\,\cos x}\, dx$$

$$= -\int_0^{\frac{\pi}{2}} \frac{\sin\alpha}{\left(\cos^2\frac{\pi}{2} + \sin^2\frac{x}{2}\right) + \cos\alpha\left(\cos^2\frac{x}{2} - \sin^2\frac{x}{2}\right)}\, dx$$

$$= -\frac{\sin\alpha}{1-\cos\alpha}\int_0^{\frac{\pi}{2}} \frac{1}{\cos^2\frac{x}{2}} \frac{\sin\alpha}{\left[\left(\frac{1+\cos\alpha}{1-\cos\alpha}\right) + \tan^2\frac{x}{2}\right]} dx$$

$$= -\frac{2\sin\alpha}{1-\cos\alpha}\int_0^{\frac{\pi}{2}} \frac{\frac{1}{2}\sec^2\frac{x}{2}}{\left[\left(\frac{2\cos^2\frac{\alpha}{2}}{2\sin^2\frac{\alpha}{2}}\right) + \tan^2\frac{x}{2}\right]} dx$$

$$= -\frac{2\left(2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\right)}{2\sin\frac{\alpha}{2}}\int_0^{\frac{\pi}{2}} \frac{1}{\left[\left(\frac{\cos\frac{\alpha}{2}}{\sin\frac{\alpha}{2}}\right)^2 + \tan^2\frac{x}{2}\right]} d\left(\tan\frac{x}{2}\right)$$

$$= -2\cot\frac{\alpha}{2}\int_0^{\frac{\pi}{2}} \frac{1}{\left[\cot^2\frac{\alpha}{2} + \tan^2\frac{x}{2}\right]} d\left(\tan\frac{x}{2}\right)$$

$$= -2\left(\tan-1\left(\tan\frac{\alpha}{2}\tan\frac{x}{2}\right)\right)\Bigg|_0^{\frac{\pi}{2}}$$

$$= -\alpha$$

Now, when $\alpha = \frac{\pi}{2}$, we have, from

$$I(\alpha) = \int_0^{\frac{\pi}{2}} \frac{\ln(1+\cos\alpha\,\cos x)}{\cos x}\,dx,\ I\left(\frac{\pi}{2}\right) = 0$$

Hence,

$$I(\alpha) = \int_{\frac{\pi}{2}}^{\alpha} -\alpha\,d\alpha$$

$$= -\frac{1}{2}\alpha^2\Big|_{\frac{\pi}{2}}^{\alpha}$$

$$= \frac{\pi^2}{8} - \frac{\alpha^2}{2},$$

which is the value of the integral $I(\alpha)$.

Example: Here, we consider the integral $\int_0^{2\pi} e^{\cos\theta}\cos(\sin\theta)\,d\theta$.

We introduce a new variable ϕ, and rewrite the integral as

$$f(\phi) = \int_0^{2\pi} e^{\phi\cos\theta}\cos(\phi\sin\theta)\,d\theta$$

Note that for $\phi = 1,\ f(\phi) = f(1) = \int_0^{2\pi} e^{\cos\theta}\cos(\sin\theta)\,d\theta$

Thus, we proceed

$$\frac{df}{d\phi} = \int_0^{2\pi} \frac{\partial}{\partial\phi}\left(e^{\phi\cos\theta}\cos(\phi\sin\theta)\right)d\theta$$

$$= \int_0^{2\pi} e^{\phi\cos\theta}(\cos\theta\cos(\phi\sin\theta) - \sin\theta\sin(\phi\sin\theta))\,d\theta$$

$$= \int_0^{2\pi} \frac{1}{\phi}\frac{\partial}{\partial\theta}\left(e^{\phi\cos\theta}\sin(\phi\sin\theta)\right)d\theta$$

$$= \frac{1}{\phi}\int_0^{2\pi} d\left(e^{\phi\cos\theta}\sin(\phi\sin\theta)\right)$$

$$= \frac{1}{\phi}\left(e^{\phi\cos\theta}\sin(\phi\sin\theta)\right)\Big|_0^{2\pi}$$

$$= 0.$$

From the equation for $f(\phi)$ we can see $f(0) = 2\pi$. So, integrating both sides of $\frac{df}{d\phi} = 0$ with respect to ϕ between the limits 0 and 1, yields

$$\int_{f(0)}^{f(1)} df = \int_0^1 d\phi = 0$$

$$\Rightarrow \qquad f(1) - f(0) = 0$$

$\Rightarrow \quad f(1) - 2\pi = 0$

$\Rightarrow \quad f(1) = 2\pi.$

which is the value of the integral $\int_0^{2\pi} e^{\cos\theta}\cos(\sin\theta)d\theta$.

Example: Find $\frac{d}{dx}\int_{\sin x}^{\cos x}\cosh t^2 dt$.

In this example, we shall simply apply the above given formula, to get

$$\frac{d}{dt}\int_{\sin x}^{\cos x}\cosh t^2 dt = \cosh(\cos^2 x)\frac{d}{dx}(\cos x) - \cosh(\sin^2 x)\frac{d}{dx}(\sin x) +$$

$$\int_{\sin x}^{\cos x}\frac{\partial}{\partial x}\cosh t^2 dt = -\cosh(\cos^2 x)\sin x - \cosh(\sin^2 x)\cos x$$

Where the derivative with respect to x of hyperbolic cosine t squared is 0. This is a simple example on how to use this formula for variable limits.

Self Assessment

Fill in the blanks:

1. The differentiation of integrals is the Lebesgue differentiation theorem, as proved by Henri Lebesgue in ....
2. The result for .... turns out to be a special case of the following result, which is based on the Besicovitch covering theorem.
3. The problem of .... is that of determining under what circumstances the mean value integral of a suitable function on a small neighbourhood of a point approximates the value of the function at that point.
4. The problem of the differentiation of integrals is much harder in an infinite-dimensional setting. Consider a separable Hilbert space (H, ⟨, ⟩) equipped with a ....

6.3 Summary

- In mathematics, the problem of differentiation of integrals is that of determining under what circumstances the mean value integral of a suitable function on a small neighbourhood of a point approximates the value of the function at that point.
- One result on the differentiation of integrals is the Lebesgue differentiation theorem, as proved by Henri Lebesgue in 1910. Consider n-dimensional Lebesgue measure λ^n on n-dimensional Euclidean space R^n.
- The result for Lebesgue measure turns out to be a special case of the following result, which is based on the Besicovitch covering theorem: if μ is any locally finite Borel measure on R^n and $f : R^n \to R$ is locally integrable with respect to μ, then

$$\lim_{r\to 0}\frac{1}{\mu(B_r(x))}\int_{B_r(x)} f(y)d\mu(y) = f(x) \text{ for } \mu\text{-almost all points } x \in R^n.$$

- The problem of the differentiation of integrals is much harder in an infinite-dimensional setting. Consider a separable Hilbert space (H, ⟨, ⟩) equipped with a Gaussian measure γ.

6.4 Keywords

Differentiation of Integrals: In mathematics, the problem of differentiation of integrals is that of determining under what circumstances the mean value integral of a suitable function on a small neighbourhood of a point approximates the value of the function at that point.

Borel measures on R^n: The result for Lebesgue measure turns out to be a special case of the following result, which is based on the Besicovitch covering theorem: if μ is any locally finite Borel measure on R^n and $f : R^n \rightarrow R$ is locally integrable with respect to μ, then

$$\lim_{r \to 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y) d\mu(y) = f(x)$$

for μ-almost all points $x \in R^n$.

6.5 Review Questions

1. Explain Differentiation of Integrals with the help of example.
2. Discuss the Theorems on the differentiation of integrals.

Answers: Self Assessment

1. 1910
2. Lebesgue measure
3. differentiation of integrals
4. Gaussian measure γ

6.6 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 7: Fundamental Theorem of Calculus

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the fundamental theorem of calculus
- Explain the primitive of a function

Introduction

In this unit we will discuss about, what is the relationship between the two notions of differentiation and integration? Now we shall try to find an answer to this question. In fact, we shall show that differentiation and Integration are intimately related in the sense that they are inverse operations of each other.

Let us begin by asking the following question: "when is a function f : [a, b] → R, the derivative of some function F : [a, b] → R?"

For example consider the function f : [–1, 1] → R defined by

$$f(x) = \begin{cases} 0 \text{ if } -1 \le x < 0 \\ i \text{ if } 0 \le x < 1 \end{cases}$$

This function is not the derivative of any function F : [–1, 1] → R. Indeed if f is the derivative of a function F : [–1, 1] → R then f must have the intermediate value property. But clearly, the function f given above does not have the intermediate value property.

Hence f cannot be the derivative of any function F : [–1, 1] → R.

However if f : [–1, 1] → R is continuous, then f is the derivative of a function F : [–1, 1] → R. This leads us to the following general theorem.

7.1 Fundamental Theorem of Calculus

Theorem 1: Let f be integrable on [a, b]. Define a function P on [a, b] as

$$F(x) = \int_a^x f(t)\, dt, \forall\ x \in [a, b].$$

Then F is continuous on [a, b]. Furthermore, if f is continuous at a point x, of [a, b], then F is differentiable at x_0 and $F'1(x_0) = f(x_0)$.

Notes

Proof: Since f is integrable on [a,b], it is bounded. In other words, there exists a positive number M such that $|f(x)| \le M, \forall\ x \in [a,b]$.

Let ε > 0 be any number. Choose x,y ∈ [a,b], x < y, such that $|x-y| < \frac{e}{M}$. Then

$$|F(y)-F(x)| = \left|\int_a^y f(t)\,dt - \int_a^x f(t)\,dt\right|$$

$$= \left|\int_a^x f(t)\,dt + \int_x^y f(t)\,dt - \int_a^x f(t)\,dt\right|$$

$$= \left|\int_x^y f(t)\,dt\right|$$

$$\le \int_x^y |f(t)|\,dt$$

$$\le \int_x^y M dt = M(y-x) < \in$$

Similarly you can discuss the case when y < x. This shows that F is continuous on [a,b]. In fact this proves the uniform continuity of F.

Now, suppose f is continuous at a point x_0 of [a, b]

We can choose some suitable h ≠ 0 such that x_0 + h ∈ [a, b].

Then,

$$F(x_0 + h) - F(x_0) = \int_a^{x_0+h} f(t)\,dt - \int_a^{x_0} f(t)\,dt$$

$$= \int_a^{x_0} f(t)\,dt + \int_{x_0}^{x_0+h} f(t)\,d(t) - \int_a^{x_0} f(t)\,dt = \int_{x_0}^{x_0+h} f(t)\,dt$$

Thus,

$$F(x_0 + h) - F(x_0) = \int_{xn}^{x_0+h} f(t)\,dt \qquad \ldots(1)$$

Now

$$\left|\frac{F(x_0+h)-F(x_0)}{h} - f(x_0)\right| = \left|\frac{1}{h}\int_{x0}^{x0+h} f(t)\,dt - \frac{1}{h} \times \int_0^{x_0+h} f(x_0)\,dt\right|$$

$$= \frac{1}{|h|}\left|\int_{x_0}^{x_0+h} [f(t)-f(x_0)]dt\right|.$$

Since f is continuous at x_0, given a number ∈ >0, 3 a number δ > 0 such that $|f(x)-f(x_0)| < \varepsilon/2$, whenever $|x-x_0| < \delta$ and x ∈ [a,b]. So, if $|h| < \delta$, then $|f(t)-f(x_0)| < \in/2$, for $t \in [x_0, x_0+h]$, and consequently

$$\left|\int_{x_0}^{x_0+h}[f(t)-f(x_0)]dt\right| \le -|h|.\ \text{Therefore}$$

$$\left|\frac{F(x_0+h)-F(x_0)}{h}-f(x_0)\right| \le \frac{\varepsilon}{2} < \varepsilon, \text{ if } |h| < \delta.$$

Therefore, $\lim_{h\to 0}\frac{F(x_0+h)-F(x)}{h}f(x_0)$, i.e., $F'(x_0)=f(x_0)$

which shows that F is differentiable at x_0 and $F'(x_0) = f(x_0)$. From Theorem 1, you can easily deduce the following theorem:

Theorem 2: Let f: [a, b] → R be a continuous function. Let F : [a, b] → R be a function defined by

$$F(x) = \int_a^x f(t)\,dt,\ x\ E[a,b].$$

Then F'(x) = f(x), a ≤ x ≤ b.

This is the first result which links the concepts of integral and derivative. It says that, if f is continuous on [a,b] then there is a function F on [a, b] such that $F'(x) = f(x), \forall\ x \in [a,b]$.

You have seen that if f: [a, b] → R is continuous, then there is a function F: [a, b] → R such that F' (x) = f(x) on [a, b]. Is such a function F unique? Clearly the answer is 'no'. For, if you add a constant to the function F, the derivative is not altered. In other words, if $G(x) = c + \int_1^x f(t)\,dt$ for a ≤ x ≤ b then also G' (x) = f(x) on [a, b].

Such a function F or G is called primitive off. We have the formal definition as follows:

7.2 Primitive of a Function

If f and F are functions defined on [a,b] such that F′(x) = f(x) for x ∈ [a,b] then F is called a 'primitive' or an 'antiderivative' of f on [a,b].

Thus from Theorem 1, you can see that every continuous function on [a,b] has a primitive. Also there are infinitely many primitives, in the sense that adding a constant to a primitive gives another primitive.

"Is it true that any two primitives differ by a constant?"

The answer to this question is yes. Indeed if F and G are two primitives of f in [a,b], then $F'(x) = G'(x) = f(x)\ \forall\ x \in [a,b]$ and therefore [F(x) – G(x)′ = 0. Thus F(x) – G(x) = k (constant), for x ∈ [a,b].

Let us consider an example.

Example: What is the primitive of f(x) = log x in [1, 2]

Solution: Since $\frac{d}{dx}(x\log x - x) = \log x\ \forall\ x \in [1,2]$, therefore F (x) = x log x – x is a primitive of f in [1, 2].

Also G(x) = x lag x – x + k, k being a constant, is a primitive of f.

According to this theorem, differentiation and integration are inverse operations.

We now discuss a theorem which establishes the required relationship between differentiation and integration. This is called the Fundamental Theorem of Calculus.

It states that the integral of the derivative of a function is given by the function itself.

The Fundamental Theorem of Calculus was given by an English mathematician Isaac Barrow [1630-1677], the teacher of great Isaac Newton.

Theorem 3: Fundamental Theorem of Calculus

If f is integrable on [a,b] and F is a primitive of f on [a,b], then $\int_a^{tb} f(x)\,dx = F(b) - F(a).$

Proof: Since f $\in$ R [a,b], therefore $\lim_{|P|-0} S(P,f) = \int_a^b f(x)\,dx$

where P = {x_0, x_1, x_2,...., x_n] is a partition of [a,b]. The Riemann sum S(P,f) is given by

$$S(P,f) = \sum_{i=1}^{n} f(t_i)\,\Delta x_i = \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1});\ x_i - 1 \le t_i \le t_i.$$

Since F is the primitive of f on [a, b], therefore F' (x) $\le$ f(x), x $\in$ [a, b].

Hence $S(P,f) = \sum_{i=1}^{n} F'(t_i)(x_i - x_{i-1}).$ We choose the points t, as follows:

By Lagrange's Mean Value theorem of Differentiability, there is a point t, in]x_{i-1}, x_i[such that

$$F(x_i) - F(x_{i-1}) = F'\,(t_i)\,(x, - x_{i-1})$$

Therefore, $S(P,f) = \sum_{i=1}^{n} [F(x_i) - F(x_{i-1})] = F(x_n) - F(x_0) = F(b) - F(a).$

Take the limit as $|P| \to 0.$ Then $\int_a^b f(x)\,dx = F(b) - F(a).$ This completes the proof.

Alternatively, the Fundamental Theorem of Calculus is also interpreted by stating that the derivative of the integral of a continuous function is the function itself.

If the derivative f of a function f is integrable on [a, b], then $\int_a^b f'(x)\,dx = f(b) - f(a).$

Applying this theorem, we can find the integral of various functions very easily.

Consider the following example:

Example: Show that $\int_0^t \sin x\,dx = 1 - \cos t.$

Solution: Since g(x) = – cos x is the primitive of f(x) = sin x in the interval [0, t], therefore

$\int_0^t \text{Sin } x\,dx = g(t) - g(o) = 1 - \cos t.$

We have, thus, reduced the problem of evaluating $\int_a^b f(x)\,dx$ to that of finding primitive of f on [a, b]. Once the primitive is known, the value of $\int_a^b f(x)\,dx$ is easily given by the Fundamental Theorem of Calculus.

You may note that any suitable primitive will serve the purpose because when the primitive is known, then the process described by the Fundamental Theorem is much simpler than other methods. However, it is just possible that the primitive may not exist while you may keep on trying to find it. It is, therefore, essential to formulate some conditions which can ensure the existence of a primitive. Thus now the next step is to find the conditions on the integral, (function to be integrated) which will ensure the existence of a primitive. One such condition is provided by the theorem.

According to theorem 2 if f is continuous in [a, b], then the function F given by

$F(x) = \int_a^x f(t)\,dt\ x \in [a, b]$ is differentiable in [a, b] and $F'(x) = f(x)\ \forall\ x \in [a, b]$

i.e. F is the primitive of f in [a, b]

The following observations are obvious from the theorems 1 and 2:

(i) If f is integrable on [a, b], then there is a function F which is associated with f through the process of integration and the domain of F is the same as the interval [a, b] over which f is integrated.

(ii) F is continuous. In other words, the process of integration generates continuous function.

(iii) If the function f is continuous on [a, b], then F is differentiable on [a, b]. Thus, the process of integration generates differentiable functions.

(iv) At any point of continuity of f, we will have f(c) = f(c) for c ε [a, b]. This means that if f is continuous on the whole of [a, b], then F will be a member of the family of primitives of f on [a, b].

In the case of continuous functions, this leads us to the notion

$$\int f(x)\,dx$$

for the family of primitives of f. Such an integral, as you know, is called an Indefinite integral of f. It does not simply denote one function, but it denotes a family of functions. Thus, a member of the indefinite integral of f will always be an antiderivative for f.

Theorem 3 gives US a condition on the function to be integrated which ensures the existence of a primitive. But how to obtain the primitives, once this condition is satisfied. In the next section, we look for the two most important techniques for finding the primitives. Before we do so, we need to study two important mean-values theorems of integrability.

Self Assessment

Fill in the blanks:

1. This function is not the derivative of any function F : [–1, 1] → R. Indeed if f is the derivative of a function F : [–1, 1] → R then f must have the ....

2. Let f : [a, b] → R be a continuous function. Let F : [a, b] → R be a function defined by ....

3. If f and F are functions defined on [a, b] such that F′(x) = f(x) for x ∈ [a, b] then F is called a 'primitive' or an '...............................' off on [a, b].

7.3 Summary

- The main thrust of this unit has been to establish the relationship between differentiation and integration with the help of the Fundamental Theorem of Calculus.
- We have discussed some important properties of the Riemann Integral. We have shown that the inequality between any two functions is preserved by their corresponding Riemann integrals; the modulus of the limit of a sum never exceeds the limit of the sum of their module and if we split the interval over which we are integrating a function into two parts, then the value of the integral over the whole will be the sum of the two integrals over the subintervals.
- Let f: [a, b] $\to$ R be a continuous function. Let F : [a, b] $\to$ R be a function defined by

$$F(x) = \int_a^x f(t)\, dt,\ x\ E[a, b].$$

Then F'(x) = f(x), a $\leq$ x $\leq$ b.

This is the first result which links the concepts of integral and derivative. It says that, if f is continuous on [a, b] then there is a function F on [a, b] such that $F'(x) = f(x), \forall\ x \in [a, b]$.

You have seen that if f: [a, b] $\to$ R is continuous, then there is a function F: [a, b] $\to$ R such that F' (x) = f(x) on [a, b]. Is such a function F unique? Clearly the answer is 'no'. For, if you add a constant to the function F, the derivative is not altered. In other words, if

$G(x) = c + \int_1^x f(t)\, dt$ for a $\leq$ x $\leq$ b then also G' (x) = f(x) on [a, b].

- It states that the integral of the derivative of a function is given by the function itself.
- The Fundamental Theorem of Calculus was given by an English mathematician Isaac Barrow [1630-1677], the teacher of great Isaac Newton.
- The following observations are obvious from the theorems 1 and 2.

  (i) If f is integrable on [a, b], then there is a function F which is associated with f through the process of integration and the domain of F is the same as the interval [a, b] over which f is integrated.

  (ii) F is continuous. In other words, the process of integration generates continuous function.

  (iii) If the function f is continuous on [a, b], then F is differentiable on [a, b]. Thus, the process of integration generates differentiable functions.

  (iv) At any point of continuity of f, we will have f(c) = f(c) for c e [a, b]. This means that if f is continuous on the whole of [a, b], then F will be a member of the family of primitives of f on [a, b].

7.4 Keywords

Primitive of a Function: If f and F are functions defined on [a, b] such that F'(x) = f(x) for x$\in$ [a, b] then F is called a 'primitive' or an 'antiderivative' off on [a, b].

Fundamental Theorem of Calculus: If f is integrable on [a, b] and F is a primitive of f on [a, b],

then $\int_a^{tb} f(x)\, dx = F(b) - F(a)$.

Notes

7.5 Review Questions

1. Find the primitive of the function f defined in [0, 2] by

$$f(x) = \begin{cases} x \text{ if } x \in [0,1] \\ I \text{ if } x \in [1,2] \end{cases}$$

2. Find $\int_0^2 f(x)\,dx$ where f is the function given in f(x) = $\begin{cases} x \text{ if } x \leftarrow [0,1] \\ 1 \text{ if } x \leftarrow [1,2] \end{cases}$

3. Evaluate $\int_1^b x^n\,dx$ where n is a positive integer.

Answers: Self Assessment

1. intermediate value property
2. $F(x) = \int_a^x f(t)\,dt,\ x\,E[a,b].$
3. antiderivative

7.6 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Sachin Kaushal, Lovely Professional University

Unit 8: Mean Value Theorem

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the first mean value theorem
- Explain the generalized first mean value theorem
- Describe the second mean value theorem

Introduction

In last unit, we discussed some mean-value theorems concerning the differentiability of a function. Quite analogous, we have two mean value theorems of integrability which we intend to discuss here. You are quite familiar with the two well-known techniques of integration namely the integration by parts and integration by substitution which you must have studied in your earlier classes.

8.1 First Mean Value Theorem

Let f : [a, b] → R be a continuous function. Then there exists c ∈ [a, b] such that

$$\int^{b} f(x)\,dx = (b-a)f(c).$$

Proof: We know that

$$m(b-a) \le \int_a^h f(x)\,dx \le M(b-a), \text{ thus}$$

$$m \le \frac{\int_a^b f(x)\,dx}{(b-a)} \le M, \text{ where}$$

m = glb {f(x) : x ∈ [a,b]}, and

M = lub {f(x) : x ∈ [a,b]}.

Notes

Since f is continuous in [a, b], it attains its bounds and it also attains every value between the bounds. Consequently, there is a point c ∈ [a, b] such that

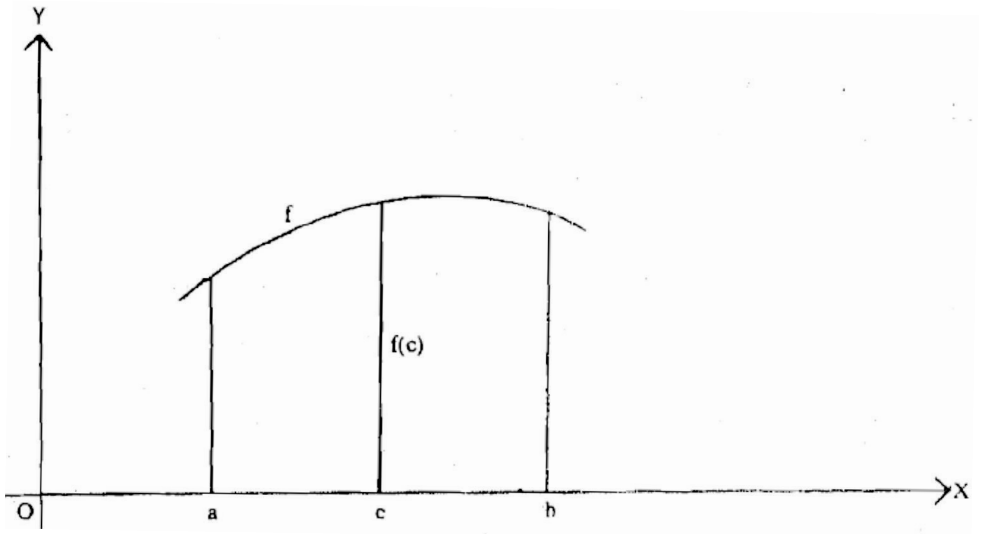
$$\int_a^b f(x)dx = f(c)(b-a),$$

which, equivalently, can be written as

$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx.$$

This theorem is usually referred to as the Mean Value theorem for integrals. The geometrical interpretation of the theorem is that for a non-negative continuous function f, the area between f, the lines x = a, x = b and the x-axis can be taken as the area of a rectangle having one side of length (b – a) and the other f(c) for some c ∈ [a, b] as shown in the Figure 25.1.

Figure 8.1



We now discuss the generalized form of the first mean value theorem.

8.2 The Generalised First Mean Value Theorem

Let f and g be any two functions integrable in [a, b]. Suppose g(x) keeps the same sign for all x ∈ [a, b]. Then there exists a number μ lying between the bounds of f such that

$$\int_a^b f(x)\,g(x)\,dx = \mu\int_a^b g(x)\,dx.$$

Proof: Let us assume that g(x) is positive over [a,b]. Since f and g are both integrable in [a, b], therefore both are bounded. Suppose that m and M are the g.l.b. and l.u.b. of f in [a, b]. Then

$$m \le f(x) \le M, \forall\, x \in [a,b].$$

Consequently,

$$mg(x) \le f(x)g(x) \le Mg(x), \forall\, x \in [a,b].$$

Therefore,

$$m\int_a^b g(x)\,dx \le \int_a^b f(x)\,g(x)\,dx \le M\int_a^b g(x)\,dx.$$

It then follows that there is a number $\mu \in [m, M]$ such that

$$\int_a^b f(x)\,g(x)\,dx = \mu\int_a^b g(x)\,dx.$$

Corollary: Let f, g be continuous functions on [a,b] and let $g(x) \ge 0$ on [a,b]. Then, there exists a $c \in [a,b]$ such that

$$\int_a^b f(x)\,g(x)\,dx = f(c)\int_a^b g(x)\,dx.$$

Proof: Since f is continuous on [a,b], so, there exists a point $c \in [a,b]$ such that

$\int_a^b f(x)\,g(x)\,dx = f(c)\int_a^b g(x)\,dx,$ where $\mu = f(c)$ is as in Theorem.

We use the first Fundamental Theorem of Calculus for integration by parts. We discuss it in the form of the following theorem.

Theorem 1: If f and g are differentiable functions Qn [a,b] such that the derivatives f'and g' are both integrable on [a,b], then

$$\int_a^b f(x)\,g'\,dx = [f(b)\,g(b) - f(a)\,g(b)] - \int_a^b f'(x)\,g(x)\,dx.$$

Proof: Since f and g are given to be differentiable on [a,b], therefore both f and g are continuous on [a,b]. Consequently both f and g are Riemann integrable on [a,b]. Hence both fg' as well as f' g are integrable.

$$fg' + f'g = (fg)'.$$

Therefore (fg)' is also integrable and consequently, we have

$$\int_a^b (fg)' = \int_a^b fg' + \int_a^b f'g.$$

By Fundamental Theorem of Calculus, we can write

$$\int_a^b (fg)' = |fg|_a^b = f(b)\,g(b) - f(a)\,g(a)$$

Hence, we have

$$\int_a^b fg' = f(b)\,g(b) - f(a)\,g(a) - \int_a^b f'g.$$

i.e.

$$\int_a^b f(x)\,g'(x)\,dx = [f(x)\,g(x)]_a^b - \int_a^b f'(x)\,g(x)\,dx.$$

This theorem is a formula for writing the integral of the product of two functions.

What we need to know is that the primitive of one of the two functions should be expressible in a simple form and that the derivative of the other should also be simple so that the product of these two is easily integrable. You may note here that the source of the theorem is the well-known product rule for differentiation.

The Fundamental Theorem of Calculus gives yet another useful technique of integration. This is known as method by Substitution also named as the change of variable method. In fact this is the reverse of the well-known chain Rule for differentiation. In other words, we compose the given function f with another function g so that the composite f o g admits an easy integral. We deduce this method in the form of the following theorem:

Theorem 2: Let f be a function defined and continuous on the range of a function g. If g' is continuously differentiable on $|c,d|$, then

$$\int_a^b f(x)\,dx = \int_c^d (f \circ g)\,(x)\,g'(x)\,dx,$$

where a = g(c) and b = g(d).

Proof: Let $F(x) = \int_a^b f(x)\,dt$ be a primitive of the function 1:

Note that the function F is defined on the range of g.

Since f is continuous, therefore, by Theorem 2, it follows that F is differentiable and F′(t) = f(t), for any t. Denote G(x) = (F o g) (x).

Then, clearly G is defined on [c,d] and it is differentiable there because both F and g are so. By the Chain Rule of differentiation, it follows that

$$G'(x) = (F \circ g)'\,(x)\,g'(x) = (f \circ g)\,(x)\,g'(x).$$

Also f o g is continuous since both f and g are continuous. Therefore, f o g is integrable.

Since g′ is integrable, therefore (f o g) g′ is also integrable. Hence

$$\begin{aligned}\int^d (f \circ g)(x)\,g'(x)\,dx &= \int_c^d G'(x)\,dx \\ &= G(d) - G(c) \qquad \text{(Why?)} \\ &= F(g\,(d)) - F(g\,(c)) \\ &= F(b) - F(a) \\ &= \int_a^b f(x)\,dx.\end{aligned}$$

you have seen that the proof of the theorem is based on the Chain Rule for differentiation. In fact, this theorem is sometimes treated as a Chain Rule for Integration except that it is used exactly the opposite way from the Chain Rule for differentiation. The Chain Rule for differentiation tells us how to differentiate a composite function while the Chain Rule for Integration or the change of variable method tells us how to simplify an integral by rewriting it as a composite function.

Thus, we are using the equalities in the opposite directions.

We conclude this section by a theorem (without Proof) known as the Second Mean Value Theorem for Integrals. Only the outlines of the proof are given.

Notes

8.3 Second Mean Value Theorem

Let f and g be any two functions integrable in $|a,b|$ and g be monotonic in $|a,b|$, then there exists $c \in |a,b|$ such that

$$\int_a^b f(x)\, g(x)\, dx \;= g(a) \int_a^c f(x)\, dx + g(b) \int_c^b f(x)\, dx$$

Proof: The proof is based on the following result known as Bonnet's Mean Value Theorem, given by a French mathematician O. Bonnet [1819–1892].

Let f and g be integrable functions in [a,b]. If ϕ is any monotonically decreasing function and positive in [a,b], then there exists a point c $\in$ [a,b] such that

$$\int_a^b f(x)\, \phi\,(x)\, dx \;= \phi\,(a) \int_a^c g(x)\, dx.$$

Let g be monotonically decreasing so that ϕ where $\phi(x)$ = g(x) – g(b), is non-negative and monotonically decreasing in [a,b]. Then there exists a number c $\in$ [a,b] such that

$$\int_a^b f(x)\,[g(x) - g(b)]\, dx \;= [g(a) - g(b)] \int_a^c f(x)\, dx$$

i.e.

$$\int_a^b f(x)\, g(x)\, dx \;= g(a) \int_a^c f(x)\, dx + g(b) \int_c^b f(x)\, dx.$$

Now let g be monotonically increasing so that –g is monotonically decreasing. Then there exists a number c $\in$ [a,b] such that

$$\int_a^b f(x)[-g(x)]\, dx \;= -g(a) \int^c f(x)\, dx - g(b) \int^b f(x)\, dx$$

i.e.

$$\int_a^b f(x)\, g(x)\, dx \;= g(a) \int_a^c f(x)\, dx + g(b) \int_c^h f(x)\, dx.$$

This completes the proof of the theorem.

There are several applications of the Second Mean Value Theorem. It is sometimes used to develop the trigonometric functions and their inverses which you may find in higher Mathematics. Here, we consider a few examples concerning the verification and application of the two Mean Value Theorems.

Example: Verify the two Mean Value Theorems for the functions f(x) = x, g(x) = e^x in the interval [–1, 1].

Solution: Verification of First Mean Value Theorem

Since f and g are continuous in [–1, 1], so they are integrable in [–1,1]. Also g(x) is positive in [–1, 1]. By first Mean Value Theorem, there is a number μ between the bounds of f such that

$$\int_{-1}^{1} f(x)\, g(x)\, dx \;= \mu \int_{-1}^{1} g(x)\, dx \text{ i.e., } \int_{-1}^{1} x\, e^x\, dx = \mu \int_{-1}^{1} e^x\, dx.$$

$$\int_{-1}^{1} x\, e^x\, dx = \left|x\, e^x\right|_{-1}^{1} - \int_{-1}^{1} e^x\, dx = \frac{2}{e} \text{ and } \int_{-1}^{1} e^x\, dx = e - \frac{1}{e}.$$

$$\therefore \quad \frac{2}{e} = \mu\left(e - \frac{1}{e}\right) \text{ i.e., } \mu = \frac{2}{e^2 - 1} = \frac{2}{(2.7)^2 - 1} = \frac{2}{6.29}$$

$\text{g.l.b.}\{f(x) | -1 \le x \le 1\} = -1$ and $\text{l.u.b.}\{f(x) | -1 \le x \le 1\} = 1$ and, so, $\mu \in [-1, 1]$. First Mean Value Theorem is verified.

Verification of Second Mean Value Theorem

As shown above, f and g are integrable in [–1, 1]. Also g is monotonically increasing in [-1, 1]. By second mean value theorem there is a points c ∈ [–1, 1] such that

$$\int_{-1}^{1} f(x)\, g(x)\, dx = g(-1)\int_{-1}^{c} f(x)\, dx + g(1)\int f(x)\, dx$$

$$\Rightarrow \quad \int_{-1}^{1} x\, e^x\, dx = \text{'I'}\ x\, dx + e\int_{c}^{1} x\, dx$$

$$\Rightarrow \quad \frac{2}{e} - \frac{1}{e}\left(\frac{c^2}{2} - \frac{1}{2}\right) + e\left(\frac{1}{2} - \frac{c^2}{2}\right).$$

Therefore $\quad c^2 = \frac{e^2 - 5}{e^2 - 1} = \frac{2.29}{6.29}$ i.e. $c = \pm\sqrt{\frac{2\ 29}{6\ 29}} \in [-1, 1]$

Thus second mean value theorem is verified.

Now we show the use of mean value theorems to prove some inequalities.

Example: By applying the first mean value theorem of Integral calculus, prove that

$$\pi/6 \le \int_{0}^{1/2} \cdot \frac{1}{\sqrt{\left[(1-x)^2(1-k^2 x^2)\right]}}\, dx \le \frac{\pi}{6}\frac{1}{\sqrt{(1 - \tfrac{1}{4}k^2)}}$$

Solution: In the first mean value theorem, take $f(x) = \frac{1}{\sqrt{(1 - k^2 x^2)}}$, $g(x) = \frac{1}{\sqrt{1 - x^2}}$, $x \in \left[0, \frac{1}{2}\right]$. Being continuous functions, f and g are integrable in $\left[0, \frac{1}{2}\right]$.

By the first mean value theorem, there is a number $\mu \in [m, M]$ such that

$$\int_{0}^{\frac{1}{2}} \frac{1}{\sqrt{\left[(1-x^2)(1-k^2 x^2)\right]}}\, dx = \int_{0}^{\frac{1}{2}} \frac{dx}{1 - x^2} = \mu\, \pi/\delta,$$

where in = g.l.b. $\left\{f(x) | 0 \le x \le \frac{1}{2}\right\}$ and $M = \text{l.u.b.}\left\{f(x) \middle| 0 \le x \le \frac{1}{2}\right\}$. Now m = 1 and $M = \frac{1}{\sqrt{1 - \frac{k^2}{4}}}$.

Therefore,

$$1 \le \mu \le \frac{1}{4} \text{ i.e. } \frac{\pi}{6} \le \frac{\mu\pi}{6} \le \frac{\pi}{6} - \frac{1}{\sqrt{1-\frac{k^2}{4}}},$$

$$\text{and; so, } \frac{\pi}{6} \le \int_0^{1/2} \frac{1}{\sqrt{\left[(1-x^2)(1-k^2x^2)\right]}}\,dx \le \frac{\pi}{6}\frac{1}{\sqrt{1-\frac{k^2}{4}}}.$$

Example: Prove that $\left|\int_p^q \frac{\sin x}{x}dx\right| \le \frac{2}{p}$, if q > p > 0.

Solution: Let $f(x) = \sin x, \phi(x) = \frac{1}{x}, x \in [p,q]$. Being continuous, these functions are integrable in [p, q]. By Bonnet form of second mean value theorem, there is a point $\xi \in$[p, q] such that

$$\int_p^q f(x)\,\phi(x)\,dx = \phi(p)\int_a^{\xi} f(x)\,dx$$

$$\text{i.e., } \int_p^q \frac{\sin x}{x}dx = \frac{1}{p}\int_p^{\xi} \sin x\,dx = \frac{1}{P}(\cos p - \cos\xi).$$

$$\text{Hence } \left|\int_p^q \frac{\sin x}{x}dx\right| \le \frac{1}{P}\left[|\cos p| + |\cos\xi|\right] \le \frac{2}{P}$$

Self Assessment

Fill in the blanks:

1. Let f : [a, b] $\to$ R be a continuous function. Then there exists c $\in$ [a, b] such that ....
2. Since f is continuous in [a, b], it attains its bounds and it also attains every value between the ....
3. The geometrical interpretation of the theorem is that for a .... function f, the area between f, the lines x = a, x = b and the x-axis can be taken as the area of a rectangle having one side of length (b-a) and the other f(c) for some c $\in$ [a, b].
4. If f and g are differentiable functions Qn [a,b] such that the derivatives f' and g' are both .... on [a, b], then

$$\int_a^b f(x)\,g'\,dx = [f(b)\,g(b) - f(a)\,g(b)] - \int_a^b f'(x)\,g(x)\,dx.$$

5. Let f and g be any two functions integrable in $|a,b|$ and g be .... then there exists $c \in |a,b|$ such that

$$\int_a^b f(x)\,g(x)\,dx \quad = g(a)\int_a^c f(x)\,dx + g(b)\int_c^b f(x)\,dx$$

Notes

8.4 Summary

- It has been proved that a continuous function has a primitive. Using the idea of a primitive, Fundamental Theorem or Calculus has been proved which shows that differentiation and integration are inverse process.
- Indefinite integral also called the integral function of an integrable function is defined and you have seen that this function is continuous. This function is differentiable whenever the integrable function is continuous. Finally in this section the First and Second Mean Value theorem have been discussed.
- The First Mean Value theorem states that if f is a continuous function in [a,b], then the value of the integral $\int_a^b f(x)\,dx$ is (b – a) times f(c) where $c \in [a, b]$. According to Generalised First Mean Value Theorem, if f and g are integrable in [a, b] and g(x) keeps the same sign, then the value $\int_a^h f(x)\,g(x).dx$ is $\int_a^b f(x)\,g(x)\,dx = \mu\int_a^b g(x)\,dx$ where μ lies between the bounds of f. But in the second mean value theorem, if out of the integrable functions f and g, g is monotonic in [a, b], then the value $\int_a^h f(x)\,g(x)\,dx$ is $g(a)\int_a^c f(x)\,dx + g(b)\int_c^b f(x)\,dx$ where c is point of [a, b].

8.6 Keywords

First Mean Value Theorem: Let f : [a, b] → R be a continuous function. Then there exists c ∈ [a, b] such that

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

The Generalised First Mean Value Theorem: Let f and g be any two functions integrable in [a, b]. Suppose g(x) keeps the same sign for all x ∈ [a, b]. Then there exists a number μ lying between the bounds off such that

$$\int_a^b f(x)\,g(x)\,dx = \mu\int_a^b g(x)\,dx.$$

Second Mean Value Theorem: Let f and g be any two functions integrable in $|a,b|$ and g be monotonic in $|a,b|$, then there exists $c \in |a,b|$ such that

$$\int_a^b f(x)\,g(x)\,dx = g(a)\int_a^c f(x)\,dx + g(b)\int_c^b f(x)\,dx$$

8.6 Review Questions

1. Show that the second mean value theorem does not hold good in the interval [–1, 1] for $f(x) = g(x) = x^2$.
2. What do you say about the validity of the first mean value theorem.

   {1, 2] for $f(x) = g(x) = x^3$.
3. Show that $\left|\int_a^b \sin x^2\,dx\right| \le \frac{1}{a}$, if $b > a > 0$.

Notes

Answers: Self Assessment

1. $\int^{b} f(x)\, dx = (b - a)f(c)$
2. bounds
3. non-negative continuous
4. integrable
5. monotonic in $|a, b|$

8.7 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 9: Lebesgue Measure

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the definition of outer measure of sets
- Define outer measure of an interval
- Explain some important properties of outer measure
- Define measurable sets
- Describe measure of countable union of measurable sets
- Measure countable intersection of measurable sets

Introduction

In last unit you have studied about mean value theorems of Riemann Stieltjes integral. In this unit we are going to study about Lebesgue outer measure of a set, measurable sets and Lebesgue measure, their important properties.

We know that the length of an interval is defined to be the difference between two end points. In this unit, we would like to extend the idea of "length" to arbitrary (or at least, as many as possible) subsets of $\mathbb{R}$. To begin with, let's recall two important results in topology.

9.1 Lindelof's Theorem

Proposition: Every open subset V of $\mathbb{R}$ is a countable union of disjoint open intervals.

Proof: For each $x \in V$, there is an open interval I_x with rational endpoints such that $x \in I_x \subseteq V$. Then the collection $\{I_x\}_{x \in V}$ is evidently countable and

$$V = \bigcup_{x \in V} I_x.$$

Next, we prove it is always possible to have a disjoint collection. Since $\{I_x\}_{x \in V}$ is a countable collection, we can enumerate the open intervals as $I_1 = (a_1, b_1)$, $I_2 = (a_2, b_2)$,.... For each $n \in \mathbb{N}$, define

$$\alpha_n = \inf\{ x \in \mathbb{R} : x \le a_n \text{ and } (x, b_n) \subseteq V\}$$

and

$$\beta_n = \sup\{ x \in \mathbb{R} : x \ge b_n \text{ and } (a_n, x) \subseteq V\}.$$

Then $\{ (\alpha_n, \beta\nu) \}_{n \in \mathbb{N}}$ is a disjoint collection of open intervals with union V.

Theorem 1 (Lindelof's Theorem): Let C be a collection of open subsets of $\mathbb{R}$. Then there is a countable sub-collection $\{O_i\}_{i \in \mathbb{N}}$ of C such that

$$\bigcup_{O \in C} O = \bigcup_{i=1}^{\infty} O_i$$

Proof: Let $U = \bigcup_{o \in C} O$. For any $x \in U$ there is $O \in C$ with $x \in O$. Take an open interval I_x with rational endpoints such that $x \in I_x \subseteq O$. Then $U = \bigcup_{x \in U} I_x$ is a countable union of open intervals. Replace I_x by the set $O \in C$ which contains it, the result follows.

9.2 Lebesgue Outer Measure

As in the Archimedean idea of finding area of a circle (approximated polygons), we define the Lebesgue outer measure $m^*: \mathcal{P}(\mathbb{R}) \to [0, \infty]$ by

$$m^*(A) = \inf\left\{ \sum_{k=1}^{\infty} \ell(I_k) : A \subseteq \bigcup_{k=1}^{\infty} I_k \text{ and each } I_k \text{ being open interval in } \mathbb{R} \right\}.$$

> *Notes* By Lindelof's Theorem, the countability of the covering is not important here.

Here are some basic properties of Lebesgue outer measure, all of them can be proved easily by the definition of m*.

(i) $m^*(A) = 0$ if A is at most countable.

(ii) m* is monotonic, i.e. $m^*(A) \le m^*(B)$ whenever $A \subseteq B$.

(iii) $m^*(A) = \inf \{m^*(O): A \subseteq O \text{ and } O \text{ is open}\}$. (*Hint:* it suffices to prove $m^*(A) \ge$ R.H.S., which is equivalent to $m^*(A) + \varepsilon >$ R.H.S. for any $\varepsilon > 0$.)

(iv) $m^*(A + x) = m^*(A)$ for all $x \in \mathbb{R}$. (Translation-invariant)

(v) $m^*(\bigcup_{k \in \mathbb{N}} A_k) \le \sum_{k=1}^{\infty} m^*(A_K)$. (Countable subadditivity)

(vi) If $m^*(A) = 0$, then $m^*(A \cup B) = m^*(B)$ and $m^*(B \setminus A) = m^*(B)$ for all $B \subseteq R$.

(vii) If $m^*(A \Delta B) = 0$, then $m^*(A) = m^*(B)$.

> *Notes* In (v), even if A_k's are disjoint, the equality may not hold.

Theorem 2: For any interval $I \subseteq R$, $m^*(I) = \ell(I)$.

Proof: We first assume $I = [a, b]$ is a closed and bounded interval. Consider the countable open interval cover $\{(a - \varepsilon, b + \varepsilon)\}$ of I, we have $m^*(I) \le b - a + 2\varepsilon$. Since $\varepsilon > 0$ is arbitrary, $m^*(I) \le b - a$.

To get the opposite result, we need to show for any $\varepsilon > 0$, $m^*(I) + \varepsilon \geq b - a$. Note that there is a countable open interval cover $\{I_k\}_{k\in\mathbb{N}}$ of I satisfying

$$m^*(I) + \varepsilon > \sum \ell(I_k).$$

By Heine-Borel Theorem, there is a finite subcover $\{I_{n_k}\}$ of $\{I_k\}$. Then

$$\sum \ell(I_{n_k}) > b - a \qquad \text{(why?)}$$

and it follows that

$$m^*(I) + \varepsilon > \sum \ell(I_k) \geq \sum \ell(I_{n_k}) > b - a.$$

Letting $\varepsilon \to 0$, $m^*(I) \geq b - a$. Hence, $m^*(I) = b - a$.

Next, we consider the case where I = (a, b), [a, b), or (a, b] which is bounded but not closed. Clearly, $m^*(I) \leq m^*(\overline{I}) = b - a$. On the other hand, if $\varepsilon > 0$ is sufficiently small then there is a closed and bounded interval $I' = [a + \varepsilon, b - \varepsilon] \subseteq I$. By monotonicity, $m^*(I) \geq m^*(I') = b - a - 2\varepsilon$. Letting $\varepsilon \to 0$ gives $m^*(I) \geq b - a$. Hence, $m^*(I) = b - a$.

Finally, if I is unbounded then the result is trivial since in that case I contains interval of arbitrarily large length.

9.3 Non-measurability

Theorem 3: Let $\mathfrak{M} \subseteq \mathcal{P}(\mathbb{R})$ be a translation-invariant σ-algebra containing all er that intervals, and $m : \mathfrak{M} \to [0, \infty]$ be a translation-invariant, countably additive measure such

$$m(I) = \ell(I) \qquad \text{for all interval I.}$$

Then there exists a set $S \notin \mathfrak{M}$.

Proof: Define an equivalent relation $x \sim y$ if and only if $x - y$ is rational. Then $\mathbb{R}$ is partitioned into disjoint cosets $[x] = \{y \in \mathbb{R} : x \sim y\}$.

By Axiom of Choice and Archimedean property of $\mathbb{R}$, there exists $S \subseteq [0,1]$ such that the intersection of S with each coset contains exactly one point.

Enumerate $\mathbb{Q} \cap [-1, 1]$ into $r_1, r_2, \ldots$. Then the sets $S + r_i$ are disjoint and

$$[0, 1] \subseteq \bigcup_{i\in\mathbb{N}} (S + r_i) \subseteq [-1, 2].$$

If S $S \in \mathfrak{M}$, then by monotonicity and countable additivity of m we have

$$1 \leq \sum_{i\in\mathbb{N}} m(S + r_i) \leq 3,$$

which is impossible since $m(S + r_i) = m(S)$ for all $i \in \mathbb{N}$.

9.4 Measurable Sets and Lebesgue Measure

As it is mentioned before, the outer measure does not have countable additivity. One may try to restrict the outer measure m* to a σ-algebra $\mathfrak{M} \subsetneq \mathcal{P}(\mathbb{R})$ such that the new measure has all the properties we wanted.

Definition (Measurability): A set $E \subseteq \mathbb{R}$ is said to be measurable if, for all $A \subseteq \mathbb{R}$, one has

(1) $$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c).$$

Since m* is known to be subadditive, (1) is equivalent to

$$m^*(A) > m^*(A \cap E) + m^*(A \cap E^c).$$

The family of all measurable sets is denoted by $\mathfrak{M}$. We will see later $\mathfrak{M}$ is a σ-algebra and translation-invariant containing all intervals. The set function m: $\mathfrak{M} \to [0, \infty]$ defined by

$$m(E) = m^*(E) \qquad \text{for all } E \in \mathfrak{M}$$

is called Lebesgue measure.

Observe that

- $E \in \mathfrak{M} \Leftrightarrow E^c \in \mathfrak{M}$.
- $\phi \in \mathfrak{M}$ and $\mathbb{R} \in \mathfrak{M}$ because $m^*(A) = m^*(A \cap \phi) + m^*(A \cap \mathbb{R})$ for all $A \subseteq \mathbb{R}$.
- $m^*(E) = 0 \Rightarrow E \in \mathfrak{M}$ because $m^*(A \cap E) + m^*(A \cap E^c) = m^*(A \cap E^c) \le m^*(A)$ for all $A \subseteq \mathbb{R}$.

Proposition: If $E_1, E_2 \in \mathfrak{M}$ then $E_1 \cup E_2 \in \mathfrak{M}$. (Therefore, $\mathfrak{M}$ is an algebra.)

Proof: For all $A \subseteq R$ one has

$$\begin{aligned} m^*(A) &= m^*(A \cap E_1) + m^*(A \cap E_1^c) \qquad (\because E_1 \in \mathfrak{M}) \\ &= m^*(A \cap E_1) + m^*(A \cap E_1^c \cap E_2) + m^*(A \cap E_1^c \cap E_2^c) \qquad \boxed{(\because E_2 \in \mathfrak{M})} \\ &= m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap (E_1 \cup E_2)^c) \end{aligned}$$

because m* is subadditive and

$$A \cap (E_1 \cup E_2) = (A \cap E_1) \cup (A \cap E_1^c \cap E_2).$$

> *Notes* Above proposition can be easily extended to a finite union of measurable sets, in fact it can be extended to a countable union. In order to do so, we need the following result.

Lemma 1: Let $E_1, E_2, \ldots, E_n$ be disjoint measurable sets. Then for all $A \subseteq R$, we have

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

Proof: Since $E_n \in \mathfrak{M}$, we have

$$\begin{aligned} m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) &= m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n\right) + m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n^c\right) \\ &= m^*(A \cap E_n) + m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) \end{aligned}$$

Repeat the process again and again, until we get

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

> *Notes* If $\{E_i\}_{i \in \mathbb{N}}$ is a sequence of disjoint measurable sets, then
>
> $$m^*\left(A \cap \left[\bigcup_{i=1}^{\infty} E_i\right]\right) = \sum_{i=1}^{\infty} m^*(A \cap E_i).$$

This is because for all $n \in \mathbb{N}$ one has

$$m^*\left(A \cap \left[\bigcup_{i=1}^{\infty} E_i\right]\right) \geq m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right)$$

$$= \sum_{i=1}^{n} m^*(A \cap E_i).$$

Letting $n \to \infty$ lead to

$$m^*\left(A \cap \left[\bigcup_{i=1}^{\infty} E_i\right]\right) \geq \sum_{i=1}^{\infty} m^*(A \cap E_i).$$

The opposite inequality follows from countable subadditivity.

Theorem 4: Let $\{E_i\}_{n \in \mathbb{N}}$ be a sequence of measurable sets, then $E = \bigcup_{i=1}^{\infty} E_i$ is also measurable. Moreover, if $E_1, E_2, \ldots$ are disjoint then $m(E) = \sum_{i=1}^{\infty} m(E_i)$.

This is called the countable additivity which can be proved by putting $A = \mathbb{R}$

Proof: We first assume $E_1, E_2, \ldots$ are disjoint. Then for all $A \subseteq R$, $n \in \mathbb{N}$ we have

$$m^*(A) = m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) + m^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)^c\right)$$

$$\geq \sum_{i=1}^{n} m^*(A \cap E_i) + m^*(A \cap E^c).$$

Letting $n \to \infty$,

$$m^*(A) \geq \sum_{i=1}^{\infty} m^*(A \cap E_i) + m^*(A \cap E^c)$$

$$= m^*(A \cap E) + m^*(A \cap E^c).$$

This proved E is measurable.

Now, if $E_1, E_2, \ldots$ are not disjoint, we let

$$F_1 = E_1, \qquad F_2 = E_2 \backslash F_1, \qquad F_3 = E_3 \backslash (F_1 \cup F_2),$$

and in general $F_k = E_k \backslash \bigcup_{i=1}^{k-1} F_i$ for $k > 2$. Then $F_1, F_2, \ldots$ are disjoint and $\bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} E_i$. Since $\mathfrak{M}$ is an algebra, $F_1, F_2, \ldots$ are all measurable. So $E = \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} F_i$ is measurable.

> *Notes* $\mathfrak{M}$ is proved to be a σ-algebra. The next result shows that all Borel sets are measurable. Recall that the family of Borel sets in $\mathbb{R}$ is, by definition, the smallest σ-algebra containing all open subsets of $\mathbb{R}$.

Theorem 5: $\mathfrak{M}$ contains all Borel subsets of $\mathbb{R}$.

Proof: It suffices to show that $(a, \infty) \in \mathfrak{M}$ for all $a \in \mathbb{R}$ (why?). Let $A \in \mathbb{R}$. We need to show that

$$m^*(A) \geq m^*(A \cap (-\infty, a]) + m^*(A \cap (a, \infty)).$$

Without loss of generality, we may assume $m^*(A) < \infty$. For convenience, let $A_1 = A \cap (-\infty, a]$ and $A_2 = A \cap (a, \infty)$. Then we need to show

$$m^*(A) + \varepsilon \geq m^*(A_1) + m^*(A_2) \qquad \text{for all } \varepsilon > 0.$$

By the definition of m*(A), there is a countable open interval cover $\{I_n\}_{n\in\mathbb{N}}$ of A with

$$m^*(A) + \varepsilon > \sum_{n=1}^{\infty} \ell(I_n).$$

Let $I_n' = I_n \cap [-\infty, a]$ and $I_n'' = I_n \cap (a, \infty)$, then $\{I_n'\}, \{I_n''\}$ are, respectively, interval covers of A_1 and A_2 (note that they may not be open interval covers). Then

$$\begin{aligned}\sum_{n=1}^{\infty} \ell(I_n) &= \sum_{n=1}^{\infty} \ell(I_n') + \sum_{n=1}^{\infty} \ell(I_n'') \\ &= \sum_{n=1}^{\infty} m^*(I_n') + \sum_{n=1}^{\infty} m^*(I_n'') \qquad (\because\ m^* = \ell \text{ for intervals}) \\ &\geq m^*\left(\bigcup_{n=1}^{\infty} I_n'\right) + m^*\left(\bigcup_{n=1}^{\infty} I_n''\right) \qquad (\because\ \text{countable subadditivity}) \\ &\geq m^*(A_1) + m^*(A_2) \qquad (\because\ \text{monotonicity})\end{aligned}$$

So, $m^*(A) + \varepsilon > \sum_{n=1}^{\infty} \ell(I_n) \geq m^*(A_1) + m^*(A_2)$ for all $\varepsilon > 0$. Letting $\varepsilon \to 0$, $m^*(A) \geq m^*(A_1) + m^*(A_2)$. This proved that $(a, \infty) \in \mathfrak{M}$.

> *Notes* Since $\mathfrak{M}$ is a σ-algebra, $(-\infty, a] \in \mathfrak{M}$ and $(-\infty, a) = \bigcup_{n=1}^{\infty} (-\infty, a - 1/n] \in \mathfrak{M}$. It follows that $(a, b) \in \mathfrak{M}$ since $(a, b) = (-\infty, b) \cap (a, \infty)$. As $\mathfrak{M}$ is a σ-algebra containing all open intervals, it must contain all open sets (recall that every open set is countable union of open intervals by Proposition). Therefore, $\mathfrak{M}$ contains all Borel sets.

Proposition: $\mathfrak{M}$ is translation invariant: for all $x \in \mathbb{R}$, $E \in \mathfrak{M}$ implies $E + x \in \mathfrak{M}$.

Proof: For all $A \in R$, we have

$$\begin{aligned}m^*(A) &= m^*(A - x) \\ &= m^*((A - x) \cap E) + m^*((A - x) \cap E^c) \\ &= m^*(((A - x) \cap E) + x) + m^*(((A - x) \cap E^c) + x) \\ &= m^*(A \cap (E + x)) + m^*(A \cap (E + x)^c)\end{aligned}$$

> *Notes* Let $E \in \mathbb{R}$ be given. Then the following statements are equivalent.
>
> 1. E is measurable.
> 2. For any $\varepsilon > 0$, there is an open set $O \supseteq E$ such that $m^*(O \setminus E) < \varepsilon$.
> 3. For any $\varepsilon > 0$, there is a closed set $F \subset E$ such that $m^*(E \setminus F) < \varepsilon$.
> 4. There is a $G \in G_\delta$ such that $E \subseteq G$ and $m^*(G \setminus E) = 0$.
> 5. There is a $F \in F_\sigma$ such that $E \supseteq F$ and $m^*(E \setminus F) = 0$.
>
> Assume $m^*(E) < \infty$, the above statements are equivalent to
>
> 6. For any $\varepsilon > 0$, there is a finite union U of open intervals such that $m^*(U \Delta E) < \varepsilon$.

Theorem 6: Littlewood's 1st Principle

Every measurable set of finite measure is nearly a finite union of disjoint open intervals, in the sense

- If E is measurable and $m(E) < \infty$, then for any $\varepsilon > 0$ there is a finite union U of open intervals such that $m^*(U \Delta E) < \varepsilon$. (Clearly, the intervals can be chosen to be disjoint.)
- If for any $\varepsilon > 0$ there is a finite union U of open intervals such that $m^*(U \Delta E) < \varepsilon$, then E is measurable. (The finiteness assumption $m^*(E) < \infty$ is not essential.)

Proof: If we can prove (1), (2), and (4) are equivalent, then it is easy to see that (2) and (3) are equivalent, because one implies another by replacing E with E^c. Similarly, (4) and (5) are equivalent.

To show (1) $\Rightarrow$ (2)

We first consider a simple case $m(E) < \infty$. For any $\varepsilon > 0$, there is a countable open interval cover $\{I_n\}$ of E such that $\sum_{n=1}^{\infty} \ell(I_n) < m(E) + \varepsilon$. Take $O = \bigcup_{n=1}^{\infty} I_n$, we see that O is open and $O \supseteq E$. Also, we have

$$m(O\setminus E) = m(O) - m(E) \leq \sum_{n=1}^{\infty} m(I_n) - m(E) < \varepsilon.$$

Here we use the assumption $m(E) < \infty$ and the countable subadditivity of m.

For the case $m(E) = \infty$, we write $E = \bigcup_{n=1}^{\infty} E_n$, where $E_n = E \cap [-n, n]$. This is a countable union of measurable sets of finite measure. By the above result there is an open set O_n such that $O_n \supseteq E_n$ and $m^*(O_n\setminus E_n) < \frac{\varepsilon}{2^n}$. Take $O = \bigcup_{n=1}^{\infty} O_n$, then O is open and $O \supseteq E$. It remains to show $m(O\setminus E) < \varepsilon$.

Note that $O\setminus E \subseteq \bigcup_{n=1}^{\infty} O_n\setminus E_n$, by countable subadditivity of m we have

$$m(O\setminus E) \leq \sum_{n=1}^{\infty} m(O_n\setminus E_n) < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

Hence, we have proved that (1) $\Rightarrow$ (2).

To show (2) $\Rightarrow$ (4)

For any $n \in \mathbb{N}$, let O_n be an open set such that $O_n \supseteq E$ and $m^*(O_n\setminus E) < 1/n$. Take $G = \bigcap_{n=1}^{\infty} O_n \in G_\delta$, then

$$m^*(G\setminus E) \leq m^*(O_n\setminus E) < \frac{1}{n}.$$

Letting $n \to \infty$, the result follows.

To show (4) $\Rightarrow$ (1)

The existence of G guarantees $E = G\setminus(G\setminus E)$ is measurable since both G and $G\setminus E$ are measurable (G is Borel set and $G\setminus E$ is of measure zero).

Hence, (1), (2), (3), (4), (5) are equivalent.

To show (2) $\Rightarrow$ (6) (with finiteness assumption $m^*(E) < \infty$)

Let $\varepsilon > 0$ be given. Let O be an open set such that $O \supseteq E$ and $m(O\setminus E) < \varepsilon/2$. Write $O = \bigcup_{n=1}^{\infty} I_n$ to be a countable union of disjoint open intervals. By the countable additivity of m, $m(O) = \sum_{n=1}^{\infty} \ell(I_n)$. Let k be a positive integer such that $\sum_{n=1}^{k} \ell(I_n) > m(O) - \varepsilon/2$. (The finiteness assumption has been used here to guarantee that $m(O) < \infty$.)

Take U = $\bigcap_{n=1}^{k} I_n$. Note that m(O \ U) = m(O) – m(U) < ε/2, so

$$m(U \Delta E) = m(U\backslash E) + m(E\backslash U)$$
$$\leq m(O\backslash E) + m(O\backslash U)$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

The finiteness assumption is essential here. The above result is false if we allow E to have infinite measure. A counter example is E = $\bigcap_{n=1}^{\infty}(2n, 2n + 1)$.

To show (6) ⇒ (2) (without finiteness assumption m*(E) < ∞)

Let ε > 0 be given and U be a finite union of open intervals. Then m*(E\U) < ε, we take an open set O′ ⊇ E\U such that m*(O′) < ε (how to do this?). Then O = U ∪ O′ is an open set containing E with m*(O\E) ≤ m*(U\E) + m*(O′) < 2 ε.

Task Let A ∈ ℝ, prove that there is a measurable set B ⊇ A with m*(A) = m*(B).

9.5 Step Functions and Simple Functions

Definition: A function ψ: [a, b] → ℝ is called step function if

$$\psi(x) = c_i \qquad (x_{i-1} < x < x_i)$$

for some partition $\{x_0, x_1,...,x_n\}$ of [a, b] and some constants $c_1, c_2,..., c_n$.

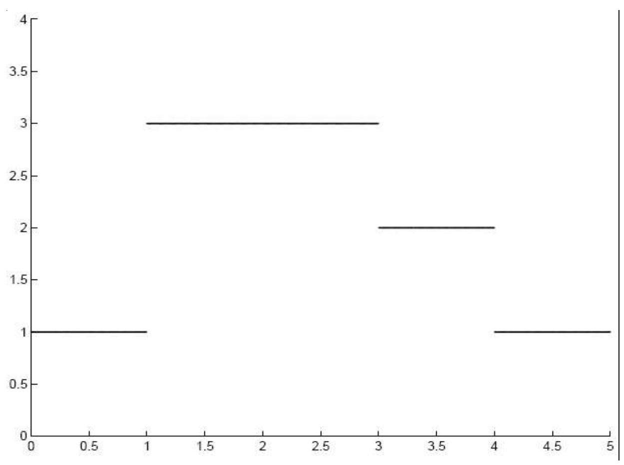
Lemma: Let ψ_1, ψ_2 be step functions on [a, b]. Then $\psi_1 \pm \psi_2$, $\alpha\psi_1+\beta\psi_2$, $\psi_1 \psi_2$, $\psi_1 \wedge \psi_2$, and $\psi_1 \vee \psi_2$ are all step functions, where α, β ∈ ℝ. Also, if $\psi_2 \neq 0$ on [a,b], then ψ_1/ψ_2 is also step function.

Note (f ∧ g)(x) = min{f(x), g(x)} and (f ∨ g)(x) = max{f(x), g(x)}.

Lemma: Let ψ be a step function on [a,b] and let ε > 0. Then there is a continuous function g on [a, b] such that ψ = g on [a, b] except on a set of measure less than ε, i.e.

$$m(\{ x \in [a, b] : \psi(x) \neq g(x)\}) < \varepsilon.$$

Figure 9.1: An Example of Step Function



Proof: Easy! One can find a piecewise linear function g with the stated property.

Definition: Let $E \in \mathfrak{M}$. A function $f: E \to \mathbb{R}$ is called a simple function if there exists $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and $E_1, E_2, \ldots, E_n \in \mathfrak{M}$ such that

(2) $$f = \sum_{i=1}^{k} a_i \chi_{E_i}$$

Note Step function is simple, $\chi_{\mathbb{Q}}$ is simple but not step function.

Proposition: Let $f: [a, b] \to \mathbb{R}$ be a simple function. For any $\varepsilon > 0$, there is a step function $\psi: [a, b] \to \mathbb{R}$ such that $f = \psi$ except on a set of measure less than ε.

Proof: Let f be given by (2), we may assume $E_1, E_2, \ldots, E_n \subseteq E$. By Littlewood's 1st Principle, there is a finite union of disjoint open intervals U_i such that $m(U_i \Delta E_i) < \varepsilon/n$. Then

$$f = \sum_{i=1}^{k} a_i \chi_{U_i} \quad \text{except on } A = \bigcup_{i=1}^{n} (U_i \Delta E_i),$$

where $m(A) < \sum_{i=1}^{n} \varepsilon/n = \varepsilon$.

Notes One can find a continuous function with the same property. Moreover, if f satisfies $m \le f \le M$ on [a, b] then ψ can be chosen such that $m \le \psi \le M$ (reason: replace ψ by $(m \vee \psi) \wedge M$ if necessary).

9.6 Measurable Functions

Definition: A function $f: E \to [-\infty, \infty]$ is said to be measurable (or measurable on E) if $E \in \mathfrak{M}$ and

$$f^{-1}((a, \infty]) \in \mathfrak{M}$$

for all $a \in \mathbb{R}$.

In fact, there is a more general definition for measurability which we will not use here. The definition goes as follows.

Definition: Let X be a measurable space and Y be a topological space. A function $f: X \to Y$ is called measurable if $f^{-1}(V)$ is a measurable set in X for every open set V inY.

Notes Simple functions, step functions, continuous functions and monotonic functions are measurable.

Proposition: Let $E \in \mathfrak{M}$ and $f : E \to [-\infty, \infty]$. Then the following four statements are equivalent:

- $f^{-1}((a, \infty]) \in \mathfrak{M}$ for all $a \in \mathbb{R}$.
- $f^{-1}([a, \infty]) \in \mathfrak{M}$ for all $a \in \mathbb{R}$.
- $f^{-1}([-\infty, a)) \in \mathfrak{M}$ for all $a \in \mathbb{R}$.
- $f^{-1}([-\infty, a]) \in \mathfrak{M}$ for all $a \in \mathbb{R}$.

Notes The above statements imply $f^{-1}(a) \in \mathfrak{M}$ for all $a \in [-\infty, \infty]$. The converse is not true.

Proof: The first one is clearly equivalent to the fourth one since $f^{-1}([a, \infty]) = E \setminus f^{-1}([-\infty, a])$. Similarly, the second and the third statements are equivalent. It remains to show the first two statements are equivalent, but this follows immediately from

$$f^{-1}([a, \infty]) = \bigcup_{n=1}^{\infty} f^{-1}\left((a - \frac{1}{n}, \infty]\right) \qquad \text{and} \qquad f^{-1}([a, \infty]) = \bigcup_{n=1}^{\infty} f^{-1}\left([a + \frac{1}{n}, \infty]\right).$$

Proposition: Let $E \in \mathfrak{M}$, $f: E \to [-\infty, \infty]$ and $g: E \to [-\infty, \infty]$. If $f = g$ almost everywhere on E then the measurability of f and g are the same.

Proof: Simply note that

$$m^*(\{x \in E : f(x) > a\} \,\Delta\, \{x \in E : g(x) > a\}) \leq m^*(\{x \in E : f(x) \neq g(x)\}) = 0.$$

This implies the measurability of the sets $\{x \in E: f(x) > a\}$ and $\{x \in E : g(x) > a\}$ are the same.

Proposition: Let f, g be measurable extend real-valued functions on $E \in \mathfrak{M}$. Then the following functions are all measurable on E:

$$f + c,\ cf,\ f \pm g,\ fg$$

where $c \in \mathbb{R}$.

Notes One may worry that cf, $f \pm g$, fg may not be defined at some points (for example, if $f = \infty$ and $g = -\infty$ then $f + g$ is meaningless). There are two ways to deal with this problem.

1. Adopt the convention $0 \cdot \infty = 0$.
2. Assume f, g are finite almost everywhere or cf, $f \pm g$, fg are meaningful almost everywhere.

Proof: We only prove $f + g$ and fg are measurable, since the others are easy or similar.

To prove $f + g$ is measurable, one should consider the set

$$\begin{aligned} E_a &= \{x \in E : f(x) + g(x) > a\} \\ &= \{x \in E : f(x) > a - g(x)\} \\ &= \bigcup_{r \in \mathbb{Q}} \{x \in E : f(x) > r > a - g(x)\} \\ &= \bigcup_{r \in \mathbb{Q}} \{x \in E : f(x) > r\} \cap \{x \in E : r > a - g(x)\} \end{aligned}$$

If $f(x) = \infty$ or $g(x) = \infty$ then $x \in E_a$ by convention. Now $E_a \in \mathfrak{M}$ because E_a is countable union of measurable sets.

Next, we prove f^2 is measurable. For $a \geq 0$,

$$\{x \in E: f^2(x) > a\} = \{x \in E : f(x) > \sqrt{a}\} \cup \{x \in E : f(x) < -\sqrt{a}\}$$

is measurable. For $a < 0$, $\{x \in E : f^2(x) > a\} = E$ is also measurable. Therefore, f^2 is measurable and it is valid even if f takes values $\pm \infty$.

Notes

So, if f and g are assumed to be finite, then

$$fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$$

is measurable on E.

Task Find two measurable functions f, g from $\mathbb{R}$ to $\mathbb{R}$ such that f o g is not measurable.

Proposition: Let $\{f_n\}_{n\in\mathbb{N}}$ be measurable extended real-valued functions on a measurable set E. Then

$$f_1 \vee f_2 \ldots \vee f_n, \qquad \sup_{n\in\mathbb{N}} f_n, \qquad \overline{\lim_{n\to\infty}} f_n$$

are all measurable on E. Similar results hold if $\vee$, sup and $\overline{\lim}$ are replaced by $\wedge$, inf, and $\underline{\lim}$.

Proof: Simply note that

$$(f_1 \vee f_2 \cdots \vee f_n)^{-1}((a,\infty)) = \bigcup_{k=1}^{\infty} f_k^{-1}((a,\infty))$$

$$\left(\sup_{n\in\mathbb{N}} f_n\right)^{-1}((a,\infty)) = \bigcup_{k=1}^{\infty} f_k^{-1}((a,\infty))$$

$$\overline{\lim_{n\to\infty}} f_n = \inf_{N\in\mathbb{N}}\left(\sup_{k\geq N} f_k\right)$$

Theorem 7: Let $E \in \mathfrak{M}$ with $m(E) < \infty$, $f: E \to [-\infty,\infty]$ be measurable and finite almost everywhere. For any $\varepsilon > 0$, there is a simple function ϕ such that

$|f - \phi| < \varepsilon$ on E except on a set of measure less than ε.

Notes If E = [a, b] is closed and bounded interval, we can find a step function g and a continuous function h play the role of ϕ. This is because simple function can be approximated by step function and step function can be approximated by continuous function.

If f satisfies an additional condition $m \leq f \leq M$, then ϕ, g, and h can be chosen to be bounded below by m and above by M.

The condition $m(E) < \infty$ in Littlewood's 2nd Principle is essential. You can see if this condition is dropped then taking f(x) = x will give a counter example.

To prove Littlewood's 2nd Principle, we introduce a lemma.

Lemma: Let $\{F_n\}_{n\in\mathbb{N}}$ be a sequence of measurable subsets of $\mathbb{R}$ (or any measure space[3]) such that

$$F_1 \supseteq F_2 \supseteq \cdots.$$

Denote $F_\infty = \bigcap_{n\in\mathbb{N}} F_n$. If $m(F_1) < \infty$ then

$$m(F_\infty) = \lim_{n\to\infty} m(F_n).$$

Proof: Write $F_1 = F_\infty \cup (F_1 \backslash F_2) \cup (F_2 \backslash F_3) \cup \cdots$ as disjoint union and use the countable additivity of m.

> *Notes* Above Lemma 4 is false if the condition $m(F_1) < \infty$ is missing.

Now, we are ready to prove Littlewood's 2nd Principle.

Proof: Let $\varepsilon > 0$ be given. Our proof is divided into two steps.

Step I: Assume $m \le f \le M$ for some $m, M \in \mathbb{R}$.

We divide [m, M] into n subintervals such that the length of each subinterval is less than ε. Symbolically, we take the partition points as follows:

$$m = y_0 < y_1 < \cdots < y_n = M \qquad \text{with } y_i - y_{i-1} < \varepsilon \text{ for } 1 \le i \le n.$$

Let $E_1 = \{x \in E : m \le f(x) \le y_1\}$, $E_2 = \{x \in E : y_1 < f(x) \le y_2\}$,..., $E_n = \{x \in E : y_{n-1} < f(x) \le M\}$. Now, take $\phi = y_1\chi_{E_1} + y_2\chi_{E_2} + \cdots + y_n\chi_{E_n}$. Since $E_1, E_2, \ldots, E_n$ are all measurable (why?), ϕ is simple and satisfies the inequality $|f - \phi| < \varepsilon$ with no exceptions.

Step II: General case.

We let

$$F_n = \{x \in E : |f(x)| \ge n\}.$$

Then $F_1 \supseteq F_2 \supseteq \cdots$. Note that $m(F_1) \le m(E) < \infty$ and $m(F_\infty) = 0$ by assumption, apply Lemma 4 there exists $N \in \mathbb{N}$ such that

$$m(F_N) < \varepsilon.$$

Now, let $f^* = (-N \vee f) \wedge N$, then $f = f^*$ on E except on a set of measure less than ε. From the result of Step I, there is a simple function ϕ such that $|f^* - \phi| < \varepsilon$ on E. Hence

$$|f - \phi| < \varepsilon \text{ on E except on a set of measure less than } \varepsilon.$$

Corollary: There is a sequence of simple functions ϕ_n such that $\phi_n \to f$ pointwisely almost everywhere on E. If E = [a, b], there are also sequence of step functions and sequence of continuous functions converging to f pointwisely almost everywhere on [a,b].

Proof: Applying Littlewood's 2nd Principle to $\varepsilon = 1/2^n$, there are simple functions ϕ_n and sets A_n with $m(A_n) < 1/2^n$ such that

$$|f - \phi_n| < \frac{1}{2^n} \qquad \text{on } E\backslash A_n.$$

Let $A = \overline{\lim}\, A_n := \bigcap_{k=1}^{\infty}\left(\bigcap_{n=k}^{\infty} A_n\right)$, then $m(A) = 0$ (why?). The proof is completed by noting that $\phi_n \to f$ pointwisely on $E\backslash A$.

> *Notes* In fact, the sequence ϕ_n can be chosen so that $\phi_n \to f$ pointwisely everywhere on E. For example, we can first divide the interval [–n, n] into $2n^2$ subintervals such that each subinterval has length 1/n, i.e. choose
>
> $$-n = y_0 < y_1 < \cdots < y_{2n^2} = n$$
>
> such that $y_i - y_{i-1} = 1/n$ for all i. Then let
>
> $$\phi_n(x) = \begin{cases} y_i & \text{if } y_i \le f(x) < y_{i+1} \text{ for some } i \\ n & \text{if } f(x) \ge n \\ -n & \text{if } f(x) < -n \end{cases}$$

Theorem 8: Littlewood's 3rd Principle/Egoroff's Theorem

Let $E \in \mathfrak{M}$ with $m(E) < \infty$, $f: E \to (-\infty, \infty)$ be measurable and $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions on E such that

$$f_n \to f \qquad \text{a.e. on } E.$$

Then for any $\eta > 0$ there is a (measurable) subset S of E with $m(S) < \eta$ such that

$$f_n \to f \qquad \text{uniformly on } E\backslash S.$$

> *Notes* Again, the condition $m(E) < \infty$ cannot be dropped. Otherwise $f_n = \chi_{[n,\infty)}$ and $f = 0$ would be a counter example.

Proof: We claim that for any $\varepsilon > 0$ and $\delta > 0$, there exists $A \subseteq E$ with $m(A) < \delta$ and $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \varepsilon \qquad \text{whenever } n \geq N \text{ and } x \in E \setminus A.$$

Be careful the above statement is not saying that $f_n \to f$ uniformly on $E \setminus A$ since A depends on ε and δ.

To prove our claim, we let

$$G_n = \{x \in E : |f_n(x) - f(x)| \geq \varepsilon\}$$

and

$$G = \overline{\lim}\, G_n := \bigcap_{n\in\mathbb{N}} E_n, \qquad \text{where } E_n = \bigcup_{k\geq n} G_k.$$

Note that if $x \in G$ then $x \in E_n$ for all $n \in \mathbb{N}$, it follows that $f_n(x) \nrightarrow f(x)$. Since the set of all x such that $f_n(x) \nrightarrow f(x)$ is of measure zero, we have $m(G) = 0$. Note also that $m(E_1) < \infty$ and E_n "decreases" to G, so $\lim m(E_n) = m(G) = 0$ by Lemma 4. There is $N \in \mathbb{N}$ such that $m(E_N) < \delta$. This N, together with $A := E_N$, proved our claim.

Now, let $\eta > 0$ be given. Apply the above result to $\varepsilon = 1/k$ and $\delta = \eta/2^k$, we obtain A_k with $m(A_k) < \eta/2^k$ and $N_k \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \frac{1}{k} \qquad \text{whenever } n \geq N_k \text{ and } x \in E \setminus A_k.$$

Let $S = \bigcup_{k\in\mathbb{N}} A_k$, then $m(S) \leq \sum_{k=1}^{\infty} m(A_k) < \eta$ and $|f_n(x) - f(x)| < 1/k$ whenever $n \geq N_k$ and $x \in E\backslash S$. Hence, $f_n \to f$ uniformly on $E\backslash S$.

Self Assessment

Fill in the blanks:

1. Every open subset V of $\mathbb{R}$ is a .... of disjoint open intervals.
2. The family of all measurable sets is denoted by $\mathfrak{M}$. We will see later $\mathfrak{M}$ is a σ-algebra and translation-invariant containing all intervals. The set function $m: \mathfrak{M} \to [0, \infty]$ defined by

   $$m(E) = m^*(E) \qquad \text{for all } E \in \mathfrak{M}$$

   is called ....
3. Let X be a .... and Y be a topological space. A function $f: X \to Y$ is called measurable if $f^{-1}(V)$ is a measurable set in X for every open set V in Y.
4. Let $E \in \mathfrak{M}$, $f: E \to [-\infty, \infty]$ and $g: E \to [-\infty, \infty]$. If $f = g$ almost everywhere on E then the .... of f and g are the same.

9.7 Summary

- The definition of outer measure of sets.
- Outer measure of an interval is its length.
- Some important properties of Outer measure.
- The definition of Measurable sets.
- Countable union of measurable sets is also measurable.
- Countable intersection of measurable sets is also measurable.
- Every Borel set is measurable.
- Littlewood's First Principle.

9.8 Keywords

Lindelof's Theorem: Let C be a collection of open subsets of $\mathbb{R}$. Then there is a countable sub-collection $\{O_i\}_{i \in \mathbb{N}}$ of C such that

$$\bigcup_{O \in C} O = \bigcup_{i=1}^{\infty} O_i$$

Lebesgue Measure: The family of all measurable sets is denoted by $\mathfrak{M}$. We will see later $\mathfrak{M}$ is a σ-algebra and translation-invariant containing all intervals. The set function m: $\mathfrak{M} \to [0, \infty]$ defined by

$$m(E) = m^*(E) \qquad \text{for all } E \in \mathfrak{M}$$

is called Lebesgue measure.

$$\bigcup_{O \in C} O = \bigcup_{i=1}^{\infty} O_i$$

Littlewood's 1st Principle: Every measurable set of finite measure is nearly a finite union of disjoint open intervals, in the sense.

Measurable Functions: A function f: $E \to [-\infty, \infty]$ is said to be measurable (or measurable on E) if $E \in \mathfrak{M}$ and

$$f^{-1}((a, \infty]) \in \mathfrak{M}$$

for all $a \in \mathbb{R}$.

9.9 Review Questions

1. Prove that the family M of measurable sets is an algebra.
2. If E_1, E_2, ….En are measurable, prove that $E_1 \cup E_2 \cup \ldots \cup E_n$ is measurable.
3. If E_1 and E_2 are measurable sets, then prove that $E_1 \cup E_2$ is also measurable.
4. Prove that properties (i) to (v) are equivalent to (vi), if m*E is finite.
5. Show that if E is measurable, then each translate E + y is also measurable.
6. Show that if E_1 and E_2 are measurable, then $m(E_1 \cup E_2) + m(E_1 | E_2) = mE_1 + mE_2$.
7. Let $\{E_i\}$ be a sequence of disjoint measurable sets and A be any set.

   Show that $m^*\left(A \cap \bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m^*(A \cap E_i)$

Answers: Self Assessment

1.	countable union	2.	Lebesgue measure
3.	measurable space	4.	measurability

9.10 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Notes

Unit 10: Measurable Functions and Littlewood's Second Principle

CONTENTS

Objectives

After studying this unit, you will be able to:

- Define measurable functions
- Discuss the Sum, difference; scalar product and product of measurable functions are measurable
- Explain Littlewood's Theorems

Introduction

In this unit we study the concept of measurability. We shall see that measurable functions are basically very robust (or strong or durable) continuous-like functions. We make "continuous-like" precise in Luzin's Theorem, which is where Littlewood got his second principle. We also study the concept of almost everywhere.

10.1 Measurable Functions

A measurable space is a pair $(X, \mathscr{S})$ where X is a set and, $\mathscr{S}$ is a σ-algebra of subsets of X. The elements of, $\mathscr{S}$ are called measurable sets. Recall that a measure space is a triple $(X, \mathscr{S}, \mu)$ where μ is a measure on $\mathscr{S}$; if we leave out the measure we have a measurable space.

In the discussion at the beginning of this unit we saw that in order to define the integral of a function $f : X \to \overline{\mathbb{R}}$, we needed to require that

$$f^{-1}(I) \in \mathscr{S} \text{ for each } I \in \mathscr{S} \text{ and } f^{-1}(a, \infty] \in \mathscr{S} \text{ for each } a \in \mathbb{R}.$$

If these properties hold, we say that f is measurable. It turns out that we can omit the first condition because it follows from the second. Indeed, since

$$f^{-1}(a, b] = f^{-1}(a, \infty] \setminus f^{-1}(b, \infty],$$

> *Notes* As a reminder, for any $A \subseteq \overline{\mathbb{R}}$, $f^{-1}(A) := \{x \in X;\ f(x) \in A\}$, so for instance $f^{-1}(a, \infty] = \{x \in X;\ f(x) \in (a, \infty)\} = \{x \in X;\ f(x) > a\}$, or $f^{-1}(a, \infty) = \{f > a\}$ if you wish to be a probabilist.

and $\mathscr{S}$ is a σ-algebra, if both right-hand sets are in $\mathscr{S}$, then so is the left-hand set. Hence, in order to define the integral of f we just need $f^{-1}[a, \infty] \in \mathscr{S}$ for each $a \in \mathbb{R}$. We are thus led to the following definition:

A function $f: X \to \overline{\mathbb{R}}$ is measurable if $f^{-1}[a, \infty] \in \mathscr{S}$ for each $a \in \mathbb{R}$.

We emphasize that the definition of measurability is not "artificial" but is required by Lebesgue's definition of the integral. If X is the sample space of some experiment, a measurable function is called a random variable; thus,

In probability, random variable = measurable function.

We note that intervals of the sort $(a, \infty]$ are not special, and sometimes it is convenient to use other types of intervals.

Proposition: For a function $f: X \to \overline{\mathbb{R}}$, the following are equivalent:

1. f is measurable.
2. $f^{-1}[-\infty, a] \in \mathscr{S}$ for each $a \in \mathbb{R}$.
3. $f^{-1}[a, \infty] \in \mathscr{S}$ for each $a \in \mathbb{R}$.
4. $f^{-1}[-\infty, a] \in \mathscr{S}$ for each $a \in \mathbb{R}$.

Proof: Since preimages preserve complements, we have

$$(f^{-1}[a, \infty])^c = f^{-1}([a, \infty]^c) = f^{-1}[-\infty, a].$$

Since σ-algebras are closed under complements, we have (1) ⇔ (2). Similarly, the sets in (3) and (4) are complements, so we have (3) ⇔ (4). Thus, we just to prove (1) ⇔ (3). Assuming (1) and writing

$$[a, \infty] = \bigcap_{n=1}^{\infty}\left[a - \frac{1}{n}, \infty\right] \Rightarrow f^{-1}[a, \infty] = \bigcap_{n=1}^{\infty} f^{-1}\left[a - \frac{1}{n}, \infty\right],$$

shows that $f^{-1}[a, \infty] \in \mathscr{S}$ since each $f^{-1}\left[a - \frac{1}{n}, \infty\right] \in \mathscr{S}$ and $\mathscr{S}$ is closed under countable intersections. Thus, (1) ⇒ (3). Similarly,

$$[a, \infty] = \bigcup_{n=1}^{\infty}\left[a + \frac{1}{n}, \infty\right] \Rightarrow f^{-1}(a, \infty] = \bigcup_{n=1}^{\infty} f^{-1}\left[a + \frac{1}{n}, \infty\right],$$

shows that (3) ⇒ (1).

As a consequence of this proposition, we can prove that measurable functions are closed under scalar multiplication. Indeed, let $f : X \to \overline{\mathbb{R}}$ be measurable and let $\alpha \in \mathbb{R}$; we'll show that α f is also measurable. Assume that $\alpha \neq 0$ (the $\alpha = 0$ case is easy) and observe that for any $a \in \mathbb{R}$,

$$(\alpha f)^{-1}[a, \infty] = \{x;\ \alpha f(x) > a\} = \begin{cases} \left\{x;\ f(x) > \dfrac{a}{\alpha}\right\} & \text{if } \alpha > 0, \\ \left\{x;\ f(x) < \dfrac{a}{\alpha}\right\} & \text{if } \alpha < 0 \end{cases}$$

$$= \begin{cases} f^{-1}\left[\dfrac{a}{\alpha}, \infty\right] & \text{if } \alpha > 0, \\ f^{-1}\left[-\infty, < \dfrac{a}{\alpha}\right] & \text{if } \alpha < 0. \end{cases}$$

By Proposition, each set on the right is measurable, thus so is $(\alpha f)^{-1}[a, \infty]$. We'll analyze more algebraic properties of measurable functions in the next section.

We now give some examples of measurable functions.

Example: Let $X = \mathbb{R}^n$ with Lebesgue measure. Then any continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is measurable because for any $a \in \mathbb{R}$, by continuity (the inverse of any open set is open),

$$f^{-1}[a, \infty] = f^{-1}(a, \infty)$$

(where we used that f does not take the value ∞) is an open subset of $\mathbb{R}^n$. Since open sets are measurable, it follows that f is measurable.

Thus, for Lebesgue measure, continuity implies measurability. However, the converse is far from true because there are many more functions that are measurable than continuous. For instance, Dirichlet's function $D : \mathbb{R} \to \mathbb{R}$,

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

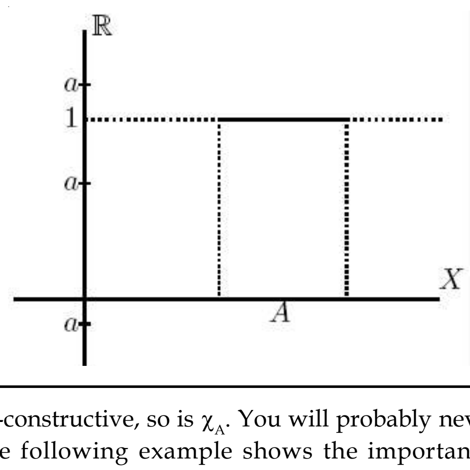
is Lebesgue measurable. Note that D is nowhere continuous. That D is measurable follows from example below and the fact that D is just the characteristic function of $\mathbb{Q} \subseteq \mathbb{R}$, and $\mathbb{Q}$ is measurable.

Example: For a general measure space X and a set $A \subseteq X$, we claim that the characteristic function $\chi_A : X \to \mathbb{R}$ is measurable if and only if the set A is measurable. Indeed, looking at Figure 27.1, we see that

$$\chi_A^{-1}[a, \infty] = \{x \in X; \chi_A(x) > a\} = \begin{cases} X & \text{if } a < 0 \\ A & \text{if } 0 \le a < 1, \\ \varnothing & \text{if } a \ge 1. \end{cases}$$

It follows that $\chi_A^{-1}[a, \infty] \in \mathscr{S}$ for all $a \in \mathbb{R}$ if and only if $A \in \mathscr{S}$, which proves the claim. In particular, there exists a non-Lebesgue measurable function on $\mathbb{R}^n$. In fact, given any non-measurable set $A \subseteq \mathbb{R}^n$, the characteristic function $\chi_A: \mathbb{R}^n \to \mathbb{R}$ is not measurable.

Figure 10.1: Graph of χ_A



Of course, since A is non-constructive, so is χ_A. You will probably never find a non-measurable function in practice. The following example shows the importance of studying extended real-valued functions, instead of just real-valued functions.

Example: Let $X = S^\infty$, where $S = \{0,1\}$, be the sample space for a Monkey-Shakespeare experiment (or any other experiment involving a sequence of Bernoulli trials). Let $f : X \to [0, \infty]$ be the number of times the Monkey types sonnet 18:

$$f(x_1, x_2, x_3, ...) = \text{the number of i's such that } x_i = 1.$$

Notice that $f = \infty$ when the Monkey types sonnet 18 an infinite number of times (in fact, as we see that $f = \infty$ on a set of measure). To show that f is measurable, write f as

$$f = \lim_{n\to\infty} f_n,$$

where f_i is the number of i's in 1, 2,..., n such that $x_i = 1$. Notice that $f_1 \le f_2 \le f_3 \le \cdots$ are non-decreasing, so it follows that for any $a \in \mathbb{R}$,

$$f(x) \le a \Leftrightarrow f_n(x) \le a \text{ for all } n \Leftrightarrow x \in \bigcap_{n=1}^{\infty} \{f_n \le a\}. \{f_n < a\}.$$

Thus,

$$f^{-1}[-\infty, a] = \bigcap_{n=1}^{\infty} f_n^{-1}[-\infty, a].$$

The set $\{f_n \le a\}$ is of the form $A_n \times S \times S \times S \times \cdots$ where $A_n \subseteq S^n$ is the subset of S^n consisting of those points with no more than a total of a entries with 1's. In particular, $\{f_n \le a\} \in \mathscr{R}(\mathscr{C})$ and hence, it belongs to $\mathscr{S}(\mathscr{C})$. Therefore, $\{f \le a\}$ also belongs to $\mathscr{S}(\mathscr{C})$, so f is measurable.

We shall return to this example when we study limits of measurable functions.

As we defined simple functions. For a quick review in the current context of our σ-algebra, $\mathscr{S}$, recall that a simple function (or $\mathscr{S}$-simple function to emphasize the σ-algebra, $\mathscr{S}$) is any function of the form

$$s = \sum_{n=1}^{N} a_n \chi_{A_n},$$

where $a_1,..., a_N \in \mathbb{R}$ and $A_1,..., A_N \in \mathscr{S}$ are pairwise disjoint. We know that we don't have to take the A_n's to be pairwise disjoint, but for proofs it's often advantageous to do so.

Theorem 1: Any Simple Function is Measurable

Proof: Let $s = \sum_{n=1}^{N} a_n \chi_{A_n}$ be a simple function where $a_1,..., a_N \in \mathbb{R}$ and $A_1,..., A_N \in \mathscr{S}$ are pairwise disjoint. If we put $A_{N+1} = X \setminus \{A_1 \cup \cdots \cup A_N\}$ and $a_{N+1} = 0$, then

$$X = A_1 \cup A_2 \cup \cdots \cup A_N \cup A_{N+1},$$

a union of pairwise disjoint sets, and $s = a_n$ on A_n, for each n = 1, 2,..., N + 1. It follows that

$$s^{-1}[a, \infty] = \{x \in X; s(x) > a\} = \bigcup_{n=1}^{N+1} \{x \in A_n; s(x) > a\}$$

$$= \bigcup_{n=1}^{N+1} \{x \in A_n; a_n > a\}.$$

Since

$$\{x \in A_n; a_n > a\} = \begin{cases} A_n & \text{if } a_n > a \\ \varnothing & \text{otherwise}. \end{cases}$$

it follows that $s^{-1}[a, \infty]$ is just a union of elements of $\mathscr{S}$. Thus, s is measurable.

Notes

10.2 Measurability and Continuity

We saw earlier that continuity implies measurability, essentially by definition of continuity in terms of open sets. It turns out that we can directly express measurability in terms of open sets.

Theorem 2: Measurability Criterion

For a function f: $X \to \overline{\mathbb{R}}$, the following are equivalent:

1. f is measurable.
2. $f^{-1}(\{\infty\}) \in \mathscr{S}$ and $f^{-1}(\mathcal{U}) \in \mathscr{S}$ for all open subsets $\mathcal{U} \subseteq \mathbb{R}$.
3. $f^{-1}(\{\infty\}) \in \mathscr{S}$ and $f^{-1}(B) \in \mathscr{S}$ for all Borel sets $B \subseteq \mathbb{R}$.

Proof: To prove that (1) ⇒ (2), observe that

$$\{\infty\} = \bigcap_{n=1}^{\infty}[n,\infty] \Rightarrow f^{-1}(\{\infty\}) = \bigcap_{n=1}^{\infty} f^{-1}[n,\infty].$$

Assuming f is measurable, we have $f^{-1}[n, \infty] \in \mathscr{S}$ for each n and since $\mathscr{S}$ is a σ-algebra, it follows that $f^{-1}(\{\infty\}) \in \mathscr{S}$. Also, if $\mathcal{U} \subseteq \mathbb{R}$ is open, then by the Dyadic Cube Theorem we can write $\mathcal{U} = \bigcup_{n=1}^{\infty} I_n$ where $I_n \in \mathscr{I}^1$ for each n. Hence,

$$f^{-1}(\mathcal{U}) = \bigcup_{n=1}^{\infty} f^{-1}(I_n)$$

By measurability, $f^{-1}(I_n) \in \mathscr{S}$ for each n, so $f^{-1}(\mathcal{U}) \in \mathscr{S}$.

To prove that (2) ⇒ (3), we don't have to worry about the preimage of ∞, so we just have to prove that $f^{-1}(B) \in \mathscr{S}$ for all Borel sets $B \subseteq \mathbb{R}$.

$$\mathscr{S}_f = \{A \in \mathbb{R}; f^{-1}(A) \in \mathscr{S}\}$$

is a σ-algebra. Assuming (2) we know that all open sets belong to $\mathscr{S}_f$. Since $\mathscr{S}_f$ is a σ-algebra of subsets of $\mathbb{R}$ and $\mathscr{B}$ is the smallest σ-algebra containing the open sets, it follows that $\mathscr{B} \subseteq \mathscr{S}_f$.

Finally we prove that (3) ⇒ (1). Let $a \in \mathbb{R}$ and note that

$$[a, \infty] = (a, \infty) \cup \{\infty\} \Rightarrow f^{-1}[a, \infty] = f^{-1}(a, \infty) \cup f^{-1}(\{\infty\}).$$

Assuming (3), we have $f^{-1}(\{\infty\}) \in \mathscr{S}$ and since $(a, \infty) \subseteq \mathbb{R}$ is open, and hence is Borel, we also have $f^{-1}(a, \infty) \in \mathscr{S}$. Thus, $f^{-1}(a, \infty] \in \mathscr{S}$, so f is measurable.

We remark that the choice of using +∞ over –∞ in the "$f^{-1}(\{\infty\}) \in \mathscr{S}$" parts of (2) and (3) were arbitrary and we could have used –∞ instead of ∞.

Consider the second statement in the theorem, but only for real-valued functions:

Measurability: A function $f : X \to \mathbb{R}$ is measurable if and only if $f^{-1}(\mathcal{U}) \in \mathscr{S}$ for each open set $\mathcal{U} \subseteq \mathbb{R}$.

One cannot avoid noticing the striking resemblance to the definition of continuity. Recall that for a topological space $(T, \mathscr{T})$, where $\mathscr{T}$ is the topology on a set T.

Continuity: A function $f : T \to \mathbb{R}$ is continuous if and only if $f^{-1}(\mathcal{U}) \in \mathscr{T}$ for each open set $\mathcal{U} \subseteq \mathbb{R}$.

Because of this similarity, one can think about measurability as a type of generalization of continuity. However, speaking philosophically, there are two very big differences between measurable functions and continuous functions as we can see by considering $X = \mathbb{R}^n$ with Lebesgue measure and its usual topology:

(i) There are a lot more measurable functions than continuous functions.

(ii) Measurable functions are closed under a lot more operations than continuous functions are.

To understand Point (i), recall from that all continuous functions on $\mathbb{R}^n$ are measurable; in contrast, there are measurable functions that are highly discontinuous (like Dirichlet's function). There are more measurable functions than continuous functions because there are a lot more measurable sets than there are open sets. For example, not only are open sets measurable but so are points, Cantor-type sets, G_δ sets, F_σ sets, etc. We shall see that, just like continuous functions, measurable functions are closed under all the usual arithmetic operations such as addition, multiplication, etc. What exemplifies Point (ii) is that measurable functions are closed under all limiting operations. For example, a limit of measurable functions is always measurable. This stands in stark contrast to continuous functions. Indeed, that the characteristic function of a Cantor set can be expressed as a limit of continuous functions. The reason that measurable functions are closed under more operations is that measurable sets are closed under operations (e.g. countable intersections and complements) that open sets are not.

Measurable functions are similar to continuous functions, but there are more of them and they are more robust. Littlewood's second principle shows exactly how "similar" measurable functions are to continuous functions.

10.3 Littlewood's Second Principle

We now continue our discussion of Littlewood's Principles where we stated the first principle;

There are three principles, roughly expressible in the following terms: Every [finite Lebesgue] measurable set is nearly a finite union of intervals; every measurable function is nearly continuous; every convergent sequence of measurable functions is nearly uniformly convergent.

—***Nikolai Luzin***

The third principle is contained in Egorov's theorem, which we'll get to in the next topic. The second principle comes from Luzin's Theorem, named after Nikolai Nikolaevich Luzin (1883-1950) who proved it in 1912 [70], and this theorem makes precise Littlewood's comment that any Lebesgue measurable function is "nearly continuous".

Theorem 3: Luzin's Theorem

Let $X \subseteq \mathbb{R}^n$ be Lebesgue measurable and let $f : X \to \mathbb{R}$ be a Lebesgue measurable function. Then given any $\varepsilon > 0$, there exists a closed set $C \subseteq \mathbb{R}^n$ such that $C \subseteq X$, $m(X\setminus C) < \varepsilon$, and f is continuous on C.

Proof: Here we follow Feldman's [38] proof that only uses Littlewood's First Principle. Luzin's theorem is commonly proved using Egorov's theorem and the fact that every measurable function is the limit of simple functions.

Step 1: We first prove the theorem only requiring that C be measurable; this proof is yet another example of the "$\frac{\varepsilon}{2^k}$–trick." Let $\{\mathcal{V}_k\}$ be a countable basis of open sets in $\mathbb{R}$; this means that every open set in $\mathbb{R}$ is a union of countably many $\mathcal{V}_k$'s. (For example, take the $\mathcal{V}_k$'s as open intervals with rational end points.) Let $\varepsilon > 0$. Then, since $f^{-1}(\mathcal{V}_k)$ is measurable, by Littlewood's First Principle there is an open set $\mathcal{U}_k$ such that

$$f^{-1}(\mathcal{V}_k) \subseteq \mathcal{U}_k \quad \text{and} \quad m(\mathcal{U}_k \setminus f^{-1}(\mathcal{V}_k)) < \frac{\varepsilon}{2^k}.$$

Now put

$$A := \bigcup_{k=1}^{\infty} (\mathcal{U}_k \setminus f^{-1}(\mathcal{V}_k)).$$

Then A is measurable and

$$m(A) \leq \sum_{k=1}^{\infty} m(\mathcal{U}_k \setminus f^{-1}(\mathcal{V}_k)) < \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon.$$

Notes

If we can prove that

$$g := f|_{X\setminus A} : X \setminus A \to \mathbb{R}$$

is continuous, then we have proven our theorem with C = X \ A (modulo the closedness condition). Since $\{\mathcal{V}_k\}$ is a basis for the topology of $\mathbb{R}$ to prove that g is continuous all we have to do is prove that for each k, $g^{-1}(\mathcal{V}_k)$ is open in $X\setminus A$. To prove this, we shall prove that

$$g^{-1}(\mathcal{V}_k) = (X\setminus A) \cap \mathcal{U}_k; \tag{3.1}$$

then, since $\mathcal{U}_k$ is an open subset of $\mathbb{R}^n$, it follows that $g^{-1}(\mathcal{V}_k)$ is open in $X\setminus A$ and we're done. Now to prove the desired equality note that, by definition of g, we have

$$g^{-1}(\mathcal{V}_k) = (X\setminus A) \cap f^{-1}(\mathcal{V}_k) \subseteq (X\setminus A) \cap \mathcal{U}_k,$$

since $f^{-1}(\mathcal{V}_k) \subseteq \mathcal{U}_k$. On the other hand, observe that

$$\begin{aligned} x \in (X\setminus A) \cap \mathcal{U}_k &\Rightarrow x \notin A,\ x \in \mathcal{U}_k \\ &\Rightarrow x \notin (\mathcal{U}_k\setminus f^{-1}(\mathcal{V}_k)),\ x \in \mathcal{U}_k \\ &\Rightarrow x \in f^{-1}(\mathcal{V}_k). \end{aligned}$$

In the second implication we used that $A = \bigcup_{j=1}^{\infty}(\mathcal{U}_j\setminus f^{-1}(\mathcal{V}_j))$ so $x \notin A$ implies, in particular, that $x \notin (\mathcal{U}_k\setminus f^{-1}(\mathcal{V}_k))$. Therefore,

$$(X\setminus A) \cap \mathcal{U}_k \subseteq (X\setminus A) \cap f^{-1}(\mathcal{V}_k),$$

which completes the proof of (3.1).

Step 2: We now require that C be closed. Given $\varepsilon > 0$ by Step 1 we can choose a measurable set $B \subseteq X$ such that $\mathfrak{m}(X\setminus B) < \varepsilon/2$ and f is continuous on B. By Littlewood's First Principle we can choose a closed set $C \subseteq \mathbb{R}^n$ such that $C \subseteq B$ and $\mathfrak{m}(B\setminus C) < \varepsilon/2$. Since

$$X\setminus C = (X \setminus B) \cup (B \setminus C),$$

we have

$$\mathfrak{m}(X \setminus C) \le \mathfrak{m}(X \setminus B) + \mathfrak{m}(B \setminus C) < \varepsilon.$$

Also, since $C \subseteq B$ and f is continuous on B, the function f is automatically continuous on the smaller set C. This completes the proof of our theorem.

We shall see that Luzin's theorem holds not just for $\mathbb{R}^n$ but for topological spaces as well.

10.4 Borel Measurability on Topological Spaces

Recall that the collection of Borel subsets of a topological space is the σ-algebra generated by the open sets. For a measurable space $(T, \mathscr{S})$ where T is a topological space with $\mathscr{S}$ its Borel subsets, we call a measurable function $f: T \to \overline{\mathbb{R}}$ Borel measurable to emphasize that the σ-algebra $\mathscr{S}$ is the one generated by the topology and it is not just any σ-algebra on T. For example, a Borel measurable function on $\mathbb{R}^n$ is a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ such that $f^{-1}(a, \infty] \in \mathscr{B}^n$ for all $a \in \mathbb{R}$.

Proposition: Any continuous real-valued function on a topological space is Borel measurable.

The proof of this proposition follows word-for-word the $\mathbb{R}^n$ case in Example, so we omit its proof. A nice thing about Borel measurability is that it behaves well under composition.

Proposition: If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is Borel measurable and $g : X \to \mathbb{R}$ is measurable, where X is an arbitrary measurable space, then the composition,

$$f \circ g : X \to \overline{\mathbb{R}}$$

is measurable.

Proof: Given $a \in \mathbb{R}$, we need to show that

$$(f \circ g)^{-1}(a, \infty] = g^{-1}(f^{-1}(a, \infty]) \in \mathscr{S}.$$

The function $f : R \to \overline{\mathbb{R}}$ is, by assumption, Borel measurable, so $f^{-1}(a, \infty] \in \mathscr{B}^1$. The function $g : X \to \mathbb{R}$ is measurable, so by Part (3) of Theorem 3.5, $g^{-1}(f^{-1}(a, \infty]) \in \mathscr{S}$. Thus, $f \circ g$ is measurable.

Example: If $g : X \to \mathbb{R}$ is measurable, and $f : \mathbb{R} \to \mathbb{R}$ is the characteristic function of the rationals, which is Borel measurable, then Proposition 3.8 shows that the rather complicated function

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

is measurable. Other, more normal looking, functions of g that are measurable include $e^{g(x)}$, $\cos g(x)$, and $g(x)^2 + g(x) + 1$.

10.5 The Concept of Almost Everywhere

Let $(X, \mathscr{S}, \mu)$ be a measure space. We say that a property holds almost everywhere (written a.e.) if the set of points where the property fails to hold is a measurable set with measure zero. For example, we say that a sequence of functions $\{f_n\}$ on X converges a.e. to a function f on X, written $f_n \to f$ a.e., if $f(x) = \lim_{n\to\infty} f_n(x)$ for each $x \in X$ except on a measurable set with measure zero. Explicitly,

$$f_n \to f \text{ a.e.} \Leftrightarrow A := \{x; f(x) \neq \lim_{n\to\infty} f_n(x)\} \in \mathscr{S} \text{ and } \mu(A) = 0.$$

For another example, given two functions f and g on X, we say that f = g a.e. if the set of points where $f \neq g$ is measurable with measure zero:

$$f = g \text{ a.e.} \Leftrightarrow A := \{x; f(x) \neq g(x)\} \in \mathscr{S} \text{ and } \mu(A) = 0.$$

If g is measurable and f = g a.e., then one might think that f must also be measurable. However, as you'll see in the following proof, to always make this conclusion we need to assume completeness.

Proposition: Assume that μ is a complete measure and let $f, g : X \to \overline{\mathbb{R}}$. If g is measurable and f = g a.e., then f is also measurable.

Proof: Assume that g is measurable and f = g a.e., so that the set $A = \{x; f(x) \neq g(x)\}$ is measurable with measure zero. Observe that for any $a \in \mathbb{R}$,

$$\begin{aligned} f^{-1}(a, \infty] &= \{x \in X; f(x) > a\} \\ &= \{x \in A; f(x) > a\} \cup \{x \in A^c; f(x) > a\} \\ &= \{x \in A; f(x) > a\} \cup \{x \in A^c; g(x) > a\} \\ &= \{x \in A; f(x) > a\} \cup (A^c \cap g^{-1}(a, \infty]). \end{aligned}$$

The first set is a subset of A, which is measurable and has measure zero, hence the first set is measurable. g is measurable, so the second set is measurable too, hence f is measurable.

For instance, this proposition holds for Lebesgue measure since Lebesgue measure is complete.

Self Assessment

Fill in the blanks:

1. A measurable space is a pair $(X, \mathscr{S})$ where X is a set and, $\mathscr{S}$ is a σ-algebra of subsets of X. The elements of, $\mathscr{S}$ are called ....
2. For Lebesgue measure, continuity implies ....
3. A function $f : T \to \mathbb{R}$ is continuous if and only if $f^{-1}(\mathcal{U}) \in \mathscr{T}$ for each open set ....
4. Measurable functions are similar to ...., but there are more of them and they are more robust.
5. .... principle shows exactly how "similar" measurable functions are to continuous functions.
6. Any continuous real-valued function on a topological space is ....
7. If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is Borel measurable and $g : X \to \mathbb{R}$ is measurable, where X is an arbitrary measurable space, then the composition, ... is measurable.

10.6 Summary

- A measurable space is a pair $(X, \mathscr{S})$ where X is a set and, $\mathscr{S}$ is a σ-algebra of subsets of X. The elements of, $\mathscr{S}$ are called measurable sets. Recall that a measure space is a triple $(X, \mathscr{S}, \mu)$ where μ is a measure on $\mathscr{S}$; if we leave out the measure we have a measurable space.

  In the discussion at the beginning of this chapter we saw that in order to define the integral of a function $f : X \to \overline{\mathbb{R}}$, we needed to require that

  $$f^{-1}(I) \in \mathscr{S} \text{ for each } I \in \mathscr{I}^1 \text{ and } f^{-1}[a, \infty] \in \mathscr{S} \text{ for each } a \in \mathbb{R}.$$

- If these properties hold, we say that f is measurable. It turns out that we can omit the first condition because it follows from the second. Indeed, since

  $$f^{-1}[a, b] = f^{-1}[a, \infty]\backslash f^{-1}[b, \infty].$$

- There are three principles, roughly expressible in the following terms: Every [finite Lebesgue] measurable set is nearly a finite union of intervals; every measurable function is nearly continuous; every convergent sequence of measurable functions is nearly uniformly convergent.
- The third principle is contained in Egorov's theorem, which we'll get to in the next section. The second principle comes from Luzin's Theorem, named after Nikolai Nikolaevich Luzin (1883-1950) who proved it in 1912 [70], and this theorem makes precise Littlewood's comment that any Lebesgue measurable function is "nearly continuous".
- Any continuous real-valued function on a topological space is Borel measurable.
- The proof of this proposition follows word-for-word the $\mathbb{R}^n$ case in Example, so we omit its proof. A nice thing about Borel measurability is that it behaves well under composition.
- If $f : \mathbb{R} \to \overline{\mathbb{R}}$ is Borel measurable and $g : X \to \mathbb{R}$ is measurable, where X is an arbitrary measurable space, then the composition,

  $$f \circ g : X \to \overline{\mathbb{R}}$$

  is measurable.

10.7 Keywords

Measurable Sets: A measurable space is a pair (X, $\mathscr{S}$) where X is a set and, $\mathscr{S}$ is a σ-algebra of subsets of X. The elements of, $\mathscr{S}$ are called measurable sets.

Measurability Criterion: For a function f: X → $\overline{\mathbb{R}}$, the following are equivalent:

1. f is measurable.
2. $f^{-1}(\{\infty\}) \in \mathscr{S}$ and $f^{-1}(\mathcal{U}) \in \mathscr{S}$ for all open subsets $\mathcal{U} \subseteq \mathbb{R}$.
3. $f^{-1}(\{\infty\}) \in \mathscr{S}$ and $f^{-1}(B) \in \mathscr{S}$ for all Borel sets $B \subseteq \mathbb{R}$.

Measurable: A function $f : X \to \mathbb{R}$ is measurable if and only if $f^{-1}(\mathcal{U}) \in \mathscr{S}$ for each open set $\mathcal{U} \subseteq \mathbb{R}$.

One cannot avoid noticing the striking resemblance to the definition of continuity. Recall that for a topological space (T, $\mathscr{T}$), where $\mathscr{T}$ is the topology on a set T.

Luzin's Theorem: Let $X \subseteq \mathbb{R}^n$ be Lebesgue measurable and let $f : X \to \mathbb{R}$ be a Lebesgue measurable function. Then given any ε > 0, there exists a closed set $C \subseteq \mathbb{R}^n$ such that $C \subseteq X$, $m(X\setminus C) < \varepsilon$, and f is continuous on C.

Borel Measurable: Any continuous real-valued function on a topological space is Borel measurable.

10.8 Review Questions

1. (a) Prove that a non-negative function f is measurable if and only if for all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$ with $0 \le k \le 2^{2n} - 1$, the sets $f^{-1}(k/2^n, (k + 1)/2^n]$ and $f^{-1}(2^n, \infty]$, are measurable.

   (b) Prove that an extended real-valued function f is measurable if and only if $f^{-1}(\{\infty\})$ and all sets of the form $f^{-1}(k/2^n, (k + 1)/2^n]$, where $k \in \mathbb{Z}$ and $n \in \mathbb{N}$, are measurable.

   (c) If $\{a_n\}$ is any countable dense subset of $\mathbb{R}$, prove that f is measurable if and only if $f^{-1}(\{\infty\})$ and all sets of the form $f^{-1}(a_m, a_n]$, where $m, n \in \mathbb{N}$, are measurable.

2. Here are some problems dealing with non-measurable functions.

   (a) Find a non-Lebesgue measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $|f|$ is measurable.

   (b) Find a non-Lebesgue measurable function $f : \mathbb{R} \to \mathbb{R}$ such that f^2 is measurable.

   (c) Find two non-Lebesgue measurable functions $f, g : \mathbb{R} \to \mathbb{R}$ such that both f + g and $f \cdot g$ are measurable.

3. Here are some problems dealing with measurable functions.

   (a) Prove that any monotone function $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable.

   (b) A function $f : \mathbb{R} \to \mathbb{R}$ is said to be lower-semicontinuous at a point $c \in \mathbb{R}$ if for any ε > 0 there is a δ > 0 such that

   $$|x - c| < \delta \Rightarrow f(c) - \varepsilon < f(x).$$

   Intuitively, f is lower-semicontinuous at c if for x near c, f(x) is either near f(c) or greater than f(c). The function f is lower-semicontinuous if it's lower-semicontinuous at all points of $\mathbb{R}$. (To get a feeling for lower-semicontinuity, show that the functions $\chi_{(0,\infty)}$, $\chi_{(-\infty,0)}$, and $\chi_{(-\infty,0)\cup(0,\infty)}$ are lower-semicontinuous at 0.) Prove that any lower-semicontinuous function is Lebesgue measurable.

   (c) A function $f : \mathbb{R} \to \mathbb{R}$ is said to be upper-semicontinuous at a point $c \in \mathbb{R}$ if for any ε > 0 there is a δ > 0 such that

   $$|x - c| < \delta \Rightarrow f(x) < f(c) + \varepsilon.$$

Notes

Intuitively, f is upper-semicontinuous at c if for x near c, f(x) is either near f(c) or less than f(c). The function f is upper-semicontinuous if it's upper-semicontinuous at all points of $\mathbb{R}$. Prove that any upper-semicontinuous function is Lebesgue measurable.

4. We can improve Luzin's Theorem as follows. First prove the

   (i) Tietze Extension Theorem for $\mathbb{R}$; named after Heinrich Tietze (1880-1964) who proved a general result for metric spaces in 1915 [98]. Let $A \subseteq \mathbb{R}$ be a non-empty closed set and let $f_0 : A \to \mathbb{R}$ be a continuous function. Prove that there is a continuous function $f_1 : \mathbb{R} \to \mathbb{R}$ such that $f_1|_A = f_0$, and if f_0 is bounded in absolute value by a constant M, then we may take f_1 to the have the same bound. Suggestion: Show that $\mathbb{R}\backslash A$ is a countable union of pairwise disjoint open intervals. Extend f_0 linearly over each of the open intervals to define f_1.

   (ii) Using Luzin's Theorem for n = 1, given a measurable function $f : X \to \mathbb{R}$ where $X \subseteq \mathbb{R}$ is measurable, prove that there is a closed set $C \subseteq \mathbb{R}$ such that $C \subseteq X$, $\mathfrak{m}(X\backslash C) < \varepsilon$, and a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that f = g on C. Moreover, if f is bounded in absolute value by a constant M, then we may take g to have the same bound as f.

5. Here are some generalizations of Luzin's Theorem.

   (i) Let μ be a σ-finite regular Borel measure on a topological space X, let $f : X \to \mathbb{R}$ be measurable, and let $\varepsilon > 0$. On "Littlewood's First Principle(s) for regular Borel measures," prove that there exists a closed set $C \subseteq X$ such that $\mathfrak{m}(X\backslash C) < \varepsilon$ and f is continuous on C.

6. Here we present Leonida Tonelli's (1885-1946) integral published in 1924 [100]. Let $f : [a, b] \to \mathbb{R}$ be a bounded function, say $|f| \leq M$ for some constant M. f is said to be quasi-continuous (q.c.) if there is a sequence of closed sets $C_1, C_2, C_3, \ldots \subseteq [a, b]$ with $\lim_{n\to\infty} \mathfrak{m}(C_n) = b - a$ and a sequence of continuous functions $f_1, f_2, f_3, \ldots$ where for each n, $f_n : [a, b] \to \mathbb{R}$, $f = f_n$ on C_n, and $|f_n| \leq M$.

   (i) Let $f : [a, b] \to \mathbb{R}$ be bounded. Prove that f is q.c. if and only if f is measurable. To prove the "if" statement, use Problem 6.

   (ii) Let $f : [a, b] \to \mathbb{R}$ be q.c. and let $\{f_n\}$ be a sequence of continuous functions in the definition of q.c. for f. Let $R(f_n)$ denote the Riemann integral of f_n and prove that the limit $\lim_{n\to\infty} R(f_n)$ exists and its value is independent of the choice of sequence $\{f_n\}$ in the definition of q.c. for f. Tonelli defines the integral of f as

$$\int_a^b f := \lim_{n\to\infty} R(f_n).$$

   It turns out that Tonelli's integral is exactly the same as Lebesgue's integral.

7. We show that the composition of two Lebesgue measurable function is not necessarily Lebesgue measurable. Let φ and M be the homeomorphism and Lebesgue measurable set, respectively. Let $g = \chi_M$. Show that $g \circ \varphi^{-1}$ is not Lebesgue measurable. Note that both φ^{-1} and g are Lebesgue measurable.

8. Prove the Banach-Sierpinski Theorem, proved in 1920 by Stefan Banach (1892-1945) and Waclaw Sierpinski (1882-1969), which states that if $f : \mathbb{R} \to \mathbb{R}$ is additive and Lebesgue measurable, then $f(x) = f(1)x$ for all $x \in \mathbb{R}$. Suggestion: Observe that

$$\mathbb{R} = \bigcup_{n=1}^{\infty} \{x \in \mathbb{R}; |f(x)| \leq n\}.$$

   Prove that for some $n \in \mathbb{N}$, the set $\{x \in \mathbb{R};\ |f(x)| \leq n\}$ has positive measure.

Answers: Self Assessment

1. measurable sets
2. measurability
3. $\mathcal{U} \subseteq \mathbb{R}$
4. continuous functions
5. Littlewood's second
6. Borel measurable
7. $f \circ g : X \to \overline{\mathbb{R}}$

10.9 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 11: Sequences of Functions and Littlewood's Third Principle

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the limsups and liminfs of sequences
- Describe operations on measurable functions
- Explain Littlewood's third principle

Introduction

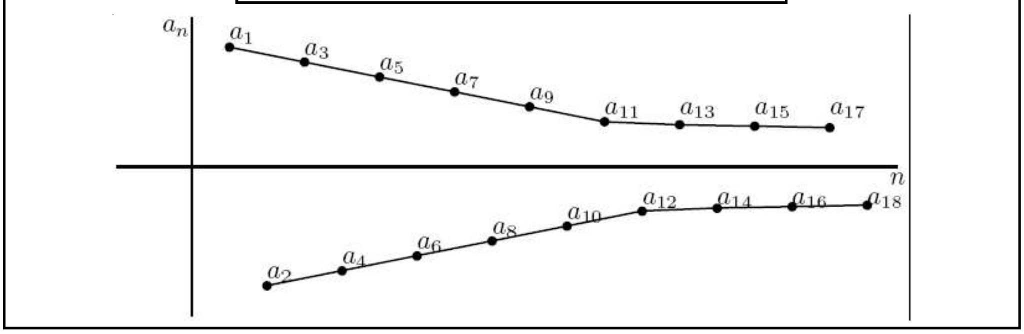
In this unit we continue our study of measurability. We show that measurable functions are very robust in the sense that they are closed under just about any kind of arithmetic or limiting operation that you can imagine: addition, multiplication, division,…, and most importantly, they are closed under just about any conceivable limiting process. We also discuss Littlewood's third principle on limits of measurable functions.

11.1 Limsups and Liminfs of Sequences

Before discussing limits of sequences of functions we need to start by talking about limits of sequences of extended real numbers.

For a sequence $\{a_n\}$ of extended real numbers, we know, in general, that lim a_n does not exist; for example, it can oscillate such as the sequence. However, for the sequence, assuming that the sequence continues the way it looks like it does, it is clear that although limit lim a_n does not exist, the sequence does have an "upper" limiting value, given by the limit of the odd-indexed a_n's and a "lower" limiting value, given by the limit of the even-indexed a_n's. Now, how do we find the "upper" (also called "supremum") and "lower" (also called "infimum") limits of $\{a_n\}$? It turns out there is a very simple way to do so, as we now explain.

Figure 11.1: A Sequence Bouncing Up and Down



Given an arbitrary sequence $\{a_n\}$ of extended real numbers, put

$$s_1 = \sup_{k\geq 1} a_k = \sup\{a_1, a_2, a_3,...\},$$

$$s_2 = \sup_{k\geq 2} a_k = \sup\{a_2, a_3, a_4,...\},$$

$$s_3 = \sup_{k\geq 3} a_k = \sup\{a_3, a_4, a_5,...\},$$

and in general,

$$s_n = \sup_{k\geq n} a_k = \sup\{a_n, a_{n+1}, a_{n+2},...\}.$$

Note that

$$s_1 \geq s_2 \geq s_3 \geq \cdots \geq s_n \geq s_{n+1} \geq \cdots$$

is an non-increasing sequence since each successive s_n is obtained by taking the supremum of a smaller set of elements. Since $\{s_n\}$ is an non-increasing sequence of extended real numbers, the limit $\lim s_n$ exists in $\overline{\mathbb{R}}$; in fact,

$$\lim s_n = \inf_n s_n = \inf\{s_1, s_2, s_3,...\},$$

as can be easily be checked. We define the lim sup of the sequence $\{a_n\}$ as

$$\limsup a_n := \inf_n s_n = \lim s_n = \lim_{n\to\infty} (\sup\{a_n, a_{n+1}, a_{n+2},...\})$$

Note that the term "lim sup" of $\{a_n\}$ fits well because $\limsup a_n$ is exactly the limit of a sequence of supremums.

Example: For the sequence a_n shown in Figure 28.1, we have

$$s_1 = a_1, \quad s_2 = a_3, \quad s_3 = a_3, \quad s_4 = a_5, \quad s_5 = a_5,...,$$

so $\limsup a_n$ is exactly the limit of the odd-indexed a_n's.

We now define the "lower" or "infimum" limit of an arbitrary sequence $\{a_n\}$. Put

$$\iota_1 = \sup_{k\geq 1} a_k = \inf\{a_1, a_2, a_3,...\},$$

$$\iota_2 = \sup_{k\geq 2} a_k = \inf\{a_2, a_3, a_4,...\},$$

$$\iota_3 = \sup_{k\geq 3} a_k = \inf\{a_3, a_4, a_5,...\},$$

and in general,

$$\iota_n = \sup_{k \geq n} a_k = \inf\{a_n, a_{n+1}, a_{n+2}, \ldots\}.$$

Note that

$$\iota_1 \leq \iota_2 \leq \iota_3 \leq \cdots \leq \iota_n \leq \iota_{n+1} \leq \cdots$$

is an non-decreasing sequence since each successive ι_n is obtained by taking the infimum of a smaller set of elements. Since $\{\iota_n\}$ is an non-decreasing sequence, the limit lim ι_n exists, and equals supn ι_n. We define the lim inf of the sequence $\{a_n\}$ as

$$\liminf a_n := \sup_n \iota_n = \lim \iota_n = \lim_{n \to \infty} (\inf\{a_n, a_{n+1}, a_{n+2}, \ldots\}).$$

Note that the term "lim inf" of $\{a_n\}$ fits well because lim inf a_n is the limit of a sequence of infimums.

Example: For the sequence a_n shown in Figure 28.1, we have

$$\iota_1 = a_2, \quad \iota_2 = a_2, \quad \iota_3 = a_4, \quad \iota_4 = a_4, \quad \iota_5 = a_6, \ldots,$$

so lim inf a_n is exactly the limit of the even-indexed a_n's.

The following lemma contains some useful properties of limsup's and liminf's. Since its proof really belongs in a lower-level analysis.

Lemma: Let $A \subseteq \overline{\mathbb{R}}$ be non-empty and let $\{a_n\}$ be a sequence of extended real numbers.

1. $\sup A = -\inf(-A)$ and $\inf A = -\sup(-A)$, where $-A = \{-a; a \in A\}$.
2. $\limsup a_n = -\liminf(-a_n)$ and $\liminf a_n = -\limsup(-a_n)$.
3. $\lim a_n$ exists as an extended real number if and only if $\limsup a_n = \liminf a_n$, in which case,

$$\lim a_n = \limsup a_n = \liminf a_n.$$

4. If $\{b_n\}$ is another sequence of extended real numbers and $a_n \leq b_n$ for all n sufficiently large, then

$$\liminf a_n \leq \liminf b_n \quad \text{and} \quad \limsup a_n \leq \limsup b_n.$$

11.2 Operations on Measurable Functions

Let $\{f_n\}$ be a sequence of extended-real valued functions on a measure space $(X, \mathscr{S}, \mu)$. We define the functions sup f_n, inf f_n, lim sup f_n, and lim inf f_n, by applying these limit operations pointwise to the sequence of extended real numbers $\{f_n(x)\}$ at each point $x \in X$. For example,

$$\limsup f_n : X \to \overline{\mathbb{R}}$$

is the function defined by

$$(\limsup f_n)(x) := \limsup(f_n(x)) \quad \text{at each } x \in X.$$

We define the limit function lim f_n by

$$(\lim f_n)(x) := \lim_{n \to \infty} (f_n(x))$$

at those points $x \in X$ where the right-hand limit exists.

We now show that limiting operations don't change measurability.

Theorem 1: Limits preserve measurability

Notes

If $\{f_n\}$ is a sequence of measurable functions, then the functions

$$\sup f_n, \quad \inf f_n, \quad \limsup f_n, \quad \text{and} \quad \liminf f_n$$

are all measurable. If the limit $\lim_{n\to\infty} f_n(x)$ exists at each $x \in X$, then the limit function $\lim f_n$ is measurable. For instance, if the sequence $\{f_n\}$ is monotone, that is, either non-decreasing or non-increasing, then $\lim f_n$ is everywhere defined and it is measurable.

Proof: To prove that $\sup f_n$ is measurable, we just have to show that $(\sup f_n)^{-1}[-\infty, a] \in \mathscr{S}$ for each $a \in \mathbb{R}$. However, this is easy because by definition of supremum, for any $a \in \mathbb{R}$,

$$\sup\{f_1(x), f_2(x), f_3(x),...\} \le a \quad \Leftrightarrow \quad f_n(x) \le a \text{ for all } n,$$

therefore

$$(\sup f_n)^{-1}[-\infty, a] = \{x; \sup f_n(x) \le a\} = \bigcap_{n=1}^{\infty} \{x; f_n(x) \le a\}$$

$$= \bigcap_{n=1}^{\infty} f_n^{-1}[-\infty, a].$$

Since each f_n is measurable, we have $f_n^{-1}[-\infty, a] \in \mathscr{S}$, so $(\sup f_n)^{-1}[-\infty, a] \in \mathscr{S}$ as well. Using an analogous argument one can show that $\inf f_n$ is measurable.

To prove that $\limsup f_n$ is measurable, note that by definition of lim sup,

$$\limsup f_n := \inf_n s_n,$$

where $s_n = \sup_{k\ge n} f_k$. Since the sup and inf of a sequence of measurable functions are measurable, we know that s_n is measurable for each n and hence $\limsup f_n = \inf_n s_n$ is measurable. An analogous argument can be used to show that $\liminf f_n$ is measurable (just note that $\liminf f_n = \sup \iota_n$ where $\iota_n = \inf_{k\ge n} f_k$).

If the limit function $\lim f_n$ is well-defined, then by Part (3) of above Lemma we know that $\lim f_n = \limsup f_n$ (= $\liminf f_n$). Thus, $\lim f_n$ is measurable.

In particular, if f is a function on X and if $f = \lim s_n$, where the s_n's are simple function (which are measurable by Theorem 3.4), then f is measurable.

Example: Let $X = S^\infty$, where $S = \{0,1\}$, a sample space for the Monkey-Shakespeare experiment (or any other sequence of Bernoulli trials), and let $f : X \to [0, \infty]$ be the random variable given by the number of times the Monkey types sonnet 18. Then

$$f(x) = \sum_{n=1}^{\infty} x_n$$

That is,

$$f = \sum_{n=1}^{\infty} \chi_{A_n} = \lim_{n\to\infty} \sum_{k=1}^{n} \chi_{A_k},$$

where $A_n = S \times S \times \cdots \times S \times \{1\} \times S \times S \times \cdots$ where $\{1\}$ is in the n-th slot. Since each A_n is measurable, it follows that each χ_{A_n} is measurable and hence so is f.

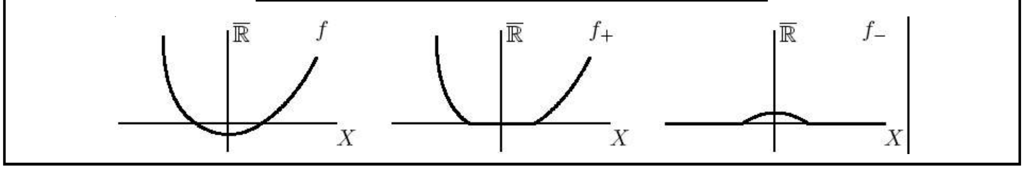
Given $f: X \to \overline{\mathbb{R}}$, we define its non-negative part $f_+: X \to [0, \infty]$ and its non-positive part $f_-: X \to [0, \infty]$ by

$$f_+ := \max\{f, 0\} = \sup\{f, 0\} \quad \text{and} \quad f_- := -\min\{f, 0\} = -\inf\{f, 0\}.$$

Notes

See Figure 11.2 for graphs of $f_{\pm}$. One can check that

Figure 11.2: Graphs of a Function f, f+, and f–



$$f = f_+ - f_- \quad \text{and} \quad |f| = f_+ + f_-.$$

Assuming f is measurable, f_+ and $-f_-$ are also measurable. Also, since measurability is preserves under scalar multiplication $f_- = -(-f_-)$ is measurable. In particular, the equality $f = f_+ - f_-$ shows that any measurable function can be expressed as the difference of non-negative measurable functions.

Theorem 2: Characterization of measurability

A function is measurable if and only if it is the limit of simple functions. Moreover, if the function is nonnegative, the simple functions can be taken to be a non-decreasing sequence of non-negative simple functions.

Proof: Consider first the non-negative case. Let f: $X \to [0, \infty]$ be measurable. For each $n \in \mathbb{N}$, consider the simple function that we constructed at the very beginning of this chapter:

$$s_n(x) = \begin{cases} 0 & \text{if } 0 \le f(x) \le \frac{1}{2^n} \\ \frac{1}{2^n} & \text{if } \frac{1}{2^n} < f(x) \le \frac{2}{2^n} \\ \frac{2}{2^n} & \text{if } \frac{2}{2^n} < f(x) \le \frac{3}{2^n} \\ \vdots & \vdots \\ \frac{2^{2n}-1}{2^n} & \text{if } \frac{2^{2n}-1}{2^n} < f(x) \le \frac{2^{2n}}{2^n} = 2^n \\ 2^n & \text{if } f(x) > 2^n. \end{cases}$$

See Figure 28.3 for an example of a function f and pictures of the corresponding s_1, s_2, and s_3. Note that s_n is a simple function because we can write

$$s_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \chi_{A_{nk}} + 2^n \chi_{B_n},$$

where

$$A_{nk} = f^{-1}\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right] \quad \text{and} \quad B_n = f^{-1}(2_n, \infty].$$

At least if we look at Figure 28.3, it is not hard to believe that in general, the sequence $\{s_n\}$ is always non-decreasing:

$$0 \le s_1 \le s_2 \le s_3 \le s_4 \le \cdots$$

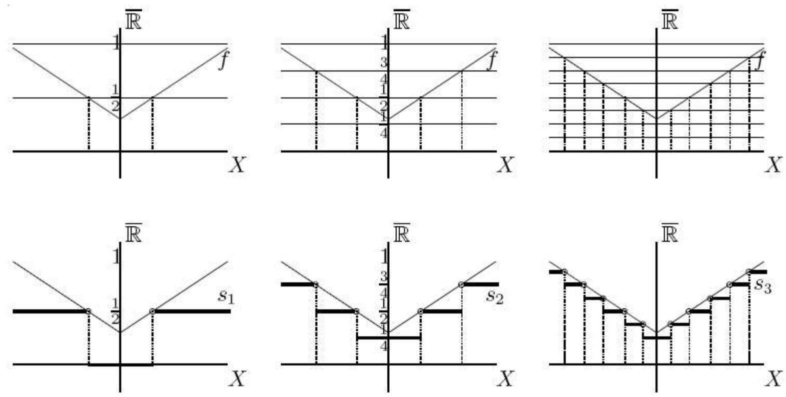
and $\lim_{n\to\infty} s_n(x) = f(x)$ at every point $x \in X$. Because this is so believable looking at Figure, we leave you the pleasure of verifying these facts.

Now let f: $X \to \overline{\mathbb{R}}$ be any measurable function; we need to show that f is the limit of simple functions. To prove this, write $f = f_+ - f_-$ as the difference of its non-negative and non-positive parts. Since $f_{\pm}$ are non-negative measurable functions, we know that f_+ and f_- can be written as limits of simple functions, say s_n^+ and s_n^-, respectively. It follows that

$$f = f_+ - f_- = \lim(s_n^+ - s_n^-)$$

is also a limit of simple functions.

Figure 11.3



Here, f looks like a "V" and is bounded above by 1. The top figures show partitions of the range of f into halves, quarters, then eights and the bottom figures show the corresponding simple functions. It is clear that $s_1 \leq s_2 \leq s_3$.

Using Theorem 2 on limits of simple functions, it is easy to prove that measurable functions are closed under all the usual arithmetic operations. Of course, the proofs aren't particularly difficult to prove directly.

Theorem 3: If f and g are measurable, then $f + g$, $f \cdot g$, $1/f$, and $|f|^p$ where $p > 0$, are also measurable, whenever each expression is defined.

Proof: We need to add the last statement for f + g and 1/f. For 1/f we need f to never vanish and for f + g we don't want f(x) + g(x) to give a non-sense statement such as $\infty - \infty$ or $-\infty + \infty$ at any point $x \in X$.

The proofs that f + g, f g, 1/f, and $|f|^p$ are measurable are all the same: we just show that each combination can be written as a limit of simple functions. By Theorem 2 we can write $f = \lim s_n$ and $g = \lim t_n$ for simple functions s_n, t_n, n = 1, 2, 3, ... Therefore,

$$f + g = \lim(s_n + t_n)$$

and

$$f\,g = \lim(s_n t_n).$$

Since the sum and product of simple functions are simple, it follows that f + g and f g are limits of simple functions, so are measurable.

To see that 1/f and $|f|^p$ are measurable, write the simple function s_n as a finite sum

$$s_n = \sum_k a_{nk} \chi_{A_{nk}},$$

where $A_{n1}, A_{n2}, \ldots \in \mathscr{S}$ are finite in number and pairwise disjoint, and $a_{n1}, a_{n2}, \ldots \in \mathbb{R}$, which we may assume are all non-zero. If we define

$$u_n = \sum_k a_{nk}^{-1} \chi_{A_{nk}} \quad \text{and} \quad v_n = \sum_k |a_{nk}|^p \chi_{A_{nk}},$$

which are simple functions, then a short exercise shows that

$$f^{-1} = \lim u_n \quad \text{and} \quad |f|^p = \lim v_n,$$

where in the first equality we assume that f is nonvanishing. This shows that f^{-1} and $|f|^p$ are measurable.

In particular, since products and reciprocals of measurable functions are measurable, whenever the reciprocal is well-defined, it follows that quotients of measurable functions are measurable, whenever the denominator is nonvanishing.

11.3 Littlewood's Third Principle

We finally come to the third of Littlewood's principles, which is

Every convergent sequence of [real-valued] measurable functions is nearly uniformly convergent, or, more precisely, in the words of Lebesgue who in 1903 stated this principle as.

Every convergent series of measurable functions is uniformly convergent when certain sets of measure ε are neglected, where ε can be as small as desired.

Lebesgue here is introducing the idea which is nowadays called "convergence almost uniformly." A sequence $\{f_n\}$ of measurable functions is said to converge almost uniformly (or "a.u." for short) to a measurable function f, denoted by

$$f_n \to f \text{ a.u.},$$

if for each $\varepsilon > 0$, there exists a measurable set A such that $\mu(A) < \varepsilon$ and $f_n \to f$ uniformly on $A^c = X \setminus A$. As a quick review, recall that $f_n \to f$ uniformly on A^c means that given any $\delta > 0$,

$$|f_n(x) - f(x)| < \delta, \quad \text{for all } x \in A^c \text{ and n sufficiently large.}$$

Note that $f_n(x)$ and f(x) are necessarily real-valued (cannot take on $\pm\infty$) on A^c. Therefore, Lebesgue is saying that

Every convergent sequence of real-valued measurable functions is almost uniformly convergent.

The following theorem, although stated by Lebesgue in 1903, is named after Dimitri Fedorovich Egorov (1869-1931) who proved it in 1911[34].

Theorem 4: Egorov's Theorem

On a finite measure space, a.e. convergence implies a.u. convergence for real-valued measurable functions. That is, any sequence of real-valued measurable functions that converges a.e. to a real-valued measurable function converges a.u. to that function.

Proof: Let f, $f_1, f_2, f_3,\ldots$ be real-valued measurable functions on a measure space X with $\mu(X) < \infty$, and assume that $f = \lim f_n$ a.e, which means there is a measurable set $A \subseteq X$ with $\mu(X \setminus A) = 0$ and $f(x) = \lim\limits_{n\to\infty} f_n(x)$ for all $x \in A$. We need to show that $f_n \to f$ a.u.

Step 1: Given $\varepsilon, \eta > 0$ we shall prove that there is a measurable set $B \subseteq X$ and an $N \in \mathbb{N}$ such that

$$\mu(B) < \eta \quad \text{and} \quad \text{for } x \in B^c, \quad |f(x) - f_n(x)| < \varepsilon \text{ for all } n > N. \tag{3.3}$$

Indeed, for each $m \in \mathbb{N}$, put

$$B_m := \bigcup_{n\geq m} \{x \in X; |f(x) - f_n(x)| \geq \varepsilon\}$$

Notice that each B_m is measurable and $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$. Also, since for all $x \in A$, we have $f_n(x) \to f(x)$ as $n \to \infty$, it follows that if $x \in A$, then $|f(x) - f_n(x)| < \varepsilon$ for all n sufficiently large. Thus, there is an m such that $x \notin B_m$, and so, $x \in A \Rightarrow x \notin B_m$ for some m. Taking contrapositives we see that $x \in B_m$ for all $m \Rightarrow x \notin A$, which is to say,

$$\bigcap_{m=1}^{\infty} B_m \subseteq X \setminus A.$$

Thus, $\mu\left(\bigcap_{m=1}^{\infty} B_m\right) \leq \mu(X \setminus A) = 0$ and therefore, since X is a finite measure space, by continuity of measures (from above), we have

Notes

$$\lim_{m\to\infty} \mu(B_m) = 0.$$

Choose N such that $\mu(B_N) < \eta$ and let $B = B_N$. Then by definition of B_N, one can check that holds. This concludes Step 1.

Step 2: We now finish the proof. Let $\varepsilon > 0$. Then by Step 1, for each $k \in \mathbb{N}$ we can find a measurable set $A_k \subseteq X$ and a corresponding natural number $N_k \in \mathbb{N}$ such that

$$\mu(A_k) < \frac{\varepsilon}{2^k} \quad \text{and} \quad \text{for } x \in A_k^c, \quad |f(x) - f_n(x)| < \frac{1}{k} \text{ for all } n > N_k.$$

Now put $A = \bigcup_{k=1}^{\infty} A_k$. Then $\mu(A) < \varepsilon$ and we claim that $f_n \to f$ uniformly on A^c. Indeed, let $\delta > 0$ and choose $k \in \mathbb{N}$ such that $1/k < \eta$. Then

$$x \in A^c = \bigcap_{j=1}^{\infty} A_j^c \Rightarrow \quad x \in A_k^c$$

$$\Rightarrow \quad |f(x) - f_n(x)| < \frac{1}{k} \text{ for all } n > N_k$$

$$\Rightarrow \quad |f(x) - f_n(x)| < \eta \text{ for all } n > N_k.$$

Thus, $f_n \to f$ a.u.

We remark that one cannot drop the finiteness assumption.

Self Assessment

Fill in the blanks:

1. Let $\{f_n\}$ be a sequence of .... on a measure space $(X, \mathscr{S}, \mu)$. We define the functions sup f_n, inf f_n, lim sup f_n, and lim inf f_n, by applying these limit operations pointwise to the sequence of extended real numbers $\{f_n(x)\}$ at each point $x \in X$.
2. If the sequence $\{f_n\}$ is ...., that is, either non-decreasing or non-increasing, then lim f_n is everywhere defined and it is measurable.
3. If f and g are measurable, then ...., and $|f|^p$ where $p > 0$, are also measurable, whenever each expression is defined.
4. Every convergent series of measurable functions is .... when certain sets of measure ε are neglected, where ε can be as small as desired.

11.4 Summary

- For a sequence $\{a_n\}$ of extended real numbers, we know, in general, that lim a_n does not exist; for example, it can oscillate. Assuming that the sequence continues the way it looks like it does, it is clear that although limit lim a_n does not exist, the sequence does have an "upper" limiting value, given by the limit of the odd-indexed a_n's and a "lower" limiting value, given by the limit of the even-indexed a_n's. Now how do we find the "upper" (also called "supremum") and "lower" (also called "infimum") limits of $\{a_n\}$? It turns out there is a very simple way to do so, as we now explain.
- Let $\{f_n\}$ be a sequence of extended-real valued functions on a measure space $(X, \mathscr{S}, \mu)$. We define the functions sup f_n, inf f_n, lim sup f_n, and lim inf f_n, by applying these limit operations pointwise to the sequence of extended real numbers $\{f_n(x)\}$ at each point $x \in X$. For example,

$$\limsup f_n : X \to \overline{\mathbb{R}}$$

is the function defined by

$(\limsup f_n)(x) := \limsup(f_n(x))$ at each $x \in X$.

- If $\{f_n\}$ is a sequence of measurable functions, then the functions

  $$\sup f_n, \quad \inf f_n, \quad \limsup f_n, \quad \text{and} \quad \liminf f_n$$

  are all measurable. If the limit $\lim_{n\to\infty} f_n(x)$ exists at each $x \in X$, then the limit function $\lim f_n$ is measurable. For instance, if the sequence $\{f_n\}$ is monotone, that is, either non-decreasing or non-increasing, then $\lim f_n$ is everywhere defined and it is measurable.
- A function is measurable if and only if it is the limit of simple functions. Moreover, if the function is non-negative, the simple functions can be taken to be a non-decreasing sequence of non-negative simple functions.

11.5 Keywords

Limits Preserve Measurability: If $\{f_n\}$ is a sequence of measurable functions, then the functions

$$\sup f_n, \quad \inf f_n, \quad \limsup f_n, \quad \text{and} \quad \liminf f_n$$

are all measurable.

Characterization of Measurability: A function is measurable if and only if it is the limit of simple functions. Moreover, if the function is nonnegative, the simple functions can be taken to be a non-decreasing sequence of nonnegative simple functions.

Uniformly Convergent: Every convergent sequence of real-valued measurable functions is almost uniformly convergent.

Egorov's Theorem: On a finite measure space, a.e. convergence implies a.u. convergence for real-valued measurable functions.

11.6 Review Questions

1. Let $A_1, A_2, \ldots$ be measurable sets and put

   $$\limsup A_n := \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \quad \text{and} \quad \liminf A_n := \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k.$$

   Let $\bar{f}$ and $\underline{f}$ be the characteristic functions of limsup A_n and liminf A_n, respectively, and for each n, let f_n be the characteristic function of A_n. Prove that

   $$\bar{f} = \limsup f_n \quad \text{and} \quad \underline{f} = \liminf f_n.$$

   (i) First prove the theorem for simple functions. Suggestion: Let f be a simple function and write $f = \sum_{k=1}^{N} a_k \chi_{A_k}$ where $X = \bigcup_{k=1}^{N} A_k$, the a_k's are real numbers, and the A_k's are pairwise disjoint measurable sets. Given $\varepsilon > 0$, there is a closed set $C_k \subseteq \mathbb{R}^n$ with $\mathfrak{m}(A_k \setminus C_k) < \varepsilon/N$ (why?). Let $C = \bigcup_{k=1}^{N} C_k$.

   (ii) We now prove Luzin's theorem for non-negative f. For nonnegative f we know that $f = \lim f_k$ where each f_k, $k \in \mathbb{N}$, is a simple function. By (i), given $\varepsilon > 0$ there is a closed set C_k such that $\mathfrak{m}(X \setminus C_k) < \varepsilon/2^k$ and f_k is continuous on C_k.

   Let $K_1 = \bigcap_{k=1}^{\infty} C_k$. Show that $\mathfrak{m}(X \setminus K_1) < \varepsilon$. Use Egorov's theorem to show that there exists a set $K_2 \subseteq K_1$ with $\mathfrak{m}(K_1 \setminus K_2) < \varepsilon$ and $f_k \to f$ uniformly on K_2. Conclude that f is continuous on K_2.

(iii) Now find a closed set $C \subseteq K_2$ such that $m(K_2 \setminus C) < \varepsilon$. Show that $m(X \setminus C) < 3\varepsilon$ and the restriction of f to C is a continuous function.

(iv) Finally, prove Luzin's theorem dropping the assumption that f is non-negative.

2. A sequence $\{f_n\}$ of real-valued measurable functions is said to be convergent in measure if there is a measurable function f such that for each $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu(\{x; |f_n(x) - f(x)| \geq \varepsilon\}) = 0.$$

(Does this remind you of the weak law of large numbers?) Prove that if $\{f_n\}$ converges in measure to a measurable function f, then f is a.e. real-valued, which means $\{x; f(x) = \pm\infty\}$ is measurable with measure zero. If $\{f_n\}$ converges to two functions f and g in measure, prove that f = g a.e. ***Suggestion:*** To see that f = g a.e., prove and then use the "set-theoretic triangle inequality": For any real-valued measurable functions f, g, h, we have

$$\{x;\ |f(x) - g(x)| \geq \varepsilon\} \subseteq \left\{x;\ |f(x) - h(x)| \geq \frac{\varepsilon}{2}\right\} \cup \left\{x;\ |h(x) - g(x)| \geq \frac{\varepsilon}{2}\right\}.$$

3. Here are some relationships between convergence a.e., a.u., and in measure.

(a) (a.u. $\Rightarrow$ in measure) Prove that if $f_n \to f$ a.u., then $f_n \to f$ in measure.

(b) (a.e. $\Rightarrow$ in measure) From Egorov's theorem prove that if X has finite measure, then any sequence $\{f_n\}$ of real-valued measurable functions that converges a.e. to a real-valued measurable function f also converges to f in measure.

(c) (In measure $\not\Rightarrow$ a.u. nor a.e.) Let X = [0,1] with Lebesgue measure. Given $n \in \mathbb{N}$, write $n = 2^k + i$ where k = 0, 1, 2,... and $0 \leq i < 2^k$, and let f_n be the characteristic function of the interval $\left[\frac{i}{2^k}, \frac{i+1}{2^k}\right]$. Draw pictures of $f_1, f_2, f_3, \ldots, f_7$. Show that $f_n \to 0$ in measure, but $\lim_{n\to\infty} f_n(x)$ does not exist for any $x \in [0, 1]$. Conclude that $\{f_n\}$ does not converge to f a.u. nor a.e.

4. A sequence $\{f_n\}$ of real-valued, measurable functions is said to be Cauchy in measure if for any $\varepsilon > 0$,

$$\mu\big(\{x;\ |f_n(x) - f_m(x)| \geq \varepsilon\}\big) \to 0, \quad \text{as } n, m \to \infty.$$

Prove that if $f_n \to f$ in measure, then $\{f_n\}$ is Cauchy in measure.

5. In this problem we prove that if a sequence $\{f_n\}$ of real-valued measurable functions is Cauchy in measure, then there is a subsequence $\{f_{n_k}\}$ and a real-valued measurable function f such that $f_{n_k} \to f$ a.u. Proceed as follows.

(a) Show that there is an increasing sequence $n_1 < n_2 < \cdots$ such that

$$\mu\big(\{x; |f_n(x) - f_m(x)| \geq \varepsilon\}\big) < \frac{1}{2^k}, \quad \text{for all } n, m \geq n_k.$$

(b) Let

$$A_m = \bigcup_{k=m}^{\infty} \left\{x; \left|f_{n_k}(x) - f_{n_{k+1}}(x)\right| \geq \frac{1}{2^k}\right\}.$$

Show that $\{f_{n_k}\}$ is a Cauchy sequence of bounded functions on the set A_m^c. Deduce that there is a real-valued measurable function f on $A := \bigcup_{m=1}^{\infty} A_m^c$ such that $\{f_{n_k}\}$ converges uniformly to f on each A_m^c.

(c) Define f to be zero on A^c. Show that $f_n \to f$ a.u.

Answers: Self Assessment

1. extended-real valued functions
2. monotone
3. f + g, f · g, 1/f
4. uniformly convergent

11.7 Further Readings

Books

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

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

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

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

S.C. Malik : Mathematical Analysis.

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

Unit 12: The Lebesgue Integral of Bounded Functions

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the Lebesgue integral of bounded functions over a set of finite measure
- Explain properties of the Lebesgue integral of bounded functions over a set of finite measure
- Describe bounded convergence theorem

Introduction

After getting basic knowledge of the Lebesgue measure theory, we now proceed to establish the Lebesgue integration theory.

In this unit, unless otherwise stated, all sets considered will be assumed to be measurable.

We begin with simple functions.

12.1 Simple Functions Vanishing Outside a Set of Finite Measure

Recall that the characteristic function χ_A for any set A is defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

A function $\varphi : E \to \mathbb{R}$ is said to be simple if there exists $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and $E_1, E_2, \ldots, E_n \subseteq E$ such that $\varphi = \sum_{i=1}^{n} a_i \chi_{E_i}$. Note that here the E_i's are implicitly assumed to be measurable, so a simple function shall always be measurable. We have another characterization of simple functions:

Proposition: A function $\varphi : E \to \mathbb{R}$ is simple if and only if it takes only finitely many distinct values $a_1, a_2, \ldots a_n$ and $\varphi^{-1}\{a_i\}$ is a measurable set for all i = 1, 2,....., n.

Proof: With the above proposition we see that every simple function φ can be written uniquely in the form

$$\varphi = \sum_{i=1}^{n} a_i \chi_{E_i}$$

where the a_i's are all non-zero and distinct, and the E_i's are disjoint. (Simply take E_i's $\varphi^{-1}\{a_i\}$ for i = 1, 2,...., n where $a_1, a_2, ..., a_n$ are all the distinct values of φ.) We say this is the canonical representation of φ.

We adopt the following notation:

Notation: A function $f : E \to \mathbb{R}$ is said to vanish outside a set of finite measure if there exists a set A with $m(A) < \infty$ such that f vanishes outside A, i.e.

$$f = 0 \quad \text{on} \quad E \backslash A$$

or equivalently f(x) = 0 for all $x \in E\backslash A$. We denote the set of all simple functions defined on E which vanish outside a set of finite measure by $S_0(E)$. Note that it forms a vector space.

We are now ready for the definition of the Lebesgue integral of such functions.

Definition: For any $\varphi \in S_0(E)$ and any $A \subseteq E$, we define the Lebesgue integral of φ over A by

$$\int_A \varphi = \sum_{i=1}^{n} a_i m(E_i \cap A)$$

where $\varphi = \sum_{i=1}^{n} a_i \chi_{E_i}$ is the canonical representation of φ. (From now on we shall adopt the convention that $0 \cdot \infty = 0$. We need this convention here because it may happen that one a_i is 0 while the corresponding $E_i C\backslash A$ has infinite measure. Also note that here A is implicitly assumed to be measurable so $m(E_i \text{ n } A)$ makes sense. We shall never integrate over non-measurable sets.)

It follows readily from the above definition that

$$\int_A \varphi = \int_A \varphi \chi_A$$

for any $\varphi \in S_0(E)$ and for any $A \subseteq E$.

We now establish some major properties of this integral (with monotonicity and linearity being probably the most important ones). We begin with the following lemma.

Lemma: Suppose $\varphi = \sum_{i=1}^{n} a_i \chi_{Ei} \in S_0(E)$ where the E_i's are disjoint. Then for any $A \subseteq E$,

$$\int_A \varphi = \sum_{i=1}^{n} a_i m(E_i \cap A)$$

holds even if the a_i's are not necessarily distinct.

Proof: If $\varphi = \sum_{j=1}^{n} b_i \chi_{Bj}$ is the canonical representation of φ, we have

1. $$B_j = \bigcup_{\{i : a_i = b_j\}}^{m} E_i$$

   for j = 1, 2,..., m and

2. $$\{1, 2, ..., n\} = \bigcup_{j=1}^{m} \{i : a_i = b_j\},$$

   where both unions are disjoint unions. Hence for any $A \subseteq E$, we have

$$\int_A \varphi = \sum_{j=1}^{m} b_j m(B_j \cap A) \qquad \text{(by definition of the integral)}$$

$$= \sum_{j=1}^{m} b_j m\left(\bigcup_{\{i:a_i=b_j\}} E_i \cap A\right) \qquad \text{(by (1))}$$

$$= \sum_{j=1}^{m} b_j \sum_{\{i:a_i=b_j\}} m(E_i \cap A) \qquad \text{(by finite additivity of m)}$$

$$= \sum_{j=1}^{m} b_j \sum_{\{i:a_i=b_j\}} a_i\, m(E_i \cap A)$$

$$= \sum_{i=1}^{n} a_i m(E_i \cap A) \qquad \text{(by (2))}$$

This complete our proof.

12.2 Properties of the Lebesgue Integral

Proposition: (Properties of the Lebesgue integral)

Suppose $\varphi + \not\varphi \in S_0(E)$. Then for any $A \subseteq E$,

(a) $\int_A (\varphi + \not\varphi) = \int_A \varphi + \int_A \not\varphi$ (Note that $\varphi + \not\varphi \in S_0(E)$ too be the vector space structure

(b) $\int_A \alpha\varphi = \alpha \int_A \varphi$ for all $\alpha \in \Upsilon$. (Note $\alpha\ \varphi \in S_0(E)$ again.)

(c) If $\alpha \leq \not\varphi$ a.e. on A then $\int_A \varphi \leq \int_A \not\varphi$.

(d) If $\varphi = \not\varphi$ a.e. on A then $\int_A \varphi = \int_A \not\varphi$.

(e) If $\varphi \geq 0$ a.e. on A and $\int_A \varphi = 0$, then $\varphi = 0$ a.e. on A.

(f) $\left|\int_A \varphi\right| \leq \int_A |\varphi|$. (Note $|\varphi| \in$ So (E) too.Why?)

Remark: (a) and (b) are known as the linearity property of the integral, while (c) is known as the monotonicity property. Furthermore, Lemma is now seen to hold by the linearity of the integral even without the disjointness assumption on the E_i's.

Proof:

(a) Let $\varphi = \sum_{i=1}^{n} a_i \chi_{Ai}$ and $\not\varphi = \sum_{j=1}^{m} b_j \chi_{Bj}$ be canonical representations of φ and $\not\varphi$ respectively.

Then noting that $\chi_{A_i} = \sum_{j=1}^{m} \chi_{A_i \cap B_j}$ for all i and $\chi_{B_j} = \sum_{i=1}^{n} \chi_{A_i \cap B_j}$

$$\varphi = \sum_{i=1}^{n} a_i \chi_{A_i} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i \chi_{A_i \cap B_j}$$

$$\not\varphi = \sum_{j=1}^{m} b_j \chi_{B_j} = \sum_{i=1}^{n}\sum_{j=1}^{m} b_j \chi_{A_i \cap B_j}$$

Consequently

$$\varphi + \not\varphi = \sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j)\chi_{A_i \cap B_j}.$$

But the $A_i \cap B_j$'s are disjoint. So by Lemma we have

$$\int_A \varphi = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i m(A_i \cap B_j \cap A)$$

$$\int_A \psi = \sum_{i=1}^{n}\sum_{j=1}^{m} b_i m(A_i \cap B_j \cap A)$$

and

$$\int_A (\varphi + \psi) = \sum_{i=1}^{n}\sum_{j=1}^{m} (a_i + b_j) m(A_i \cap B_j \cap A).$$

Hence $\int_A (\varphi + \psi) = \int_A \varphi + \psi$

(b) If $\alpha = 0$ the result is trivial; if not, then let $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ be the canonical representation of φ. We see that $\alpha\varphi = \sum_{i=1}^{n} \alpha a_i \chi_{A_i}$ is the canonical representation of $\alpha\varphi$ and hence the result follows.

(c) Since $\int_A \varphi - \int_A \psi = \int_A (\varphi - \psi)$ by linearity, it suffices to show $\int_A \phi \geq 0$ whenever $\varphi \geq 0$ a.e. on A. This is easy, since if $a_1, a_2, \ldots, a_n$ are the distinct values of ϕ, then

$$\int_A \phi \sum_{\{i:a_i<0\}} a_i m(\phi^{-1}\{a_i\} \cap A) + \sum_{\{i:a_i\geq 0\}} a_i m(\phi^{-1}\{a_i\} \cap A) \geq \sum_{\{i:a_i<0\}} a_i \cdot 0 = 0$$

where the inequality follows from the fact that $m(\phi^{-1}\{a_i \cap A\}) = 0$ for all $a_i < 0$.

(d) This is immediate from (c).

(e) Since it is given that $\varphi \geq 0$ a.e. on A, it suffices to show $m(\{x : \varphi(x) > 0\} \cap A) = 0$.

Suppose not, then there exists $a > 0$ such that $m(\{x : \varphi(x) = a\} \cap A) > 0$ so $\int_A \varphi \geq a \cdot m(\{x : \varphi(x) = a\} \cap A) > 0$. This leads to a contradiction.

(f) This follows directly from monotonicity since $-|\varphi| \leq \varphi \leq |\varphi|$.

12.3 Bounded Measurable Functions Vanishing Outside a Set of Finite Measure

Resembling the construction of the Riemann integral, we define the upper and lower Lebesgue integrals.

Definition: Let $f : E \to \mathbb{R}$ be a bounded function which vanish outside a set of finite measure. For any $A \subseteq C$, we define the upper integral and the lower integral of f on A by

$$\overline{\int_A} f = \inf\left\{\int_A \psi : f \leq \psi \text{ on } A, \psi \in So(E)\right\}$$

$$\underline{\int_A} f = \sup\left\{\int_A \varphi : f \leq \varphi \text{ on } A, \varphi \in So(E)\right\}$$

If the two values agree we denote the common value by $\int_A f$. (Again the set A is implicitly assumed to be measurable so that $\int_A \psi$ and $\int_A \varphi$ make sense.)

Note that both the infimum and the supremum in the definitions of the upper and lower integrals exist because f is bounded and vanishes outside a set of finite measure. It is evident that for the

Notes

functions have their upper and lower integrals both equal to their integral as defined in the last section. In other words, if $\varphi \in S_0(E)$ then $\overline{\int}_A \varphi = \underline{\int}_A \varphi = \int_A \varphi$, where the last integral is as defined in the last section. It is also clear that $-\infty < \underline{\int}_A f = \overline{\int}_A f < \infty$ whenever they are defined; we investigate when $\underline{\int}_A f = \overline{\int}_A f$.

Proposition: Let f be as in the above definition. Then $\underline{\int}_A f = \overline{\int}_A f$ for all $A \subseteq E$ if and only if f is measurable.

Proof: ($\Leftarrow$) Let f be a bounded measurable function defined on E which vanishes outside F with $F \subseteq E$ and $m(F) < \infty$. Then for each positive integer n there are simple functions φ_n, $\psi_n \in S_0(E)$ vanishing outside F such that $\varphi_n \leq f \leq \psi_n$ and $0 \leq \psi_n - \varphi_n \leq 1/n$ E on E (Why?). Hence for any $A \subseteq E$, we have

$$0 \leq \overline{\int}_A f - \underline{\int}_A f \quad \text{(subtraction makes sense since both integrals are finite)}$$

$$\leq \overline{\int}_A \psi_n - \underline{\int}_A \varphi_n \quad \text{(definition of } \overline{\int}_A f \text{ and } \underline{\int}_A f\text{)}$$

$$= \int_A (\psi_n - \varphi_n)$$

$$= \int_{A\cap F} (\psi_n - \varphi_n) \quad (\varphi_n = \psi_n = 0 \text{ outside F})$$

$$\leq \int_F (\psi_n - \varphi_n) \quad (\psi_n - \varphi_n \geq 0 \text{ on F and } A \cap F \subseteq F)$$

$$\leq m(F)/n \quad (1,\ \psi_n - \varphi_n\ 1/n \text{ on F})$$

for all n. Letting $n \to \infty$ we have $\underline{\int}_A f = \overline{\int}_A f$. ($m(F) < \infty$)

($\Leftarrow$) Suppose $\overline{\int}_A f = \underline{\int}_A f$ for any $A \subseteq E$. Then $\overline{\int}_E f = \underline{\int}_E f$. Denote the common value by L. Then for all positive integers n there exists φ_n, ψ_n $S_0(E)$ such that $\varphi_n \leq f \leq \psi_n$ on E and $L - 1/n \leq \int_E \varphi_n \leq \int_E \psi_n \leq L + 1/n$. Let $\varphi = \sup_n \varphi_n$ and $\psi = \inf_n \psi_n$. We shall show $\varphi = \psi$ a.e. on E. (Then the desired conclusion follows since then $\varphi < f < \psi$ on E implies that $\varphi = f = \psi$ a.e. on E and hence f is measurable.) To show that $\varphi = \psi$ a.e. on E, let $\Delta = \{x \in E : \varphi(x) \neq \psi(x)\}$ and $\Delta_i = \{x \in E : \psi(x) - \varphi(x) > 1/i\}$. Then $\Delta = \bigcup_{i=1}^{\infty} \Delta_i$. We wish to show $m(\Delta) = 0$, which will be true if we can show $m(\Delta_i) = 0$ for all i. Now for any i and n, since $\psi_n - \varphi_n \geq \psi - \varphi > 1/i$ on Δ_i, we have

$$\frac{1}{i} m(\Delta_i) = \int_{\Delta_i} \frac{1}{i} \quad \text{(by definition of the integral)}$$

$$\leq \int_{\Delta_i} (\psi_n - \varphi_n)$$

$$\leq \int_E (\psi_n - \varphi_n) \quad (\psi_n - \varphi_n \geq 0 \text{ on E and } \Delta i \subseteq E)$$

$$\leq \int_E \psi_n - \int_E \varphi_n$$

$$\leq 2/n$$

Letting $n \to \infty$ we have $m(\Delta_i) = 0$ for all i, completing our proof.

Notation: We shall denote the set of all (real-valued) bounded measurable functions defined on E which vanishes outside a set of finite measure by $B_0(E)$.

So from now on for $f \in B_0(E)$, we have

$$\int_A f = \inf\left\{\int_A \psi : f \le \psi \in So\,(E)\right\} = \sup\left\{\int_A \varphi : f \ge \varphi \in So\,(E)\right\}$$

for any $A \subseteq E$.

Note also that $B_0(E)$ is a vector lattice, by which we mean it is a vector space partially ordered by $\le$ (such that $f \le g$ if and only if $f(x) \le g(x)$ for all $x \in E$) and every two elements of it (say $f, g \in B_0(E)$) have a least upper bound in it (namely $f \vee g \in B_0(E)$). (Why is it a least upper bound?)

We have the following nice proposition concerning the relationship between the Riemann and the Lebsegue integrals.

Proposition: If $f : [a, b] \to \mathbb{R}$ is Riemann integrable on the closed and bounded interval [a, b], then $f \in B_0([a, b])$ and

(3) $\qquad (\mathcal{R})\ \int_a^b f = (\mathcal{L})\int_{[a,b]} f,$

where the $(\mathcal{R})$ and $(\mathcal{L})$ represents Riemann integral and Lebesgue integral respectively.

Proof: Since step functions defined on closed and bounded interval [a, b] are simple and have the same Lebesgue and Riemann integral over [a, b] (why?), we see from the definitions

$$(\mathcal{R}) = \underline{\int}_a^b f = \sup\left\{\int_a^b \varphi : f \ge \text{step on } [a, b]\right\}$$

$$(\mathcal{L}) = \underline{\int}_{[a,b]} f = \sup\left\{\int_{[a,b]} \varphi : f \ge \varphi \text{ simple on } [a, b]\right\}$$

$$(\mathcal{L}) = \overline{\int}_{[a,b]} f = \inf\left\{\int_{[a,b]} \psi : f \le \psi \text{ simple on } [a, b]\right\}$$

$$(\mathcal{R}) = \overline{\int}_a^b f = \inf\left\{\int_a^b \psi : f \le \psi \text{ step on } [a, b]\right\}$$

that

(4) $\qquad (\mathcal{R}) = \underline{\int}_a^b f \le (\mathcal{L})\underline{\int}_{[a,b]} f \le (\mathcal{L})\overline{\int}_{[a,b]} f \le (\mathcal{R})\overline{\int}_a^b f$

whenever the four quantities exist. Now if f is Riemann integrable over [a, b], then f is bounded on [a,b]. Since [a,b] is of finite measure, we see that all four quantities in (4) exist. In that case $(\mathcal{R})\underline{\int}_a^b f = (\mathcal{R})\overline{\int}_a^b f$ as well so all four quantities in (4) are equal, which implies that f is measurable (so $f \in B_0([a, b])$) and (3) holds.

Proposition: Properties of the Lebesgue integral

Suppose $f, g \in B_0(E)$. Then $f + g$, αf, $|f| \in B_0(E)$, and for any $A \subseteq E$, we have

(a) $\quad \int_A (f+g) = \int_A f + \int_A g$

(b) $\quad \int_A \alpha f = \alpha \int_A f$ for all $\alpha \in \mathbb{R}$.

(c) $\quad \int_A f = \int_E f \chi_A$

(d) $\quad$ If $B \subseteq A$ then $\int_B f + \int_{A \setminus B} f$.

(e) $\quad$ If $B \subseteq A$ and $\int_B f \le \int_A f$

Notes

(f) If $f \le g$ a.e. on A then $\int_A f \le \int_A g$.

(g) If $f = g$ a.e. on A then $\int_A f = \int_A g$.

(h) If $f \ge 0$ a.e. on A and $\int_A f = 0$, then $f = 0$ a.e. on A.

(i) $\left|\int_A f\right| \le \int_A |f|$.

Proof: We prove only (h); the others are easy and left as an exercise.

(h) For each positive integer n let $A_n = \{x \in A : f(x) \ge 1/n\}$. Then

$$0 = \int_A f \ge \int_{An} f \quad \text{(by (e))}$$

$$\ge \int_{An} \frac{1}{n} \quad \text{(by (f))}$$

$$= \frac{1}{n} m(A_n) \quad \text{(by (by definition of the integral)}$$

$$\ge 0$$

so $m(A_n) = 0$. Since this holds for all n, we see from $f^{-1}(0, \infty) \cap A = \bigcup_{n=1}^{\infty} A_n$ that $0 \le m(f^{-1}(0, \infty) \cap A) \le \sum_{n=1}^{\infty} m(A_n) = 0$. So $m(f^{-1}(0, \infty) \cap A) = 0$. Together with $f \ge 0$ a.e. on A.

Theorem: Bounded Convergence Theorem

Suppose $m(E) < \infty$, and $\{f_n\}$ is a sequence of measurable functions defined and uniformly bounded on E by some constant $M > 0$, i.e.

$$|f_n| \le M \quad \text{for all n on E.}$$

If $\{f_n\}$ converges to a function f (pointwisely) a.e. on E, then f is also bounded measurable on E, $\lim_{n\to\infty} \int_E f_n$ exists (in $\mathbb{R}$) and is given by

(5) $$\lim_{n\to\infty} \int_E f_n = \int_E f$$

Proof: Under the given assumptions it is clear that f, being the pointwise limit of $\{f_n\}$ a.e. on E, is bounded (by M) and measurable on E. We wish to show $\lim_{n\to\infty} \int_E f_n$ exists and (5) holds. The result is trivial if $m(E) = 0$. So assume $m(E) > 0$ and let $\varepsilon > 0$ be given. Then for each natural number i let

$$E_i = \{x \in, E : |f_j(x) - f(x)| \ge \varepsilon/2m(E) \text{ for some } j \ge i\}.$$

Then $\{E_i\}$ is a decreasing sequence of sets with $m(E_1) \le m(E) < \infty$. So

$$m(E_i) \downarrow m\left(\bigcap_{i=1}^{\infty} E_i\right) = 0,$$

the last equality follows from the fact that

$$m\left(\bigcap_{i=1}^{\infty} E_i\right) \le m(\{x \in E : f_n(x) \nrightarrow f(x)\}) = 0$$

Choose N large enough such that $m(E_N) < \varepsilon/4M$ and let $A = E_N$. Then $|f_n - f| < \varepsilon/2m(E)$ everywhere on $E\backslash A$ for all $n \ge N$, and hence whenever $n \ge N$ we have

$$\left|\int_E f_n - \int_E f\right| \le \int_E \left|\int_n - f\right| \quad \text{(by linearity and (i))}$$

Notes

$$= \int_{E\setminus A}|f_n - f| + \int_A |f_n - f| \text{ (by (e))}$$

$$\le \int_{E\setminus A}\frac{\varepsilon}{2m(E)} + \int_{E\setminus A} 2M \text{ (by our choice of N and that } n \ge N)$$

$$= \frac{\varepsilon m(E\setminus A)}{2m(E)} + 2Mm(A)$$

$$\le \frac{\varepsilon}{2} + 2M\frac{\varepsilon}{4M}$$

$$= \varepsilon.$$

Hence $\lim_{n\to\infty}\int_E f_n$ exists (in $\mathbb{R}$) and (5) holds.

(Alternatively when $\varepsilon > 0$ is given, by Littlewood's 3rd Principle we can choose a subset A of E with $m(A) < \varepsilon/4M$ such that $\{f_n\}$ converges uniformly to f on E\A. Then choose N large enough such that $|f_n - f| < \varepsilon/2m(E)$ everywhere on E\A for all $n \ge N$, we see that whenever $n \ge N$, we have (as in the above)

$$\left|\int_E f_n - \int_E f\right| < \varepsilon.$$

Hence $\lim_{n\to\infty}\int_E f_n$ exists (in $\mathbb{R}$) and (5) holds.)

Notes The first argument is just an adaptation of the proof of Littlewood's 3rd Principal to the present situation.

Self Assessment

Fill in the blanks:

1. A function $\varphi : E \to \mathbb{R}$ is simple if and only if it takes only finitely many distinct values a_1, a_2,.... a_n and $\varphi^{-1}\{a_i\}$ is a .... for all i = 1, 2,....., n.
2. A function $f : E \to \mathbb{R}$ is said to vanish outside a set of .... if there exists a set A with $m(A) < \infty$ such that f vanishes outside A, i.e.

   $f = 0$ on $E\setminus A$
3. Let f be as in the above definition. Then $\underline{\int}_A f = \overline{\int}_A f$ for all $A \subseteq E$ if and only if f is ....
4. If $f : [a, b] \to \mathbb{R}$ is .... on the closed and bounded interval [a, b], then $f \in B_0([a, b])$ and $(\mathcal{R})\int_a^b f = (\mathcal{L})\int_{[a,b]} f$, where the $(\mathcal{R})$ and $(\mathcal{L})$ represents Riemann integral and Lebesgue integral respectively.
5. Suppose $m(E) < \infty$, and $\{f_n\}$ is a sequence of measurable functions defined and uniformly bounded on E by some constant $M > 0$, i.e. .... for all n on E.

12.4 Summary

- Recall that the characteristic function χ_A for any set A is defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

- A function $\varphi : E \to \mathbb{R}$ is said to be simple if there exists $a_1, a_2, ..., a_n \in \mathbb{R}$ and $E_1, E_2, ..., E_n \subseteq E$ such that $\varphi = \sum_{i=1}^{n} a_i \chi_{E_i}$. Note that here the E_i's are implicitly assumed to be measurable, so a simple function shall always be measurable. We have another characterization of simple functions:
- A function $\varphi : E \to \mathbb{R}$ is simple if and only if it takes only finitely many distinct values a_1, a_2, ... a_n and $\varphi^{-1}\{a_i\}$ is a measurable set for all i = 1, 2,....., n.

  (a) $\int_A (\varphi + \psi) = \int_A \varphi + \int_A \psi$ (Note that $\varphi + \psi \in S_0(E)$ too by the vector space structure

  (b) $\int_A \alpha\varphi = \alpha \int_A \varphi$ for all $\alpha \in \Upsilon$. (Note $\alpha\ \varphi \in S_0(E)$ again.)

  (c) If $\alpha \leq \psi$ a.e. on A then $\int_A \varphi \leq \int_A \psi$.

  (d) If $\varphi = \psi$ a.e. on A then $\int_A \varphi = \int_A \psi$.

  (e) If $\varphi \geq 0$ a.e. on A and $\int_A \varphi = 0$, then $\varphi = 0$ a.e. on A.

  (f) $\left|\int_A \varphi\right| \leq \int_A |\varphi|$. (Note $|\varphi| \in$ So (E) too.Why?)

- Bounded Convergence Theorem Suppose $m(E) < \infty$, and $\{f_n\}$ is a sequence of measurable functions defined and uniformly bounded on E by some constant $M > 0$, i.e.

  $|f_n| \leq M$ for all n on E.

12.5 Keyword

Bounded Convergence Theorem: Suppose $m(E) < \infty$, and $\{f_n\}$ is a sequence of measurable functions defined and uniformly bounded on E by some constant $M > 0$, i.e.

$|f_n| \leq M$ for all n on E.

12.6 Review Questions

1. Show that if $A, B \subseteq E$, $A \cap B = \emptyset$ and $\varphi \in S_0(E)$, then $\int_{A\cup B} \varphi = \int_A \varphi + \int_B \varphi$.
2. Show that if $\varphi \in S_0(E)$ vanishes outside F, then $\int_A \varphi = \int_{A\cap F} \varphi$ for any $A \subseteq E$.
3. Show that if $A \subseteq B \subseteq E$ and $0 \leq \varphi \in S_0(E)$, then $\int_A \varphi \leq \int_B \varphi$.
4. Find an example to show that the assumption $m(E) < \infty$ cannot be dropped in the Bounded Convergence Theorem.
5. Prove or disprove the following: Let E be of finite or infinite measure. If $\{f_n\}$ is a sequence of uniformly bounded measurable functions on E which vanishes outside a set of finite measure and converges pointwisely to $f \in B_0(E)$ a.e. on E, then $\lim_{n\to\infty} \int_E f_n = \int_E f$. (Compare with the statement of the Bounded Convergence Theorem.)

Answers: Self Assessment

1.	measurable set	2.	finite measurable
3.	measurable	4.	Riemann integrable
5.	$|f_n| \leq M$		

12.7 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 13: Riemann's and Lebesgue

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss Riemann's and Lebesgue
- Explain the small subsets of R
- Discuss the functions outside small set

Introduction

In last unit you have studied about the Lebesgue integral of bounded functions. In this unit we are going to study about the definition and the difference of Riemann's and Lebesgue.

13.1 Riemann vs. Lebesgue

Measure theory helps us to assign numbers to certain sets and functions to a measurable set we may assign its measure, and to an integrable function we may assign the value of its integral. Lebesgue integration theory is a generalization and completion of Riemann integration theory. In Lebesgue's theory, we can assign numbers to more sets and more functions than what is possible in Riemann's theory. If we are asked to distinguish between Riemann integration theory and Lebesgue integration theory by pointing out an essential feature, the answer is perhaps the following.

Riemann integration theory $\mapsto$ finiteness.

Lebesgue integration theory $\mapsto$ countable infiniteness.

Riemann integration theory is developed through approximations of a finite nature (e.g.: one tries to approximate the area of a bounded subset of $\mathbb{R}^2$ by the sum of the areas of finitely many rectangles), and this theory works well with respect to finite operations– if we can assign numbers to finitely many sets $A_1,\ldots, A_n$ and finitely many functions $f_1,\ldots,f_n$, then we can assign numbers to $A_1\cup\cdots\cup A_n$, $f_1 +\cdots+ f_n$, $\max\{f_1,\ldots, f_n\}$, etc. The disadvantage of Riemann integration theory is that it does not behave well with respect to operations of a countably infinite nature - there may not be any consistent way to assign numbers to $\bigcup_{n=1}^{\infty} A_n$, $\sum_{n=1}^{\infty} f_n$, $\lim_{n \to \infty} f_n$, $\sup\{f_n : n \in \mathbb{N}\}$, etc. even if we can assign numbers to the sets $A_1, A_2, \ldots$, and functions $f_1, f_2 \ldots$. Lebesgue integration theory rectifies this disadvantage to a large extent.

In Riemann integration theory, we proceed by considering a partition of the domain of a function, where as in Lebesgue integration theory, we proceed by considering a partition of the range of the function – this is observed as another difference. Moreover, while Riemann's theory is restricted to the Euclidean space, the ideas involved in Lebesgue's theory are applicable to more general spaces, yielding an abstract measure theory. This abstract measure theory intersects with many branches of mathematics and is very useful. There is even a philosophy that measures are easier to deal with than sets.

13.2 Small Subsets of $\mathbb{R}^d$

It is possible to think about many mathematical notions expressing in some sense the idea that a subset $Y \subset \mathbb{R}^d$ is a small set (or a big set) with respect to $\mathbb{R}^d$. We will discuss this a little as a warm-up. We will also use this opportunity to introduce Lebesgue outer measure.

Suppose you have a certain notion of smallness or bigness for a subset of $\mathbb{R}^d$. Then there are some natural questions. Two sample questions are:

1. If $Y \subset \mathbb{R}^d$ is big, is $\mathbb{R}^d \setminus Y$ small?
2. If $Y_1, Y_2, \ldots \subset \mathbb{R}^d$ are small, is $\bigcup_{n=1}^{\infty} Y_n$ small?

For instance, consider the following two elementary notions. Saying that $Y \subset \mathbb{R}^d$ is unbounded is one way of saying Y is big, and saying that $Y \subset \mathbb{R}^d$ is a finite set is one way of saying Y is small. Note that the complement of an unbounded set can also be unbounded and a countable union of finite sets need not be finite. So here we have negative answers to the above two questions.

Task Find an uncountable collection $\{Y_\alpha : \alpha \in I\}$ of subsets of $\mathbb{R}$ such that Y_α's are pairwise disjoint, and each Y_α is bounded neither above nor below.

To discuss some other notions of smallness, we introduce a few definitions.

Definitions:

(i) We say $Y \subset \mathbb{R}^d$ is a discrete subset of $\mathbb{R}^d$ if for each $y \in Y$, there is an open set $U \subset \mathbb{R}^d$ such that $U \cap Y = \{y\}$. For example, $\{1/n: n \in \mathbb{N}\}$ is a discrete subset of $\mathbb{R}$.

(ii) A subset $Y \subset \mathbb{R}^d$ is nowhere dense in $\mathbb{R}^d$ if $\text{int}[\,\overline{Y}\,] = \emptyset$, or equivalently if for any non-empty open set $U \subset \mathbb{R}^d$, there is a nonempty open set $V \subset U$ such that $V \cap Y = 0$. For example, if $f: \mathbb{R} \to \mathbb{R}$ is a continuous map, then its graph $G(f) := \{(x, f(x)) : x \in \mathbb{R}\}$ is nowhere dense in $\mathbb{R}^2$ ($\because$ G(f) is closed and does not contain any open disc).

(iii) A subset $Y \subset \mathbb{R}^d$ is of first category in $\mathbb{R}^d$ if Y can be written as a countable union of nowhere dense subsets of $\mathbb{R}^d$; otherwise, Y is said to be of second category in $\mathbb{R}^d$. For example, $Y = \mathbb{Q} \times \mathbb{R}$ is of first category in $\mathbb{R}^2$ since Y can be written as the countable union $Y = \bigcup_{r \in \mathbb{Q}} Y_r$, where $Y_r := \{r\} \times \mathbb{R}$ is nowhere dense in $\mathbb{R}^2$.

(iv) (The following definition can be extended by considering ordinal numbers, but we consider only non-negative integers). For $Y \subset \mathbb{R}^d$ and integer $n \geq 0$, define the nth derived set of Y inductively as $Y^{(0)} = Y$, $Y^{(n+1)} = \{\text{limit points of } Y^{(n)} \text{ in } \mathbb{R}^d\}$. We say $Y \subset \mathbb{R}^d$ has derived length n if $Y^{(n)} \neq \emptyset$ and $Y^{(n+1)} \neq \emptyset$; and we say Y has infinite derived length if $Y^{(n)} \neq \emptyset$ for every

Notes

integer $n \geq 0$. For example, $\mathbb{Q}$ has infinite derived length (since $\overline{\mathbb{Q}} = \mathbb{R}$), and $\{(1/m,1/n): m,n \in \mathbb{N}\}$ has derived length 2.

(v) We say $A \subset \mathbb{R}^d$ is a d-box if $A = \pi_{j=1}^{d} I_j$, where I_j's are bounded intervals. The d-dimensional volume of a d-box A is $Vol_d(A) = \pi_{j=1}^{d} |I_j|$. For example, $Vol_3([1, 4) \times [0,1/2] \times (-1,3]) = 6$.

(vi) The d-dimensional Jordan outer content $\mu^*_{j,d}[Y]$ of a bounded subset $Y \subset \mathbb{R}^d$ is defined as $\mu^*_{j,d}[Y] = \{\sum_{n=1}^{k} \inf Vold(An) : k \in \mathbb{N}$, and A_n's are d-boxes with $Y \subset \bigcup_{n=1}^{k} A_n\}$.

(vii) The d-dimensional Lebesgue outer measure $\mu^*_{L,d}[Y]$ of an arbitrary set $Y \subset \mathbb{R}^d$ is defined as $\mu^*_{L,d}[Y] = \inf\{\sum_{n=1}^{\infty} Vol_d(A_n) : A_n$'s are d-boxes with $Y \subset \bigcup_{n=1}^{\infty} A_n\}$.

We have that $\mu^*_{L,d}[Y] \leq \mu^*_{J,d}[Y]$ for any bounded set $Y \subset \mathbb{R}^d$, and $\mu^*_{L,d}[A] = \mu^*_{J,d}[A] = Vol_d(A)$ for any d-box $A \subset \mathbb{R}^d$.

Proof: Any finite union $\bigcup_{n=1}^{k} A_n$ of d-boxes can be extended to an infinite union $\bigcup_{n=1}^{\infty} A_n$ of d-boxes without changing the total volume by taking A_n's to be singletons for $n > k$. This observation yields that $\mu^*_{L,d}[Y] \leq \mu^*_{J,d}[Y]$. It is easy to see $\mu^*_{J,d}[A] = Vol_d(A)$ if A is a d-box. It remains to show $\mu^*_{L,d}[A] \geq Vol_d(A)$ when A is a d-box. First suppose A is closed. Then A is compact by Heine-Borel. Let $\varepsilon > 0$ and let $A_1, A_2, \ldots \subset \mathbb{R}^d$ be d-boxes such that $A \subset \bigcup_{n=1}^{\infty} A_n$ and $\sum_{n=1}^{\infty} Vol_d(A_n) < \mu^*_{L,d}[A] + \varepsilon$. For each $n \in \mathbb{N}$, let B_n be an open d-box with $A_n \subset B_n$ and $Vol_d(B_n) < Vol_d(A_n) + \varepsilon/2^n$. Then $\{B_n: n \in \mathbb{N}\}$ is an open cover for the compact set A. Extracting a finite subcover, we have $Vol_d(A) \leq \sum_{n=1}^{k} Vol_d(B_n) \leq \sum_{n=1}^{\infty} (Vol_d(A_n) + \varepsilon/2^n) < \mu^*_{L,d}[A] + 2\varepsilon$. Thus $\mu^*_{L,d}[A] = Vol_d(A)$ for closed d-boxes. Now if B is an arbitrary d-box and $\varepsilon > 0$, then there is a closed d-box $A \subset B$ with $Vol_d(B) - \varepsilon < Vol_d(A) = \mu^*_{L,d}[A] \leq \mu^*_{L,d}[B]$.

Other basic properties of Lebesgue outer measure and Jordan outer content are given below.

(i) $\mu^*_{L,d}[\emptyset] = 0$.

(ii) [Monotonicity] $\mu^*_{L,d}[X] \leq \mu^*_{L,d}[Y]$ if $X \subset Y \subset \mathbb{R}^d$.

(iii) [Translation-invariance] $\mu^*_{L,d}[Y + x] = \mu^*_{L,d}[Y]$ for every $Y \subset \mathbb{R}^d$ and every $x \in \mathbb{R}^d$.

(iv) [Countable subadditivity] If $Y_1, Y_2, \ldots \subset \mathbb{R}^d$ and $Y = \bigcup_{n=1}^{\infty} Y_n$, then $\mu^*_{L,d}[Y] \leq \sum_{n=1}^{\infty} \mu^*_{L,d}[Y_n]$.

(v) $\mu^*_{L,d}[Y] = 0$ for every countable set $Y \subset \mathbb{R}^d$.

(vi) Forany $Y \subset \mathbb{R}^d$, we have $\mu^*_{L,d}[Y]$

$= \{\sum_{n=1}^{\infty} \inf Vold(An) : A_n$'s are closed d-boxes with $Y \subset \bigcup_{n=1}^{\infty} A_n\}$

$= \{\sum_{n=1}^{\infty} \inf Vold(An) : A_n$'s are open d-boxes with $Y \subset \bigcup_{n=1}^{\infty} A_n\}$.

(vii) For any $Y \subset \mathbb{R}^d$, we have $\mu^*_{L,d}[Y] = \mu^*_{L,d} \inf\{[U] : Y \subset U$ and U is open in $\mathbb{R}^d\}$.

(viii) $\mu^*_{L,d}[\mathbb{R}^d] = \infty$.

(ix) If $X,Y \subset \mathbb{R}^d$ are such that $dist(X,Y) := \inf\{||x - y|| : x \in X, y \in Y\} > 0$, then $\mu^*_{L,d}[X \cup Y] = \mu^*_{L,d}[X] + \mu^*_{L,d}[Y]$.

Proof: (i), (ii) and (iii) are clear. To prove (iv), without loss of generality we may assume $\sum_{n=1}^{\infty} \mu^*_{L,d}[Y_n] < \infty$. Given $\varepsilon > 0$, there exist d-boxes A(n,k) such that $Y_n \subset \bigcup_{k=1}^{\infty} A(n,k)$ and $\sum_{k=1}^{\infty} Vol_d(A(n,k)) < \mu^*_{L,d}[Y_n] + \varepsilon/2^n$. Then $Y \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} A(n,k)$ and we have the estimate $\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} Vol_d(A(n,k)) \leq \sum_{n=1}^{\infty} (\mu^*_{L,d}[Y_n] + \varepsilon/2^n) = (\sum_{k=1}^{\infty} \mu^*_{L,d}[Y_n]) + \varepsilon$.

Now (v) follows from (iv) since singletons have Lebesgue outer measure zero (or we can see it directly by noting that singletons are d-boxes with zero volume). The first part of (vi) is clear since any d-box and its closure have equal volume. To get the second part, note that if $A_1, A_2, \ldots$ are d-boxes and $\varepsilon > 0$, there exist open d-boxes $B_1, B_2, \ldots$ such that $A_n \subset B_n$ and $Vol_d(B_n) < Vol_d(A_n) + \varepsilon/2^n$. We may deduce (vii) using part (vi).

Now we prove (ix). From countable subadditivity, we have $\mu^*_{L,d}[X \cup Y] \leq \mu^*_{L,d}[X] + \mu^*_{L,d}[Y]$. To prove the other inequality, we may assume $\mu^*_{L,d}[X \cup Y] < \infty$. Let $\delta = \text{dist}(X,Y)$. Given $\varepsilon > 0$, find d-boxes $A_1, A_2, \ldots$ such that $X \cup Y \subset \bigcup_{n=1}^{\infty} A_n$ and $\sum_{n=1}^{\infty} \text{Vol}_d(A_n) < \mu^*_{L,d}[X \cup Y] + \varepsilon$. By partitioning the d-boxes into smaller d-boxes and throwing away the unnecessary ones, we may assume that $\text{diam}[A_n] < \delta$ and $(X \cup Y) \cap A_n \neq \emptyset$ for every $n \in \mathbb{N}$. Let $\in = \{n \in \mathbb{N}: X \cap A_n \neq \emptyset\}$ and $\Gamma' = \{n \in \mathbb{N}: Y \cap A_n \neq \emptyset\}$. Then $\mathbb{N} = \in \cup \Gamma'$ is a disjoint union, $X \subset \bigcup_{n \in \Gamma} A_n$, and $Y \subset \bigcup_{n \in \Gamma} A_n$. Hence $\mu^*_{L,d}[X] + \mu^*_{L,d}[Y] \leq \sum_{n \in \Gamma} \text{Vol}_d(A_n) + \sum_{n \in \Gamma} \text{Vol}_d(A_n) = \sum_{n=1}^{\infty} \text{Vol}_d(A_n) < \mu^*_{L,d}[X \cup Y] + \varepsilon$.

(i) $\mu^*_{J,d}[\emptyset] = 0$.

(ii) [Monotonicity] $\mu^*_{J,d}[X] \leq \mu^*_{L,d}[Y]$ if $X \subset Y$ are bounded subsets of $\mathbb{R}^d$.

(iii) [Translation-invariance] $\mu^*_{L,d}[Y + x] = \mu^*_{L,d}[Y]$ for every bounded set $Y \subset \mathbb{R}^d$ and every $x \in \mathbb{R}^d$.

(iv) [Finite subadditivity] If $X, Y \subset \mathbb{R}^d$ are bounded subsets, then $\mu^*_{J,d}[X \cup Y] \leq \mu^*_{J,d}[X] + \mu^*_{J,d}[Y]$.

(v) $\mu^*_{J,d}[Y] = 0$ for every finite set $Y \subset \mathbb{R}^d$.

(vi) For any bounded set $Y \subset \mathbb{R}^d$, we have $\mu^*_{J,d}[Y]$

$= \inf\{\sum_{n=1}^{k} \text{Vol}_d(A_n) : k \in \mathbb{N}, \text{ and } A_n\text{'s are closed d-boxes with } Y \subset \bigcup_{n=1}^{k} A_n\}$

$= \inf\{\sum_{n=1}^{k} \text{Vol}_d(A_n) : k \in N, \text{ and } A_n\text{'s are open d-boxes with } Y \subset \bigcup_{n=1}^{k} A_n\}$

$= \inf\{\sum_{n=1}^{k} \text{Vol}_d(A_n) : k \in N, \text{ and } A_n\text{'s are pairwise disjoint d-boxes with } Y \subset \bigcup_{n=1}^{k} A_n\}$.

(vii) If $X, Y \subset \mathbb{R}^d$ are bounded sets with $\text{dist}(X,Y) := \inf\{\| x - y \| : x \in X, y \in Y\} > 0$, then $\mu^*_{J,d}[X \cup Y] = \mu^*_{J,d}[X] + \mu^*_{J,d}[Y]$.

(viii) For any bounded set $Y \subset \mathbb{R}^d$, we have $\mu^*_{J,d}[\bar{Y}] = \mu^*_{J,d}[Y]$.

Proof: To prove (viii), use the first expression for $\mu^*_{J,d}[Y]$ in (vi) and note that a finite union of closed sets is closed.

Example: Let $Y = \mathbb{Q}^d \cap [0, 1]^d$. Note that $\mu^*_{L,d}[Y] = 0 \neq 1 = \mu^*_{L,d}[\bar{Y}]$. But we have $\mu^*_{J,d}[Y] = \mu^*_{J,d}[\bar{Y}] = 1$. So the Jordan outer content of a bounded countable set need, not be zero. This example also shows that $\mu^*_{L,d}[Y] < \mu^*_{J,d}[Y]$ is possible for a bounded set, and that the Jordan outer content does not satisfy countable subadditivity for bounded sets (since the Jordan outer content of a singleton is zero). If $X = [0, 1]^d \backslash Y$, then $\mu^*_{J,d}[X] = 1$ since $\bar{X} = [0, 1]^d$ and hence $\mu^*_{J,d}[X] + \mu^*_{J,d}[Y] = 2 \neq 1 = \mu^*_{J,d}[X \cup Y]$.

Some ways of saying that $Y \subset \mathbb{R}^d$ is a small set:

(i) Y is a countable set.

(ii) Y is a discrete subset of $\mathbb{R}^d$.

(iii) Y is contained in a vector subspace of $\mathbb{R}^d$ of dimension $\leq d - 1$.

(iv) Y is nowhere dense in $\mathbb{R}^d$.

(v) Y is of first category in $\mathbb{R}^d$.

(vi) Y has finite derived length.

(vii) Y is a bounded set with $\mu^*_{J,d}[Y] = 0$.

(viii) $\mu^*_{L,d}[Y] = 0$.

It is good to investigate various possible implications between pairs of notions given above.

Task If Y is a discrete subset of $\mathbb{R}^d$, then Y is countable. [*Hint:* Let $\mathbb{B} = \{B(x, 1/n): x \in \mathbb{Q}^d, n \in \mathbb{N}\}$. Then, $\mathbb{B}$ is countable and for each $y \in Y$, there is $B \in \mathbb{B}$ such that $B \cap Y = \{y\}$.]

Let $K \subset [0,1]$ be the middle-third Cantor set. Then, K is an uncountable, nowhere dense compact set with $\mu_{J,1}[K] = \mu^*_{L,1}[K] = 0$. Moreover, K has no isolated points.

Proof: We recall the construction of K. Let $K_0 = [0,1]$, $K_1 = [0,1/3] \cup [2/3,1]$, $K_2 = [0,1/9] \cup [2/9,1/3] \cup [2/3, 7/9] \cup [8/9,1]$, and so on. That is, K_n is the disjoint union of 2^n closed subintervals of [0,1], each having length $1/3^n$, and K_{n+1} is obtained from K_n by removing the middle-third open intervals from each of these 2^n closed intervals. The middle-third Cantor set K is defined as $K = \bigcap_{n=0}^{\infty} K_n$. Being the intersection of compact sets, K is compact. Since the maximal length of an interval contained in K_n is $(1/3)^n$, K does not contain any open interval, and hence K is nowhere dense. Also, since $K \subset K_n$, the above description yields $\mu^*_{J,1}[K] \le (2/3)^n$. So $\mu^*_{J,1}[K] = 0$ and hence $\mu^*_{L,1}[K] = 0$ also.

It may be verified that $K = \{\sum_{n=1}^{\infty} x_n/3^n : x_n \in \{0, 2\}\}$. That is, K is precisely the set of those $x \in [0,1]$ whose ternary expansion (i.e., base 3 expansion) $x = 0.x_1x_2 \ldots$ contains only 0's and 2's. Hence K is bijective with $\{0, 2\}^{\mathbb{N}}$ which is uncountable.

We show K has no isolated point. Let $x \in K$ and let U be a neighborhood of x. Choose n large enough so that one of the 2^n closed intervals constituting K_n, say J_n, satisfies $x \in J_n \subset U$. Let $y \in J_n \setminus \{x\}$ be an end point of J_n. This end point is never removed in the later construction, so $y \in K_m$ for every $m \ge n$. Thus $y \in K \cap (U \setminus \{x\})$.

Notes It may be noted that for $x \in K$, the base 3 expansion $x = 0.x_1x_2 \ldots$ is eventually constant iff x is an end point of a removed open interval. This helps to see that K contains points other than the end points of the (countably many) removed open intervals.

The following theorem is relevant while considering big and small sets in a topological sense.

Task If Y is contained in a vector subspace W of $\mathbb{R}^d$ with $\dim(W) \le d - 1$, then Y is a nowhere dense subset of $\mathbb{R}^d$. [*Hint:* W is closed in $\mathbb{R}^d$ ($\because$ fix a basis for W and argue with the coefficients of each basis vector separately) and W does not contain any open ball of $\mathbb{R}^d$.]

Baire Category Theorem: Let (X, ρ) be a complete metric space and let $U_n \subset X$ be open and dense in X for $n \in \mathbb{N}$. Then, $\bigcap_{n=1}^{\infty} U_n$ is also dense in X. In particular, $\bigcap_{n=1}^{\infty} U_n \ne \emptyset$.

Proof: Let $V \subset X$ be a nonempty open set. It suffices to show $V \cap (\bigcap_{n=1}^{\infty} Un) \ne \emptyset$. Since U_1 is open and dense, $U_1 \cap V$ is a nonempty open set. Let B_1 be an open ball in X such that $\overline{B_1} \subset U_1 \cap V$ and $\text{diam}[B_1] < 1$. Since U_2 is open and dense, $B_1 \cap U_2 = \emptyset$. Let $B_2 \subset X$ be an open ball with $\overline{B_2} \subset B_1 \cap U_2$ and $\text{diam}[B_2] < 1/2$. In general, let $B_{n+1} \subset X$ be an open ball with $\overline{B_{n+1}} \subset B_n \cap U_{n+1}$ and $\text{diam}[B_{n+1}] < 1/(n+1)$. If x_n is the center of the ball B_n, then we note that for every $n, m \ge k$ we have $x_n, x_m \in B_k$ and hence $\rho(x_n, x_m) \le \text{diam}[B_k] < 1/k$. So (x_n) is a Cauchy sequence. Since (X, ρ) is complete, there is $x \in X$ such that $(x_n) \to x$. Now, for any n, we have $x_m \in \overline{B_n}$ for $m \ge n$ and hence $x \in \overline{B_n}$. Thus $x \in \bigcap_{n=1}^{\infty} \overline{B_n} \in V \cap (\bigcap_{n=1}^{\infty} U_n)$.

Notes (i) By considering the complements of U_n's in the above, we get the following conclusion: if (X, ρ) is a complete metric space, then X cannot be written as a countable union of nowhere dense (closed) subsets of X. That is, X is of second category in itself. (ii) Since $\mathbb{R}^d$ is a complete metric space with respect to the Euclidean metric, $\mathbb{R}^d$ cannot be written as a countable union of nowhere dense (closed) subsets of $\mathbb{R}^d$. (iii) From a topological point of view, a first category subset is considered as a small set and a dense G_δ subset is considered as a big set because of Baire Category Theorem. However, a set that is topologically big (small) need not be measure theoretically big (small). (iv) The uncountability of the middle-third Cantor set can be proved with the help of Baire Category Theorem also.

We observe in the following the distinction between topological bigness (smallness) and measure theoretical bigness (smallness).

Task For any $Y \subset \mathbb{R}^d$, the set $Y \backslash Y^{(1)}$ is discrete and hence countable. In particular, every uncountable subset of $\mathbb{R}^d$ has a limit point in $\mathbb{R}^d$. [*Hint:* Let $y \in Y \backslash Y^{(1)}$. If $B(y, 1/n) \cap Y$ contains a point other than y for every $n \in \mathbb{N}$, then $y \in Y^{(1)}$, a contradiction.]

(i) For every $\varepsilon > 0$, there is a dense open set $U \subset \mathbb{R}^d$ such that $\mu^*_{L,d}[U] < \varepsilon$.

(ii) There is a dense G_δ subset $Y \subset \mathbb{R}^d$ with $\mu^*_{L,d}[Y] = 0$.

(iii) There is an F_σ set $X \subset \mathbb{R}^d$ of first category with $\mu^*_{L,d}[X] = \infty$ and $\mu^*_{L,d}[\mathbb{R}^d \backslash X] = 0$.

(iv) For every closed d-box A and every $\varepsilon > 0$, there is anywhere dense compact set $K \subset \mathbb{R}^d$ such that $K \subset A$ and $\mu^*_{L,d}[K] > Vol_d(A) - \varepsilon$.

Proof: (i) Write $\mathbb{Q}^d = \{x_1, x_2, \ldots\}$. For each $n \in \mathbb{N}$, let A_n be an open d-box with $x_n \in A_n$ and $Vol_d(A_n) < \varepsilon/2^n$. Put $U = \bigcap_{n=1}^{\infty} A_n$.

(ii) Let $U_n \subset \mathbb{R}^d$ be a dense open subset with $\mu^*_{L,d}[U_n] < 1/n$ and put $Y = \bigcap_{n=1}^{\infty} U_n$.

(iii) Let Y be as in (ii) and take $X = \mathbb{R}^d \backslash Y$.

(iv) LetUbeasin(i)andletK = $A \backslash U$.

The next result shows that the Lebesgue outer measure does not satisfy finite additivity (and hence it does not satisfy countable additivity), even though it satisfies countable subadditivity.

Let $X = \mathbb{Q}^d \cap [0, 1]^d$. Then, there is a subset $Y \subset [0, 1]^d$ satisfying the following:

(i) The translations $Y + x$ are pairwise disjoint for $x \in X$.

(ii) There exist finitely many distinct elements $x_1, \ldots, x_n \in X$ such that $\mu^*_{L,d}[\bigcup_{i=1}^{n}(Y + x_i)] \neq \sum_{i=1}^{n} \mu^*_{L,d}[Y + x_i]$.

Proof: Define an equivalence relation on $[0, 1]^d$ by the condition that $a \sim b$ iff $a - b \in \mathbb{Q}^d$. By the axiom of choice, we can form a set $Y \subset [0, 1]^d$ whose intersection with each equivalence class is a singleton.

(i) We verify that $(Y + r) \cap (Y + s) = \emptyset$ for any two distinct $r, s \in X$. If $(Y + r) \cap (Y + s) = \emptyset$ for $r, s \in X$, then there are $a, b \in Y$ such that $a + r = b + s$. Now we have $a - b = s - r \in \mathbb{Q}^d$, and hence $a \sim b$. So we must have $a = b$ by the definition of Y, and then necessarily $r = s$.

(ii) If $z \in \mathbb{R}^d$, there is $r \in \mathbb{Q}^d$ such that $z - r \in [0, 1]^d$. Then there is $y \in Y$ such that $y \sim z - r$ and so there is $r' \in \mathbb{Q}^d$ such that $y + r' = z - r$ or $z = y + r + r'$. This shows that $\mathbb{R}^d = \bigcup_{r \in \mathbb{Q}^d} (Y + r)$. By [102](iii) and [102](viii), we conclude that $\mu^*_{L,d}[Y] > 0$. Let $\delta = \mu^*_{L,d}[Y]$ and $n \in \mathbb{N}$ be such that $n\delta > 2^d$. Choose distinct elements $x_1, \ldots, x_n \in X$. Then $\sum_{i=1}^{n} \mu^*_{L,d}[Y + x_i] = n\delta > 2^d$ again by translation invariance. On the other hand, $Y + X \in [0,2]^d$ and therefore $\mu^*_{L,d}[\bigcup_{i=1}^{n}(Y+x_i)] < 2^d$.

Note The construction above is due to Vitali, and hence the set Y is called a Vitali set.

13.3 About Functions Behaving Nicely Outside a Small Set

There are a few classical results in Analysis with conclusion of the following form: "… the function has nice behavior outside a small set". We will consider some such results here.

We know that a function that is the pointwise limit of a sequence of continuous functions may not be continuous. For instance, $f : [0, 1] \to \mathbb{R}$ given by $f(1) = 1$ and $f(x) = 0$ for $x < 1$ is the pointwise limit of (f_n), where $f_n : [0, 1] \to \mathbb{R}$ is $f_n(x) = x^n$.

Definition: Let X, Y be metric spaces and let $f: X \to Y$ be a function. Then the oscillation $\omega(f, x)$ of f at a point $x \in X$ is defined as $\omega(f, x) = \lim_{\delta \to 0+} \text{diam}[f(B(x, \delta))]$. Clearly, f is continuous at x iff $\omega(f, x) = 0$.

Task Let X,Y be metric spaces and let $f: X \to Y$ be a function. Then the set $\{x \in X: f$ is continuous at $x\}$ is a G_δ subset of X. [*Hint:* The given set is equal to $\bigcup_{n=1}^{\infty} U_n$, where $U_n = \{x \in X: \omega(f, x) < 1/n\}$, and U_n is open.]

Let (X, ρ_1) be a complete metric space, (Y, ρ_2) be an arbitrary metric space, and let (f_n) be a sequence of continuous functions from X to Y, converging pointwise to a function $f: X \to Y$. Then the set $\{x \in X : f$ is continuous at $x\}$ is a dense G_δ subset of X.

Proof: Let $\varepsilon > 0$ and $D_\varepsilon = \{x \in X : \omega(f, x) > \varepsilon\}$. We know that D_ε is a closed set. We claim that D_ε is nowhere dense in X. Let $U \subset X$ be a nonempty open set. We have to find a nonempty open set $V \subset U$ such that $D_\varepsilon \cap V = \emptyset$.

For $n \in \mathbb{N}$, let $K_n = \{x \in X: \rho_2(f_n(x), fj(x)) \le \varepsilon/8$ for every $j \ge n\}$. Then K_n is a closed set and $X = \bigcup_{n=1}^{\infty} K_n$. The continuity of the distance function ρ_2 implies that $\rho_2(f_n(x), f(x)) \le \varepsilon/8$ for every $x \in K_n$. Let $U_1 \subset X$ be a nonempty open set with $\overline{U_1} \subset U$. Since $(\overline{U_1}, \rho_1)$ is a complete metric space, there is $n \in N$ such that $U_2 := \text{int}[K_n \cap \overline{U_1}] \neq \emptyset$. Let $b \in U_2$ and $V \subset U_2$ be an open set with $\text{diam}[f_n(V)] \le \varepsilon/8$. For any $x, y \in V$, we have $\rho_2(f(x), f(y)) \le \rho_2(f(x), f_n(x)) + \rho_2(f_n(x), f_n(b)) + \rho_2(f_n(b), f_n(y)) + \rho_2(f_n(y), f(y)) \le \varepsilon/8 + \varepsilon/8 + \varepsilon/8 + \varepsilon/8 = \varepsilon/2$. Hence $\text{diam}[f(V)] \le \varepsilon/2$ and therefore $\omega(f, x) \le \varepsilon/2$ for every $x \in V$. This shows $D_\varepsilon \cap V = \emptyset$, proving our claim.

The claim implies that $D := \bigcup_{n=1}^{\infty} D_1/n$ is an F_σ set of first category in X. This completes the proof since $\{x \in X : f$ is continuous at $x\} = X \setminus D$, and X is a complete metric space.

We know that the derivative of a differentiable real function need not be continuous. However, we can say the following.

Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. Then there exists a sequence (g_n) of continuous functions from $\mathbb{R}$ to $\mathbb{R}$ converging to f' pointwise. Consequently, $\{x \in \mathbb{R}: f'$ is continuous at $x\}$ is a dense G_δ subset of $\mathbb{R}$.

Notes

Proof: Since f is differentiable, f is continuous. Define $g_n(x) = [f(x + 1/n) - f(x)]/(1/n)$ and use [108].

Now we will show that a monotone real function (increasing or decreasing) is continuous and differentiable at most of the points.

Let $-\infty \le a < b \le \infty$ and let f: (a, b) → ℝ be a monotone function. Then, Y = {x ∈ (a, b) : f is discontinuous at x} is a countable set (possibly empty).

Proof: Suppose f is increasing. If x ∈ Y, then necessarily f(x–) < f(x+), and we may choose a rational number between f(x–) and f(x+). This gives a one-one map from Y to Q.

Definition: A collection Γ of non-degenerate intervals is a Vitali cover for a set X ⊂ ℝ if for each ε > 0, the subcollection {I ∈ Γ : 0 < |I| < ε} is also a cover for X.

[Vitali's covering lemma] Let X ⊂ ℝ be such that $\mu^*_{L,1}[X] < \infty$ and let Γ be a collection of intervals forming a Vitali cover for X. Then,

(i) There are countably many pairwise disjoint intervals $I_1, I_2, \ldots \in \Gamma$ such that $\mu^*_{L,1}[X \setminus \cup_n I_n] = 0$.

(ii) For every ε > 0, there exist finitely many pairwise disjoint intervals $I_1, \ldots, I_k \in \Gamma$ with the property that $\mu^*_{L,1}[X \setminus \bigcup_{n=1}^{k} I_n] < \varepsilon$.

Proof: Write $\mu^* = \mu^*_{L,1}$ for simplicity.

(i) With out loss of generality assume that every I ∈ Γ is a (non-degenerate) closed interval. Choose an open set U ⊂ ℝ such that X ⊂ U and μ*[U] < ∞. Every x ∈ X has a neighbourhood contained in U. Hence Γ′ = {I ∈ Γ : I ⊂ U} is also a vitali cover for X. We will choose the intervals I_n inductively. Let $\delta_0 = \sup\{|J| : J \in \Gamma'\}$ (note that $\delta_0 < \mu^*[U] < \infty$) and let $I_1 \in \Gamma'$ be any interval with $|I_1| > \delta_0/2$. Suppose that we have chosen pairwise disjoint intervals $I_1, \ldots, I_n \in \Gamma'$. If $X \subset \bigcup_{i=1}^{n} I_i$, then we are done. Else, any $x \in X \setminus \bigcup_{i=1}^{n} I_i$ is at a positive distance from the closed set $\bigcup_{i=1}^{n} I_i$. Let $\delta_n = \sup\{|J| : J \in \Gamma' \text{ and } I_i \cap J = \emptyset \text{ for } 1 \le i \le n\}$. Then $0 < \delta_n \le \mu^*[U] < \infty$. Let $I_{n+1} \in \Gamma'$ be an interval with $|I_{n+1}| > \delta_n/2$. We will show that the sequence (I_n) does the job.

Observation: For every J ∈ Γ, there is n ∈ ℕ such that $I_n \cap J \ne \emptyset$ ($\because \Sigma|I_n| \le \mu^*[U] < \infty$ so that $(|I_n|) \to 0$, and hence there is n ∈ ℕ such that $|I_n| < |J|/2$).

Let $Y = X \setminus \bigcup_{n=1}^{\infty} I_n$ and ε > 0. We claim that μ*[Y] < ε. Let c_n be the midpoint of I_n and let $Y_n \subset \mathbb{R}$ be the closed interval with midpoint c_n and $|Y_n| = 6|I_n|$ (this Y_n may not be in Γ′). Let k ∈ ℕ be so that $\sum_{n=k+1}^{\infty} |I_n| < \varepsilon/6$. If x ∈ Y, then in particular x does not belong to the closed set $\bigcup_{n=1}^{k} I_n$. Choose J ∈ Γ′ with x ∈ J and $I_n \cap J = \emptyset$ for 1 ≤ n ≤ k. By our observation above, $I_m \cap J \ne \emptyset$ for some m ≥ k + 1. Let m be the smallest such number. Then $|J| \le \delta_{m-1} < 2|I_m|$ and hence $|x - c_m| \le |J| + |I_m| \le 3|I_m|$. Therefore, $x \in Y_m$. We have shown that $Y \subset \bigcup_{n=k+1}^{\infty} Y_n$. Since $\sum_{n=k+1}^{\infty} |Y_n| \le 6\sum_{n=k+1}^{\infty} |I_n| < \varepsilon$, we have proved that μ*[Y] < ε.

Now, note that the argument given above actually shows that for every ε > 0, there is k ∈ ℕ such that $\mu^*[X \setminus \bigcup_{n=1}^{k} I_n] < \varepsilon$. Hence we have established (ii) also.

When a mathematical problem is difficult, it is a good idea to divide the problem into many subcases and to treat each case separately. If f : (a, b) → ℝ is a function, then the four Dini derivatives of f at a point x ∈ (a, b) are defined as follows.

$$D^+f(x) = \limsup_{h \to 0+} \frac{f(x+h) - f(x)}{h} \qquad \text{[upper right derivative]}$$

$$D_+f(x) = \liminf_{h \to 0+} \frac{f(x+h) - f(x)}{h} \qquad \text{[lower right derivative]}$$

Notes

$$D^-f(x) = \limsup_{h\to 0-} \frac{f(x+h)-f(x)}{h} \qquad \text{[upper left derivative]}$$

$$D_-f(x) = \liminf_{h\to 0-} \frac{f(x+h)-f(x)}{h} \qquad \text{[lower left derivative].}$$

Note Here, $\limsup_{h\to 0+} \frac{f(x+h)-f(x)}{h} := \lim_{y\to 0+}\left[\sup_{0<h<y} \frac{f(x+h)-f(x)}{h}\right]$, and similarly the others.

Example: Let f: $(-1,1) \to \mathbb{R}$ be $f(0) = 0$ and $f(x) = x \sin(1/x)$ for $x = 0$. Then, $D^+f(0) = 1 = D^-f(0)$ and $D_+f(0) = -1 = D_-f(0)$ so that f is not differentiable at 0.

Notes That f is differentiable at x iff all the four Dini derivatives are equal and real (i.e., different from $\pm\infty$). Since $D_+f(x) \le D^+f(x)$ and $D_-f(x) < D^-f(x)$ by definition, we also see that f is differentiable at x iff the four Dini derivatives are real numbers satisfying $D^+f(x) \le D_-f(x)$ and $D^-f(x) \infty D_+f(x)$.

[Lebesgue's differentiation theorem] Let $-\infty \le a < b \le \infty$, let $f : (a, b) \to \mathbb{R}$ be a monotone function and let $Y = \{x \in (a, b) : f \text{ is not differentiable at } x\}$. Then $\mu^*_{L,1}[Y] = 0$.

Proof: Since (a, b) can be written as a countable union of bounded open intervals, we may as well assume (a, b) itself is bounded. Assume f is increasing and write $\mu^* = \mu^*_{L,1}$. By the remark above, $Y = Y_1 \cup Y_2$, where $Y_1 = \{x \in (a, b) : D_- + f(x) < D^+f(x)\}$ and $Y_2 = \{x \in (a, b) : D_+f(x) < D^-f(x)\}$. We will only show that $\mu^*[Y_1] = 0$; the case of Y_2 is similar.

Let $\Gamma = \{(r, s) \in \mathbb{Q}^2 : r < s\}$, let $X(r, s) = \{x \in (a,b): D_-f(x) < r < s < D^+f(x)\}$ and note that $Y_1 = U_{(rs)\in\Gamma} X(r, s)$. Hence it suffices to show there $\mu^*[X(r, s)] = 0$ for every $(r, s) \in \Gamma$. Fix $(r, s) \in \Gamma$, write $X = X(r, s)$ and let $\varepsilon > 0$ be arbitrary. Choose an open set $U \subset (a, b)$ such that $X \subset U$ and $\mu^*[U] < \mu^*[X] + \varepsilon$.

Since $D_-f < r$ on X, for each $x \in X$ and $\delta > 0$ we can find a non-degenerate closed interval $I(x, \delta) = [x - \alpha, x] \subset U$ such that $0 < \alpha < \delta$ and $f(x) - f(x - \alpha) < r\alpha$. Then $\Gamma = \{I(x, \delta) : x \in X, \delta > 0\}$ is a Vitali cover for X. By Vitali's lemma, we can find finitely many pairwise disjoint intervals $I_1,\ldots,I_k \in \Gamma$ such that $\mu^*[X \setminus \bigcup_{n=1}^{k} |I_n|] < \varepsilon$.

Let $V = \bigcup_{n=1}^{k} \text{int}[I_n]$. Then, V is open, $V \subset U$, and $\mu^*[X] - \varepsilon < \mu^*[V] \le \mu^*[U] < \mu^*[X] + \varepsilon$. Let $X' = V \cap X$. Since $D^+f > s$ on X, and hence on X′, for each $y \in X'$ and $\delta > 0$ we can find a non-degenerate closed interval $J(y, \delta) = [y, y + \beta] \subset V$ (hence $J(y, \delta) \subset I_n$ for some $n \in \{1,\ldots,k\}$) such that $0 < \beta < \delta$, and $f(y + \beta) - f(y) > s\beta$. Then $\Gamma = \{J(y, \delta) : y \in X', \delta > 0\}$ is a Vitali cover for X′. Again by Vitali's lemma, we can find finitely many pairwise disjoint intervals $J_1,\ldots, J_m \in \Gamma$ such that $\mu^*[X' \setminus \bigcup_{j=1}^{m} |J_j|] < \varepsilon$. Then $\bigcup_{j=1}^{m} |J_j| \ge \mu^*[X'] - \varepsilon \ge \mu^*[X] - 2\varepsilon$.

Write $I_n = [x_n - \alpha_n, x_n]$ and $J_j = [y_j, y_j + \beta_j]$. For each $n \in \{1,\ldots,k\}$, let $D_n = \{j \in \{1,..,m\} : J_j \subset I_n\}$. Then $\{1,\ldots, m\}$ is the disjoint union of D_n's.

Note that $\sum_{j\in D_n} (f(y_j + \beta_j) - f(y_j)) \le f(xn) - f(x_n - \alpha_n)$ for each $n \in \{1,\ldots,k\}$ since f is increasing. Summing over n, we get $\sum_{j=1}^{m} (f(y_j + \beta_j) - f(y_j)) \le \sum_{n=1}^{k} (f(x_n) - f(x_n - \alpha_n))$, and hence $\sum_{j=1}^{m} s\beta_j < \sum_{n=1}^{k} r\alpha_n$, or $s(\sum_{j=1}^{m} |J_j|) < r(\sum_{n=1}^{k} |I_n|)$. From the earlier estimates we conclude that $s(\mu^*[X] - 2\varepsilon) < r(\mu^*[X] + \varepsilon)$. Since $\varepsilon > 0$ was arbitrary and $r < s$, we must have $\mu^*[X] = 0$.

The conclusion is, it can be extended to more general class of real functions.

Definition: If f: [a, b] → ℝ is a function and P = {$a_0 = a \le a_1 \le \ldots \le a_{n-1} \le a_n = b$} is a partition of [a, b], let $V_a^b(f, P) = \sum_{i=1}^{n} |f(a_i) - f(a_{i-1})|$. Define the total variation of f as $V_a^b(f) = \sup\{V_a^b(f, P) : P$ is a partition of [a, b]}. We say f is of bounded variation if $V_a^b(f) < \infty$. It is easy to see that if f is of bounded variation, then f is bounded (∵ if x ∈ [a, b], take P = {a ≤ x ≤ b} to see that $|f(x) - f(a)| \le V_a^b(f)$).

Examples:

(i) If f : [a, b] → ℝ is a monotone function, then $V_a^b(f, P) = |f(b) - f(a)|$ for any partition P of [a,b] and hence $V_a^b(f) = |f(b) - f(a)| < \infty$. So f is of bounded variation.

(ii) Suppose f : [a,b] → ℝ is Lipschitz continuous (this happens if f is C^1) with Lipschitz constant λ > 0. Then, it may be seen that $V_a^b(f) \le \lambda(b - a) < \infty$ and hence f is of bounded variation.

Example: A (uniformly) continuous function f : [a, b] → ℝ need not be of bounded variation. Let f: [0, 1] → ℝ be the (uniformly) continuous function defined as f(0) = 0 and f(x) = x sin (1/x) if x ∈ (0,1). Let $a_k = 2/k\pi \in [0,1]$ for k ∈ ℕ. Observe that $|f(a_{2k}) - f(a_{2k-1})| = |0 - a_{2k-1}| = a_{2k-1}$. Let m ∈ ℕ and P = {$0 \le a_{2m} \le a_{2m-1} \le \ldots \le a_1 \le 1$}. Then $V_0^1(f, P) \ge \sum_{k=1}^{m} |f(a_{2k}) - f(a_{2k-1})| = \sum_{k=1}^{m} a_{2k-1} = (2/\pi) \sum_{k=1}^{m} (2K - 1)^{-1} \to \infty$ as m → ∞. Hence $V_0^1(f) = \infty$, and thus f is not of bounded variation. This example also shows that bounded ⇏ bounded variation.

Notes If f, g: [a, b] → ℝ are of bounded variation, r, s ∈ ℝ, and h : [a, b] → ℝ is defined as h = r f(x) + sg(x), then $V_a^b(h) \le |r|\, V_a^b(f) + |s|\, V_a^b(g) < \infty$. Hence {f : [a,b] → ℝ : f is of bounded variation } is a real vector space (in fact, it is a normed space with the norm $\|f\| = |f(a)| + V_a^b(f)$).

A function f : [a, b]→ ℝ is of bounded variation iff there exist monotone functions g, h : [a, b]→ ℝ such that f(x) = g(x) – h(x) for every x ∈ [a,b]. Consequently, for any function f : [a, b] → ℝ of bounded variation, we have $\mu^*_{L,1}$ [{x ∈ [a, b] : f is not differentiable at x}] = 0.

Proof: Suppose f = g – h, where g, h are monotone. Since g, h are of bounded variation, f is also of bounded variation since the collection of functions of bounded variation on [a, b] is a vector space. Conversely assume that f is of bounded variation and define g : [a, b] → ℝ as $g(x) = V_a^x(f)$. Then g is monotone increasing. Let h = g – f, and consider points x < y in [a, b]. We have $g(y) - g(x) = V_x^y(f) \ge |f(y) - f(x)| \ge f(y) - f(x)$, and therefore h(y) ≥ h(x). Thus h is also monotone increasing. Clearly, f = g – h.

Let [a, b] be a compact interval. Then for a function f : [a, b] → ℝ, we have the following implications: f is Lipschitz continuous ⇒ f is absolutely continuous ⇒ f is of bounded variation. Consequently, if f is either Lipschitz continuous or absolutely continuous, then $\mu^*_{L,1}$ [Y] = 0, where Y = {x ∈ [a, b] : f is not differentiable at x}.

Note However, there is a limit to these type of results; there are continuous functions f : [a, b] → ℝ which are not differentiable at any point.

Now we mention a characterization of Riemann integrable functions in terms of small sets. For simplicity, we restrict ourselves to dimension one, even though the corresponding result is true in higher dimensions as well. If f : [a, b] → ℝ is a bounded function and if P = {$a_0 = a < a_1 \le \cdots \le a_{n-1}$

$\leq a_n = b\}$ is a partition of [a, b], let $M_i = \sup\{f(x) : a_{i-1} \leq x \leq a_i\}$ and $m_i = \inf\{f(x) : a_{i-1} \leq x \leq a_i\}$. The upper and lower Riemann sums with respect to the partition P are defined as $U(f,P) = \sum_{i=1}^{n} M_i(a_i - a_{i-1})$ and $L(f, P) = \sum_{i=1}^{n} m_i(a_i - a_{i-1})$. A bounded function f: $[a, b] \to \mathbb{R}$ is said to be Riemann integrable if for every $\varepsilon > 0$ there is a partition P of [a, b] such that $U(f,P) - L(f,P) < \varepsilon$. The following characterization says that a Riemann integrable function is not very different from a continuous function.

Let $f : [a, b] \to \mathbb{R}$ be a bounded function and let $Y = \{x \in [a, b] : f \text{ is not continuous at } x\}$. Then, f is Riemann integrable iff $\mu^*_{L,1}[Y] = 0$.

Proof: Let $\omega(f,x)$ be the oscillation of f at x defined earlier.

$\Rightarrow$: Since $Y = \bigcup_{k=1}^{\infty} Y_k$, where $Y_k = \{x \in [a, b] : \omega(f, x) \geq 1/k\}$, it suffices to show $\mu^*_{L,1}[Y_k] = 0$ for every $k \in \mathbb{N}$. Fix $k \in \mathbb{N}$ and let $\varepsilon > 0$. Let $P = \{a_0 = a \leq a_1 \leq \cdots \leq a_{n-1} \leq a_n = b\}$ be a partition of [a, b] with $U(f,P) - L(f,P) < \varepsilon/2k$. Let $A_i = (a_{i-1}, a_i)$ and $\Gamma = \{1 \leq i \leq n : Y_k \cap A_i \neq \emptyset\}$. Note that $M_i - m_i \geq 1/k$ for $i \in \Gamma$. Write $Y_k = Y_k' \cap Y_k''$, where $Y_k' = Y_k \cap (\bigcup_{i\in\Gamma} A_i)$ and $Y_k'' = Y_k \cap \{a_1,\ldots, a_n\}$. We have $\varepsilon/2k > U(f, P) - L(f, P) \geq \sum_{i\in\Gamma}(M_i - m_i)|A_i| \geq 1/k \sum_{i\in\Gamma}|A_i|$ and hence $\sum_{i\in\Gamma}|A_i| < \varepsilon/2$. And since Y_k'' is a finite set, there are finitely many intervals $B_1,\ldots,B_m$ such that $Y_k'' \subset \bigcup_{j=1}^{m}|Bj$ and $\sum_{j=1}^{m}|B_j| < \varepsilon/2$. Thus $Y_k \subset [\bigcup_{i\in\Gamma} A_i] \cup [\bigcup_{j=1}^{m} B_j]$ and $\sum_{i\in\Gamma}|A_i| + \sum_{j=1}^{m}|B_j| < \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $\mu^*_{L,1}[Y_k] = 0$.

$\Leftarrow$: Let $\varepsilon > 0$ be given. We have to find a partition P of [a, b] such that $U(f, P) - L(f, P) < \varepsilon$. Let $\lambda = \sup\{|f(x)| : x \in [a, b]\}$ and let $\varepsilon' = \varepsilon/[2\lambda + 2(b - a)]$. For each $x \in [a, b]\backslash Y$, choose an open interval $A(x) \subset \mathbb{R}$ containing x such that $|f(x) - f(z)| < \varepsilon'$ for every $z \in [a, b] \cap \overline{A(x)}$ by continuity. Also choose countably many open intervals B_m such that $Y \subset \bigcup_{m=1}^{\infty} B_m$ and $\bigcup_{m=1}^{\infty}|B_m| < \varepsilon'$. Then $\{A(x) : x \in [a,b]\backslash Y\} \cap \{B_m : m \in \mathbb{N}\}$ is an open cover for the compact set [a, b]. Extract a finite subcover $\{A(x_j) : 1 \leq j \leq p\} \cup \{B_m : 1 \leq m \leq q\}$. The end points inside [a, b] of these finitely many intervals determine a partition $P = \{a_0 = a \leq a_1 \leq \cdots \leq a_{n-1} \leq a_n = b\}$ of [a, b]. Observe that for each $i \in \{1,\ldots,n\}$, we have $[a_{i-1}, a_i] \subset \overline{A(x_j)}$ for some $j \in \{1,\ldots,p\}$, or $[a_{i-1}, a_i] \subset \overline{B_m}$ for some $m \in \{1,\ldots,q\}$. Let $\Gamma = \{1 \leq i \leq n : [a_{i-1}, a_i] \subset \overline{A(x_j)}$ for some $j\}$ and $\Gamma' = \{1,\ldots,n\}\backslash\Gamma$. Note that $M_i - m_i < 2\varepsilon'$ if $i \in \Gamma$. Hence $U(f, P) - L(f,P) \leq \sum_{i\in\Gamma}(M_i - m_i)(a_i - a_{i-1}) + \sum_{i\in\Gamma}(M_i - m_i)(a_i - a_{i-1}) \leq 2\varepsilon' \sum_{i\in\Gamma}(a_i - a_{i-1}) + 2\lambda \sum_{i\in\Gamma}(a_i - a_{i-1}) \leq 2\varepsilon' \sum_{i=1}^{n}(a_i - a_{i-1}) + 2\lambda \sum_{m=1}^{q}|B_m| < 2\varepsilon'(b - a) + 2\lambda\varepsilon' = \varepsilon$.

A corollary is that any bounded function $f : [a, b] \to \mathbb{R}$ with at most countably many points of discontinuity (in particular, any continuous function) is Riemann integrable. The higher dimensional generalization can be stated as follows.

Let $A \subset \mathbb{R}^d$ be a d-box, let $f : A \to \mathbb{R}$ be a bounded function and let Y be the set $\{x \in A : f$ is not continuous at $x\}$. Then, f is Riemann integrable iff $\mu^*_{L,d}[Y] = 0$.

Example: Let $f : [0,1] \to \mathbb{R}$ be $f(0) = 0$ and $f(x) = \sin(1/x)$ for $x = 0$. Even though the graph of f has infinitely many ups and downs (in fact, f is not of bounded variation), f is Riemann integrable since f is bounded and is discontinuous only at one point, namely 0.

Definition: Let X be a set and $A \subset X$. The characteristic function $\chi_A : X \to \mathbb{R}$ of the subset A is defined as $\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \in X \setminus A. \end{cases}$

Example: We discuss an example that illustrates the main drawback of Riemann integration theory. Write $[0,1] \cap \mathbb{Q} = \{r_1, r_2,\ldots\}$, let $f_n : [0, 1] \to \mathbb{R}$ be the characteristic function of $\{r_1,\ldots, r_n\}$, and let $f : [0,1] \to \mathbb{R}$ be the characteristic function of $[0,1] \cap \mathbb{Q}$. We have $0 \leq f_1 \leq f_2 \leq \cdots \leq f \leq 1$ and the sequence (f_n) converges to f pointwise. Each f_n is Riemann integrable since f_n is discontinuous only at finitely many points. But f is discontinuous at every point of [0,1], and the Lebesgue outer measure of [0,1] is positive. Hence f is not Riemann integrable by [115]. Thus even the pointwise limit of a uniformly bounded, monotone sequence of Riemann integrable functions need not be Riemann integrable.

(i) Let f: [0, 1] → $\mathbb{R}$ be the characteristic function of [0, 1] ∩ $\mathbb{Q}$. Since f is not continuous at any point, it is not possible to realize f as the pointwise limit of a sequence of continuous functions from [0, 1] to $\mathbb{R}$, in view of [108].

(ii) Let (f_n) be a sequence of continuous functions from [a, b] to $\mathbb{R}$ converging pointwise to a function f : [a, b] → $\mathbb{R}$, and let Y = {x ∈ [a, b] : f is not continuous at x}. From [108] we know that Y is an F_σ set of first category in [a, b]. But Y can have positive outer Lebesgue measure by [106]. Hence f may not be Riemann integrable. Thus even the pointwise limit of a sequence of continuous functions may not be Riemann integrable (of course, we did not give an example).

(iii) Lebesgue integration theory is developed not just for the sake of making the characteristic function of [0,1] ∩ $\mathbb{Q}$ integrable. The limit theorems in Lebesgue's theory allow us to integrate the pointwise limit of a sequence of integrable functions, and to interchange limit and integration, under very mild hypothesis. Moreover, the powerful tools in Lebesgue's theory make many proofs simpler (e.g.: the proof of the change of variable theorem in d-dimension), and provide us with new ways of dealing with functions (e.g.: L^p spaces). Also, as we will see later, in Lebesgue's theory we have a more satisfactory version of the Fundamental Theorem of Calculus (describing differentiation and integration as inverse operations of each other).

13.4 σ-algebras and Measurable Spaces

A d-box in $\mathbb{R}^d$ has a well-defined d-dimensional volume. We may ask whether it is possible to define the notion of a d-dimensional value for all subsets of $\mathbb{R}^d$. Of course, we would like to have consistency conditions such as monotonicity and countable additivity.

Question: Can we have a function μ : P($\mathbb{R}^d$) → [0, ∞] such that

(i) $\mu[A] = \mathrm{Vol}_d(A)$ if $A \subset \mathbb{R}^d$ is ad-box,

(ii) [Monotonicity] $\mu[A] \le \mu[B]$ for subsets A, B of $\mathbb{R}^d$ with $A \subset B$,

(iii) [Countable additivity] $\mu[\bigcup_{n=1}^{\infty} A_n] = \sum_{n=1}^{\infty} \mu[A_n]$ if A_n's are pairwise disjoint subsets of $\mathbb{R}^d$?

Notes We know that the Lebesgue outer measure $\mu^*_{L,d}$ does not satisfy countable additivity. The key observation of Lebesgue's theory is that $\mu^*_{L,d}$ will satisfy all the three conditions stated above if we restrict $\mu^*_{L,d}$ to a slightly smaller collection $\mathcal{A} \subset \mathbb{P}(\mathbb{R}^d)$ by discarding some pathological subsets of $\mathbb{R}^d$. In order to describe the structure of this smaller collection $\mathcal{A}$, it is convenient to proceed in an abstract manner, which we do below.

Definition: Let X be a nonempty set. A collection $\mathcal{A} \in \mathbb{P}(X)$ of subsets of X is said to be a σ-algebra on X if the following hold:

(i) $\emptyset, X \in \mathcal{A}$.

(ii) $A \in \mathcal{A} \Rightarrow X \backslash A \in \mathcal{A}$.

(iii) $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

If $\mathcal{A}$ is a σ-algebra on X, then (X, $\mathcal{A}$) is called a measurable space.

Notes

Example: {Ø, X} and $\mathbb{P}(X)$ are trivial examples of σ-algebras on any nonempty set X. The following are some σ-algebras on $\mathbb{R}^d$ (verify):

$\mathcal{A}_1 = \{A \subset \mathbb{R}^d: A \text{ or } \mathbb{R}^d \setminus A \text{ is countable}\}$,

$\mathcal{A}_2 = \{A \subset \mathbb{R}^d: A \text{ or } \mathbb{R}^d \setminus A \text{ is of first category in } \mathbb{R}^d\}$,

$\mathcal{A}_3 = \{A \subset \mathbb{R}^d: \mu^*_{L,d}[A] = 0 \text{ or } \mu^*_{L,d}[\mathbb{R}^d \setminus A] = 0\}$.

$\mathcal{A}_4 = \{A \subset \mathbb{R}^d: [0, 1]^d \subset A \text{ or } [0, 1]^d \subset \mathbb{R}^d \setminus A\}$.

Definition: Let X be a nonempty set and $\mathcal{C} \subset \mathbb{P}(X)$ be a collection of subsets of X. A σ-algebra $\mathcal{A}$ on X is said to be generated by $\mathcal{C}$ if $\mathcal{A}$ is the smallest σ-algebra on X containing $\mathcal{C}$. Here, $\mathcal{A}$ exists and is unique since $\mathcal{A}$ is precisely the intersection of all σ-algebras on X containing $\mathcal{C}$ (note that there is at least one σ-algebra on X containing $\mathcal{C}$, namely $\mathbb{P}(X)$).

Definition: Let X be a metric space. Then the σ-algebra on X generated by the collection of all open subsets of X is called the Borel σ-algebra on X, and is denoted as $\mathcal{B}(X)$ (or just $\mathcal{B}$, if X is clear from the context). The subsets of X belonging to $\mathcal{B}(X)$ are called Borel subsets of X. For example, open subsets, closed subsets, G_δ subsets and F_σ subsets of X are Borel subsets of X.

[Characterizations of the Borel σ-algebra on $\mathbb{R}^d$] Consider the following collections of subsets of $\mathbb{R}^d$:

$\mathcal{C}_1 = \{A \subset \mathbb{R}^d: A \text{ is closed}\}$,

$\mathcal{C}_2 = \{A \subset \mathbb{R}^d: A \text{ is compact}\}$,

$\mathcal{C}_3 = \{A \subset \mathbb{R}^d: A \text{ is closed d-box}\}$,

$\mathcal{C}_4 = \{A \subset \mathbb{R}^d: A \text{ is an opend-box}\}$,

$\mathcal{C}_5 = \{A \subset \mathbb{R}^d: A \text{ is ad-box}\}$,

$\mathcal{C}_6 = \{A \subset \mathbb{R}^d: A \text{ is an open ball}\}$,

$\mathcal{C}_7 = \{f^{-1}(W) : f : \mathbb{R}^d \to \mathbb{R} \text{ is continuous and } W \subset \mathbb{R} \text{ is open}\}$.

If $\mathcal{A}_i$ is the σ-algebra on $\mathbb{R}^d$ generated by $\mathcal{C}_i$ for $1 \le i \le 7$, then $\mathcal{A}_i = \mathcal{B}(\mathbb{R}^d)$ for $1 \le i \le 7$.

Proof: Clearly $\mathcal{A}_3 \subset \mathcal{A}_2 \subset \mathcal{A}_1 = \mathcal{B}(\mathbb{R}^d)$. Since $(a, b) = \bigcup_{n=n_0}^{\infty} [a + 1/n, b - 1/n]$ (where n_0 is chosen so that $a + 1/n_0 \le b - 1/n_0$), it follows that any open d-box°is a countable union of closed d-boxes, and therefore $\mathcal{A}_4 \subset \mathcal{A}_3$. Since $[a,b] = \bigcup_{n=1}^{\infty} (a - 1/n, b + 1/n)$, $[a, b) = \bigcup_{n=1}^{\infty} (a - 1/n, b)$, and $(a, b] = \bigcup_{n=1}^{\infty} (a, b + 1/n)$, we deduce that any d-box is a countable intersection of open d-boxes, and hence $\mathcal{A}_4 = \mathcal{A}_5$. Since any open set in $\mathbb{R}^d$ can be written as a countable union of open d-boxes as well as a countable union of open balls, we have $\mathcal{A}_4 = \mathcal{A}_6 = \mathcal{B}(\mathbb{R}^d)$. Thus $\mathcal{A}_i = \mathcal{B}(\mathbb{R}^d)$ for $1 \le i \le 6$.

By the definition of continuity, we have $\mathcal{A}_7 \subset \mathcal{B}(\mathbb{R}^d)$. If $U \subset \mathbb{R}^d$ is an open set different from $\mathbb{R}^d$, let $A = \mathbb{R}^d \setminus U$ and define $f : \mathbb{R}^d \to \mathbb{R}$ as $f(x) = \text{dist}(x, A) := \inf\{\| x - a \| : a \in A\}$. Then f is continuous, and $A = f^{-1}(0)$ because A is closed. Now, $U = f^{-1}(\mathbb{R}\setminus\{0\})$ and $\mathbb{R}\setminus \{0\}$ is open in $\mathbb{R}$. Hence $\mathcal{B}(\mathbb{R}^d) \subset \mathcal{A}_7$, completing the proof.

Topological Remarks:

(i) If X is a separable metric space, then any base or subbase for the topology of X will generate the Borel σ-algebra $\mathcal{B}(X)$.

(ii) In the above characterization we used implicitly the fact that $\mathbb{R}^d$ is second countable and locally compact. If a metric space X fails to be second countable or locally compact, then the σ-algebra generated by all compact subsets of X will only be a proper sub-collection of $\mathcal{B}(X)$. For example, try to figure out what happens for the spaces ($\mathbb{R}$, discrete metric) (which is not second countable), and ($\mathbb{Q}$, Euclidean metric) (which is not locally compact).

Next our aim is to determine the cardinality of $\mathcal{B}(\mathbb{R}^d)$. We need some set-theoretic preparation.

Definition: An order $\leq$ on a set X is a partial order if (i) $x \leq x$ for every $x \in X$, (ii) $x \leq y$ and $y \leq x \Rightarrow x = y$ for every $x, y \in X$, (iii) $x \leq y$ and $y \leq z \Rightarrow x \leq z$ for every $x, y, z \in X$. We say $(X, \leq)$ is a totally ordered set if $\leq$ is a partial order and any two elements of X are comparable. We say $(X, \leq)$ is a well-ordered set if $(X, \leq)$ is totally ordered and any nonempty subset $Y \subset X$ has a least element in Y.

Examples:

(i) Let X be the collection of all nonempty subsets of $\mathbb{N}$. Define an order $\leq$ on X as $A \leq B$ iff the minimum of A is less than or equal to the minimum of B. Then this is not a partial order since the second condition fails.

(ii) If X is any nonempty set, then $\mathbb{P}(X)$ with inclusion as order is partially ordered, but in general not totally ordered.

(iii) $\mathbb{R}$ with the usual order is totally ordered, but not well-ordered since the subset (0,1) does not contain a least element.

(iv) $\mathbb{N}$ with the usual order is well-ordered.

Well-ordering principle (equivalent to the axiom of choice): Any nonempty set admits a well-ordering.

Now we describe the construction of some ordinal numbers. Start with an uncountable set X such that card(X) = card($\mathbb{R}$), and let $\leq$ be a well-ordering on X. Let θ denote the least element of X. By adding one extra element to X if necessary, we may also assume that $(X, \leq)$ has a largest element, say θ'. For each $\beta \in X$, let $L_\beta = \{\alpha \in X : \alpha < \beta\}$ be the left section of β in X. Let $Y = \{\beta \in X : L_\beta$ is uncountable$\}$. Then $Y \neq \emptyset$ since $\theta' \in Y$. So Y has a least element, say Ω. Then L_Ω is uncountable, but L_β is countable for every $\beta < \Omega$. Here, Ω is called the first uncountable ordinal, and each $\beta \in L_\Omega$ is called a countable ordinal number since each $\beta \in L_\Omega$ represents the type of a countable well-ordered set through L_β.

Fact: If $A \subset L_\Omega$ is a nonempty countable set, then A has a least upper bound in L_Ω. [*Proof:* If $B = \bigcup_{\beta \in A} L_\beta$, then B is countable and hence $L_\Omega \setminus B \neq \emptyset$. The least element of $L_\Omega \setminus B$ is the least upper bound of A]

If $\alpha \in L_\Omega$, then the least element of the nonempty set $\{\beta \in L_\Omega : \alpha < \beta\}$ will be denoted as $\alpha + 1$. Note that there are no elements between α and $\alpha + 1$ in L_Ω. On the other hand, given $\beta \in L_\Omega$, there may or may not exist $\alpha \in L_\Omega$ such that $\alpha + 1 = \beta$. For example, if $\beta \in L_\Omega$ is the least upper bound of the countable set $\{\theta, \theta + 1, \theta + 2, \ldots\}$ (where recall that θ is the least element of L_Ω), then there is no $\alpha \in X$ with $\alpha + 1 = \beta$. We say $\beta \in L_\Omega$ is a limit ordinal if there is no $\alpha \in L_\Omega$ with $\alpha + 1 = \beta$.

card($\mathcal{B}(\mathbb{R}^d)$) = card($\mathbb{R}$).

Proof: We will use transfinite induction (i.e., induction with respect to ordinal numbers) by using L_Ω described above. Recall that we denoted the least element of L_Ω by the symbol θ. To start the induction process, let $\mathcal{A}_\theta = \{U \subset \mathbb{R}^d : U \text{ is open}\}$. Let $\beta \in L_\Omega$ and assume that we have defined $\mathcal{A}_\alpha$ for every $\alpha \in L_\beta$. If β is a limit ordinal, define $\mathcal{A}_\beta = \bigcup_{\alpha<\beta} A_\alpha$. If $\beta = \alpha + 1$ for some $\alpha \in L_\Omega$, let $\mathcal{A}_\alpha' = \{A \subset \mathbb{R}^d : \mathbb{R}^d \setminus A \in \mathcal{A}_\alpha\}$, and $\mathcal{A}_\beta = \{A \subset \mathbb{R}^d : A$ is a countable union of members from $\mathcal{A}_\alpha \cup \mathcal{A}_\alpha'\}$. This defines $\mathcal{A}_\beta$ for every $\beta \in L_\Omega$. Finally, put $\mathcal{A} = \bigcup_{\beta<\Omega} \mathcal{A}_\beta$. From our construction, it is clear that $\mathcal{A} \subset \mathcal{B}(\mathbb{R}^d)$.

We verify that $\mathcal{A}$ is a σ-algebra on $\mathbb{R}^d$. It suffices to check only the third property. So consider $A_1, A_2, \ldots \in \mathcal{A}$. Then there are $\beta_1, \beta_2, \ldots \in L_\Omega$ such that $A_n \in \mathcal{A}_{\beta_n}$ for every $n \in \mathbb{N}$. By the Fact mentioned above, the countable set $\{\beta_n : n \in \mathbb{N}\}$ has a least upper bound, say δ in L_Ω. Then $A_n \in \mathcal{A}_\delta$

Notes

for every $n \in \mathbb{N}$ and hence $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_{\delta+1} \subset \mathcal{A}$. Thus $\mathcal{A} \subset \mathcal{B}$ is a σ-algebra on $\mathbb{R}^d$ containing all open subsets of $\mathbb{R}^d$. Hence $\mathcal{A} = \mathcal{B}(\mathbb{R}^d)$.

Now, it suffices to show that card($\mathcal{A}$) = card($\mathbb{R}$). Since there is an open ball of radius 1 centered at each point of $\mathbb{R}^d$, we have card($\mathcal{A}$) ≥ card($\mathcal{A}_0$) ≥ card($\mathbb{R}$). So it suffices to establish that card($\mathcal{A}$) ≤ card($\mathbb{R}$). Since card(L_Ω) = card($\mathbb{R}$) and $\mathcal{A} = \bigcup_{\beta<\Omega} \mathcal{A}_\beta$, it is enough to show that card($\mathcal{A}_\beta$) ≤ card($\mathbb{R}$) for each $\beta \in L_\Omega$.

Let $\mathcal{D}$ be the collection of all open balls in $\mathbb{R}^d$ with rational radius and center in $\mathbb{Q}^d$. Then $\mathcal{D}$ is countable, and any open set $U \subset \mathbb{R}^d$ can be written as a countable union of members of $\mathcal{D}$. Hence card($\mathcal{A}_0$) ≤ card($\mathcal{D}^{\mathbb{N}}$) = card($\mathbb{R}$). Let $\beta \in L_\Omega$ and suppose we have proved that card($\mathcal{A}_\alpha$) ≤ card($\mathbb{R}$) for every $\alpha < \beta$. If β is a limit ordinal, then $\mathcal{A}_\beta = \bigcup_{\beta<\Omega} \mathcal{A}_\alpha$ is a countable union and hence card($\mathcal{A}_\beta$) ≤ card($\mathbb{R}$). If there is α with α + 1 = β, then any $A \in \mathcal{A}_\beta$ can be written as $A = \bigcup_{n=1}^{\infty} A_n$ with $A_n \in A_\alpha \cup \mathcal{A}_\alpha'$. This gives a one-one map from $\mathcal{A}_\beta$ into $(A_\alpha \cup \mathcal{A}_\alpha')^{\mathbb{N}}$. Hence card($\mathcal{A}_\beta$) ≤ card($(A_\alpha \cup \mathcal{A}_\alpha')^{\mathbb{N}}$) ≤ card($\mathbb{R}$). This completes the proof.

Corollary: For any uncountable set $Y \subset \mathbb{R}^d$, there is $A \subset Y$ such that A is not a Borel subset of$\mathbb{R}^d$.

Proof: We have card($\mathcal{B}(\mathbb{R}^d)$) = card($\mathbb{R}$) = card(Y) < card($\mathbb{P}$(Y)).

Definition: Let $\mathcal{N}(\mathbb{R}^d) = \{A \subset \mathbb{R}^d: \mu_{L,d}^*[A] = 0\}$. The members of $\mathcal{N}(\mathbb{R}^d)$ are called Lebesgue null sets. The σ-algebra $\mathcal{L}(\mathbb{R}^d)$ on $\mathbb{R}^d$ generated by $\mathcal{B}(\mathbb{R}^d) \cup \mathcal{N}(\mathbb{R}^d)$ is called the Lebesgue σ-algebra on $\mathbb{R}^d$, and members of $\mathcal{L}(\mathbb{R}^d)$ are called Lebesgue measurable subsets of $\mathbb{R}^d$.

card($\mathcal{N}(\mathbb{R}^d)$) = card($\mathcal{L}(\mathbb{R}^d)$) = card($\mathbb{P}(\mathbb{R}^d)$) > card($\mathbb{R}$). Hence, $\mathcal{N}(\mathbb{R}^d) \subsetneq \mathcal{B}(\mathbb{R}^d) \subsetneqq \mathcal{L}(\mathbb{R}^d)$.

Proof: Let K be the middle-third Cantor set. Then, for any subset $A \subset K$, we have $\mu_{L,1}^*[A] \le \mu_{L,1}^*[K] = 0$. So $\mu_{L,d}^*[A] = 0$ also. This shows that $\mathbb{P}(K) \subset \mathcal{N}(\mathbb{R}^d) \subset \mathcal{L}(\mathbb{R}^d)$. And card($\mathbb{P}$(K)) = card($\mathbb{P}(\mathbb{R})$) since K is an uncountable subset of $\mathbb{R}$.

[Translation invariance] (i) $A + x \in \mathcal{B}(\mathbb{R}^d)$ for every $A \in \mathcal{B}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$.

(ii) $A + x \in \mathcal{N}(\mathbb{R}^d)$ for every $A \in \mathcal{N}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$.

(iii) $A + x \in \mathcal{L}(\mathbb{R}^d)$ for every $A \in \mathcal{L}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$.

Proof: First let us mention a general principle that will be used at many places. To establish that the members of a certain σ-algebra $\mathcal{D}$ on a set X satisfies a certain property P, it suffices to do the following: show that the collection {A ⊂ X: A satisfies property P} is a σ-algebra, and then find a suitable collection $\mathcal{C} \subset \mathbb{P}(X)$ generating $\mathcal{D}$ and show that every member of $\mathcal{C}$ satisfies the property P.

Let $\mathcal{A} = \{A \subset \mathbb{R}^d: A + x \in \mathcal{B}(\mathbb{R}^d) \text{ for every } x \in \mathbb{R}^d\}$. It is easy to check that $\mathcal{A}$ is a σ-algebra containing all d-boxes. And recall that the collection of all d-boxes generates $\mathcal{B}(\mathbb{R}^d)$. This proves (i). Next, statement (ii) is a consequence of the translation invariance property of the Lebesgue outer measure, and (iii) follows from (i) and (ii) by applying the principle mentioned above.

We will give other characterizations of the Lebesgue measurable sets shortly, and we will also show that $\mathcal{L}(\mathbb{R}^d) \neq \mathbb{P}(\mathbb{R}^d)$.

Self Assessment

Fill in the blanks:

1. .... is developed through approximations of a finite nature (e.g.: one tries to approximate the area of a bounded subset of $\mathbb{R}^2$ by the sum of the areas of finitely many rectangles).
2. While Riemann's theory is restricted to the Euclidean space, the ideas involved in .... are applicable to more general spaces, yielding an abstract measure theory.

Notes

3. The .... $\mu^*_{j,d}[Y]$ of a bounded subset $Y \subset \mathbb{R}^d$ is defined as $\mu^*_{j,d}[Y] = \inf\{\sum_{n=1}^{k} Vol_d(A_n) : k \in \mathbb{N}$, and A_n's are d-boxes with $Y \subset \bigcup_{n=1}^{k} A_n\}$.

4. The construction above is due to Vitali, and hence the set Y is called a ....

5. Let X, Y be metric spaces and let f: X → Y be a function. Then the oscillation ω (f, x) of f at a point x ∈ X is defined as $\omega(f, x) = \lim_{\delta \to 0+} \text{diam}\,[f(B(x, \delta))]$. Clearly, f is .... at x iff ω (f, x) = 0.

6. Let $-\infty \le a < b \le \infty$ and let f: (a, b) → ℝ be a .... Then, Y = {x ∈ (a, b) : f is discontinuous at x} is a countable set (possibly empty).

7. Let X be a metric space. Then the σ-algebra on X generated by the collection of all open subsets of X is called the .... on X, and is denoted as $\mathcal{B}(X)$ (or just $\mathcal{B}$, if X is clear from the context).

8. If $A \subset L_\Omega$ is a .... set, then A has a least upper bound in L_Ω. [*Proof:* If $B = \bigcup_{\beta \in A} L_\beta$, then B is countable and hence $L_\Omega \setminus B \ne \emptyset$. The least element of $L_\Omega \setminus B$ is the least upper bound of A].

13.5 Summary

- Measure theory helps us to assign numbers to certain sets and functions – to a measurable set we may assign its measure, and to an integrable function we may assign the value of its integral. Lebesgue integration theory is a generalization and completion of Riemann integration theory. In Lebesgue's theory, we can assign numbers to more sets and more functions than what is possible in Riemann's theory. If we are asked to distinguish between Riemann integration theory and Lebesgue integration theory by pointing out an essential feature, the answer is perhaps the following.

  (i) We say $Y \subset \mathbb{R}^d$ is a discrete subset of $\mathbb{R}^d$ if for each y ∈ Y, there is an open set $U \subset \mathbb{R}^d$ such that U ∩ Y = {y}. For example, {1/n: n ∈ ℕ} is a discrete subset of ℝ.

  (ii) A subset $Y \subset \mathbb{R}^d$ is nowhere dense in $\mathbb{R}^d$ if $\text{int}[\overline{Y}] = \emptyset$, or equivalently if for any nonempty open set $U \subset \mathbb{R}^d$, there is a nonempty open set V ⊂ U such that V ∩ Y = 0. For example, if f: ℝ → ℝ is a continuous map, then its graph G(f) := {(x, f(x)):x ∈ ℝ} is nowhere dense in $\mathbb{R}^2$ (∵ G(f) is closed and does not contain any open disc).

  (iii) A subset $Y \subset \mathbb{R}^d$ is of first category in $\mathbb{R}^d$ if Y can be written as a countable union of nowhere dense subsets of $\mathbb{R}^d$; otherwise, Y is said to be of second category in $\mathbb{R}^d$. For example, Y = ℚ × ℝ is of first category in $\mathbb{R}^2$ since Y can be written as the countable union $Y = \bigcup_{r \in \mathbb{Q}} Y_r$, where Y_r := {r} × ℝ is nowhere dense in $\mathbb{R}^2$.

  (iv) (The following definition can be extended by considering ordinal numbers, but we consider only non-negative integers). For $Y \subset \mathbb{R}^d$ and integer n ≥ 0, define the nth derived set of Y inductively as $Y^{(0)} = Y$, $Y^{(n+1)}$ = {limit points of $Y^{(n)}$ in $\mathbb{R}^d$}. We say $Y \subset \mathbb{R}^d$ has derived length n if $Y^{(n)} \ne \emptyset$ and $Y^{(n+1)} \ne \emptyset$; and we say Y has infinite derived length if $Y^{(n)} \ne \emptyset$ for every integer n ≥ 0. For example, ℚ has infinite derived length (since $\overline{\mathbb{Q}} = \mathbb{R}$), and {(1/m,1/n): m,n ∈ ℕ} has derived length 2.

  (v) We say $A \subset \mathbb{R}^d$ is a d-box if $A = \prod_{j=1}^{d} I_j$, where I_j's are bounded intervals. The d-dimensional volume of a d-box A is $Vol_d(A) = \prod_{j=1}^{d} |I_j|$. For example, $Vol_3([1, 4) \times [0,1/2] \times (-1,3]) = 6$.

  (vi) The d-dimensional Jordan outer content $\mu^*_{j,d}[Y]$ of a bounded subset $Y \subset \mathbb{R}^d$ is defined as $\mu^*_{j,d}[Y] = \inf\{\sum_{n=1}^{k} Vol_d(A_n) : k \in \mathbb{N}$, and A_n's are d-boxes with $Y \subset \bigcup_{n=1}^{k} A_n\}$.

Notes

(vii) The d-dimensional Lebesgue outer measure $\mu^*_{L,d}[Y]$ of an arbitrary set $Y \subset \mathbb{R}^d$ is defined as $\mu^*_{L,d}[Y] = \inf\{\sum_{n=1}^{\infty} \text{Vol}_d(A_n) : A_n\text{'s are d-boxes with } Y \subset \bigcup_{n=1}^{\infty} A_n\}$.

- If $f : [a, b] \to \mathbb{R}$ is a function and $P = \{a_0 = a \le a_1 \le \ldots \le a_{n-1} \le a_n = b\}$ is a partition of $[a, b]$, let $V_a^b(f, P) = \sum_{i=1}^{n} |f(a_i) - f(a_{i-1})|$. Define the total variation of f as $V_a^b(f) = \sup\{V_a^b(f, P) : P$ is a partition of $[a, b]\}$. We say f is of bounded variation if $V_a^b(f) < \infty$. It is easy to see that if f is of bounded variation, then f is bounded ($\because$ if $x \in [a, b]$, take $P = \{a \le x \le b\}$ to see that $|f(x) - f(a)| \le V_a^b(f)$).

13.6 Keywords

Riemann Integration Theory: Riemann integration theory $\mapsto$ finiteness.

Lebesgue Integration Theory: Lebesgue integration theory $\mapsto$ countable infiniteness.

Baire Category Theorem: Let (X, ρ) be a complete metric space and let $U_n \subset X$ be open and dense in X for $n \in \mathbb{N}$. Then, $\bigcap_{n=1}^{\infty} U_n$ is also dense in X. In particular, $\bigcap_{n=1}^{\infty} U_n \neq \emptyset$.

Lebesgue's Differentiation Theorem: Let $-\infty \le a < b \le \infty$, let $f : (a, b) \to \mathbb{R}$ be a monotone function and let $Y = \{x \in (a, b) : f \text{ is not differentiable at } x\}$. Then $\mu^*_{L,1}[Y] = 0$.

Borel●algebra: If X is a separable metric space, then any base or subbase for the topology of X will generate the Borel σ-algebra $\mathcal{B}(X)$.

Well-ordering Principle: Well-ordering principle (equivalent to the axiom of choice): Any non-empty set admits a well-ordering.

13.7 Review Questions

1. If $f, g : [a, b] \to \mathbb{R}$ are of bounded variation, then fg is of bounded variation. [*Hint:* Let $M > 0$ be such that $|f|, |g| \le M$. Now, subtracting and adding the term $f(a_i)g(a_{i-1})$, note that $|(fg)(a_i) - (fg)(a_{i-1})| \le |f(a_i)||g(a_i) - g(a_{i-1})| + |f(a_i) - f(a_{i-1})||g(a_{i-1})|$ and hence $V_a^b(fg) \le M(V_a^b(f) + V_a^b(g))$.]

2. If $f : [a, b] \to \mathbb{R}$ is a function and $c \in [a, b]$, then $V_a^b(f) = V_a^c(f) + V_c^b(f)$. [*Hint:* If P_1 is a partition of $[a, c]$ and P_2 is a partition of $[c, b]$, then $V_a^c(f, P_1) + V_c^b(f, P_2) = V_a^b(f, P_1 \cup P_2) \le V_a^b(f)$. Conversely, if P is a partition of $[a, b]$, first refine it by inserting c and then divide into partitions P_1 of $[a, c]$ and P_2 of $[c, b]$. Check that $V_a^b(f, P) \le V_a^c(f, P_1) + V_c^b(f, P_2) \le V_a^c(f) + V_c^b(f)$.]

3. Let $f: [a, b] \to \mathbb{R}$ be a bounded function. If f is either monotone or of bounded variation, then f is Riemann integrable.

4. If $f, g; [a,b] \to \mathbb{R}$ are Riemann integrable, then $h := \max\{f, g\}$ is also Riemann integrable. [*Hint:* The set of discontinuities of h is contained in $\{x : f \text{ is not continuous at } x\} \cup \{x : g \text{ is not continuous at } x\}$.]

5. If $\mathcal{A}$ is a σ-algebra on a set X show that

   (i) $A \setminus B, A \Delta B \in \mathcal{A}$ if $A, B \in \mathcal{A}$,

   (ii) $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$ if $A_1, A_2, \ldots \in \mathcal{A}$.

6. Let $\mathcal{A} = \{A \subset \mathbb{R}^d : A$ is a countable (possibly finite or empty) union of d-boxes$\}$. Is $\mathcal{A}$ a σ-algebra on $\mathbb{R}^d$? [*Hint:* Let d = 1. Consider $\mathbb{Q}$ and $\mathbb{R}\setminus\mathbb{Q}$, or the middle-third Cantor set and its complement.]

7. Are the following σ-algebras on $\mathbb{R}^d$: $\mathcal{A}_1 = \{A \subset \mathbb{R}^d : A \text{ or } \mathbb{R}^d \setminus A \text{ is open in } \mathbb{R}^d\}$ and $\mathcal{A}_2 = \{A \subset \mathbb{R}^d : A \text{ or } \mathbb{R}^d \setminus A \text{ is dense in } \mathbb{R}^d\}$?

Notes

8. If A_α's are σ-algebras on a set X, then $\bigcap_\alpha \mathcal{A}_\alpha := \{A \subset X : A \subset \mathcal{A}_\alpha \text{ for every } \alpha\}$ is also a σ-algebra on X.

9. Show that $\mathcal{B}(\mathbb{R})$ is generated by each of the following collections: $\{(a, \infty) : a \in \mathbb{R}\}$, $\{[a, \infty) : a \in \mathbb{R}\}$, $\{(-\infty, b) : b \in \mathbb{R}\}$, $\{(-\infty, b] : b \in \mathbb{R}\}$, $\{(a, b) : a < b \text{ and } a, b \in \mathbb{Q}\}$.

10. (i) If $\text{card}(X) \le \text{card}(\mathbb{R})$, then $\text{card}(X^{\mathbb{N}}) \le \text{card}(\mathbb{R})$. (ii) If $\text{card}(J) \le \text{card}(\mathbb{R})$ and $\text{card}(X_\beta) \le \text{card}(\mathbb{R})$ for each $\beta \in J$, then, $\text{card}(\bigcup_{\beta \in J} X_\beta) \le \text{card}(\mathbb{R})$. [*Hint:* (i) Assume X = (0,1). Define a one-one map $f : (0, 1)^{\mathbb{N}} \to (0,1)$ as follows. If $x = (x_n) \in (0, 1)^{\mathbb{N}}$ and if $x_n = 0.x_{n,1}x_{n,2}\cdots$, then $f(x) = 0.x_{1,1}x_{1,2}x_{2,1}x_{1,3}x_{2,2}x_{3,1}\cdots$. (ii) Let $g: \mathbb{R} \to J$ and $h_\beta : \mathbb{R} \to X_\beta$ be surjections. Then $f: \mathbb{R}^2 \to \bigcup_{\beta \in J} X_\beta$ defined as $f(x, y) = h_{g(y)}(x)$ is a surjection, and $\text{card}(\mathbb{R}^2) = \text{card}(\mathbb{R})$.]

Answers: Self Assessment

1. Riemann integration theory
2. Lebesgue's theory
3. d-dimensional Jordan outer content
4. Vitali set
5. continuous
6. monotone function
7. Borel σ-algebra
8. non-empty countable

13.8 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 14: The Integral of a Non-negative Function

Notes

CONTENTS

Objectives

After studying this unit, you will be able to:

- Discuss the integral of a non-negative function
- Explain Properties of the integral of non-negative functions
- Describe Monotone convergence theorem and
- Definition of Integrable function over a measurable set

Introduction

In this unit we are going to study about the definition and the properties of the integral of non-negative functions and some important theorems.

14.1 Integration of Non-negative Measurable Functions

We integrate non-negative measurable functions through approximation by bounded measurable functions vanishing outside a set of finite measure, which we studied earlier.

Definition: For a non-negative measurable function $f : E \to [0, \infty]$ (where E is a set which may be of finite or infinite measure), we define

$$\int_A f = \sup\left\{\int_A \varphi : \varphi \le f \text{ on } A, \varphi \in B_0(E)\right\}$$

for any $A \subseteq E$.

Note that for non-negative bounded measurable functions vanishing outside a set of finite measure, this definition agrees with the old one. Also note that we allow the functions to take infinite value here.

We verify the monotonicity and linearity of such integrals.

Proposition: Suppose $f, g : E \to [0, \infty]$ are non-negative measurable and $A \subseteq E$.

(a) If $f \le g$ a.e. on A then $\int_A f \le \int_A g$.

(b) For $\alpha > 0$, $f + g$ and αf are non-negative measurable functions too and

Notes

$$\int_A (f+g) = \int_A f + \int_A g$$

$$\int_A \alpha f = \alpha \int_A f$$

Proof:

(a) This is clearly true, for if $\varphi \in B_0(E)$ and $\varphi \le f$ on A, then $\varphi \le g$ on A so $\int_A \varphi \le \int_A g$ by definition of $\int_A g$. Taking supremum over all such φ's, we get $\int_A f \le \int_A g$.

(b) The assertion on $\int_A \alpha f$ can be proved using supremum arguments similar to that in (a) by noting that for $\alpha > 0$ and $\varphi \in B_0(E)$, $\varphi/\alpha \le f$ on A whenever $\varphi \le \alpha f$ on A, and $\alpha\varphi \le \alpha f$ on A whenever $\varphi \le f$ on A.

To verify $\int_A (f+g) = \int_A f + \int_A g$, note that if φ, $\not\varphi \in B_0(E)$ and $\varphi \le f$, $\not\varphi \le g$ on A, then $\varphi + \not\varphi \in B_0(E)$ and $\varphi + \not\varphi \le f + g$ on A so

$$\int_A (f+g) \ge \int_A (\varphi + \not\varphi) \quad \text{(by definition of } \int_A (f+g))$$

$$= \int_A \varphi + \int_A \not\varphi$$

take supremum over all such φ's and $\not\varphi$'s we have $\int_A (f+g) \ge \int_A f + \int_A g$. For the opposite inequality, note that if $\phi \in B_0(E)$ with $\phi \le f + g$ on A, then write $\varphi = \min\{\phi, f\}$ and $\not\varphi = \phi - \varphi$ we see that φ, $\not\varphi \in B_0(E)$ (note (i) $-M \le \varphi \le \phi \le M$ if $|\phi| \le M$ so φ is bounded on E; (ii) $\not\varphi = \phi - \varphi$ is bounded on E because both ϕ and φ are; (iii) measurability of φ, $\not\varphi$ is clear; and (iv) from $\varphi = \min\{\phi, f\}$ and $\not\varphi = \max\{0, \phi - f\}$ we see that φ, $\not\varphi = 0$ whenever $\phi = 0$ so φ, $\not\varphi$ vanishes outside a set of finite measure). Further, we have $\varphi \le f$, $\not\varphi \le g$ on A. Hence

$$\int_A \phi = \int_A \varphi + \int_A \not\varphi$$

$$\le \int_A f + \int_A g$$

Taking supremum over all such ϕ's we get $\int_A (f+g) \le \int_A f + g$

Theorem 1: Fatou's Lemma

Suppose $\{f_n\}$ is a sequence of non-negative measurable functions defined on E and $\{f_n\}$ converges (pointwisely) to a non-negative function f a.e. on E. Then

$$\int_E f \le \liminf_{n\to\infty} \int_E f_n$$

Proof: Let $h \in B_0(E)$ and $h \le f$ on E. Then there exists $A \subseteq E$ with $m(A) < \infty$ such that $h = 0$ outside A. Let $h_n = \min\{f_n, h\}$ on A, we have h_n is uniformly bounded and measurable on A : in fact if $|h| \le M$ on E, then $h_n = \min\{f_n, h\} > \min\{0, h\} \ge -M$ and $h_n = \min\{f_n, h\} \le h \le M$ so $|h_n| \le M$ on A Further, with the observation that $\min\{a, b\} = (a + b - |a - b|)/2$ for all real a, b we have

$$h_n = \frac{f_n + h - |f_n - h|}{2} \to \frac{f + h - |f - h|}{2} = \min\{f, h\} = h$$

on A. Since $m(A) < \infty$, we can conclude by Bounded Convergence Theorem that $\int_A h = \lim_{n\to\infty} \int_A h_n$.

So assuming $h_n = 0$ on $E\backslash A$, we have

$$\int_E h = \int_A h \lim_{n\to\infty} \int_A h_n = \lim_{n\to\infty} \int_E h_n \le \liminf_{n\to\infty} \int_E f_n$$

where the first equality follows from h = 0 on E/A and the last line $h_n \le f_n$ on E for all n. Taking supremum over all such h's, we get the desired inequality.

Theorem 2: Monotone Convergence Theorem

If $\{f_n\}$ is an increasing sequence of non-negative measurable functions defined on E (increasing in the sense that $f_n \le f_{n+1}$ for all n on E) and $f_n \to f$ a.e. on E, then

$$\int_E f_n \uparrow \int_E f$$

by which it means $\{\int_E f_n\}$ is an increasing sequence with limit $\int_E f$.

In symbol,

$$0 \le fn \uparrow f \text{ a.e. on } E \Rightarrow \int_E f_n \uparrow \int_E f$$

Proof:

$$\int_E f \le \liminf_{n\to\infty} \int_E f_n \le \limsup_{n\to\infty} \int_E f_n \le \int_E f,$$

the first inequality follows from Fatou's Lemma, the last inequality follows from $f_n \le f$ on E for all n. Hence $\int_E f_n \uparrow \int_E f$. (That $\int_E f_n$ increases as n increases is immediate from monotonicity of such integrals.)

Corollary: Extension of Fatou's lemma

If $\{f_n\}$ is a sequence of non-negative measurable functions on E, then $\int_E \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int_E f_n$.

Proposition: Suppose f is a non-negative measurable function defined on E such that $\int_E f < \infty$. Then for all ε > 0, there is a δ > 0 such that

$$\int_E f < \varepsilon$$

whenever $A \subseteq E$ with m(A) < δ.

Proof: The result clearly holds if f is bounded on E. Suppose now f is not necessarily bounded, we see that $(f \wedge n) \uparrow f$ so by Monotone Convergence Theorem

$$\int_A f = \lim_{n\to\infty} \int_A (f \wedge n)$$

for all $A \subseteq E$. Note that by assumption $\int_E f < \infty$ so both sides of the equality above are finite. Hence if ε > 0 is given, then there is a N such that $\left|\int_A f - \int_A (f \wedge N)\right| < \varepsilon$.

Take δ = ε/2N, we see that

$$\int_A f \le \left|\int_A f - \int_A f(f \wedge N)\right| + \int_A (f \wedge N) \le \varepsilon/2 + Nm(A) \le \varepsilon/2 + N\delta < \varepsilon$$

whenever $A \subseteq E$ with m(A) < δ. So we are done.

14.2 Extended Real-valued Integrable Functions

Here we integrated non-negative measurable functions, and we wish to drop the non-negative requirement. Recall that it is a natural requirement that our integral be linear, and now we can integrate a general non-negative measurable function, so it is tempting to define the integral of a general (not necessarily non-negative) measurable function f to be $\int f^+ - \int f^-$ where f+ = f V0 and f^- = (–f) V0, since f^+, f^- are non-negative measurable and they sum up to f. But it turns out that we cannot always do that, because it may well happen that $\int f^+$ and $\int f^-$ are both infinite, in which case their difference would be meaningless. (Remember that ∞ – ∞ is undefined.) So we need to restrict ourselves to a smaller class of functions than the collection of all measurable functions when we drop the non-negative requirement and come to the following definition.

Definition: For f : E → [–∞, ∞], denote f+ = f V0 and f⁻ = (–f) V0. Then f is said to be integrable if and only if both $\int_E f^+$ and $\int_E f^-$ are finite, in which case we define the integral of f by

$$\int_A f = \int_A f^+ - \int_A f^-$$

for any $A \subseteq E$

Notation: We shall denote the class of all (extended real-valued) integrable functions defined on E by C(E).

Note that in the above definition, f^+ and f^- are both non-negative measurable, so for any set $A \subseteq E$, $\int_A f^+$ and $\int_A f^-$ are both defined. Furthermore, $\int_A f^+ \le \int_E f^+ < \infty$ and similarly $\int_A f^- < \infty$ so their difference makes sense now. Also note that for non-negative integrable functions this definition agrees with our old one.

We provide an alternative characterization of integrable functions.

Proposition: A measurable function f defined on E is integrable if and only if $\int_E |f| < \infty$ so.

Proof: Just note that $|f| = f^+ + f^-$.

We proceed to investigate the structure of $\mathcal{L}$(E). We want to say it is a vector lattice. But we have to be careful here: Given f, g ∈ $\mathcal{L}$(E) it may well happen that f(x) = + ∞ and g(x) = –∞ for some x ∈ E and then f + g cannot be defined by f(x) + g(x) at that x. Luckily there cannot be too many such x's, in the sense that the set of all such x's is of measure zero. In fact every integrable function is finite. We know that the values of a function on a set of measure zero are not important as far as integration is concerned. (This was observed as in the case of bounded measurable functions vanishing outside a set of finite measure; the reader should verify this for the case of general integrable functions as well.) So that eliminates our previous worries: more precisely, let us agree from now on two functions f,g: E → [–∞, ∞] are said to be equal (write f = g) if and only if they take the same values a.e.on E, and f + g shall mean a function whose value at x is equal to f(x) + g(x) for a.e.x ∈ E. Also say f ≤ g if and only if f(x) ≤ g(x) for a.e. x ∈ E. Then we have the following proposition.

Proposition: $\mathcal{L}$(E) forms a vector lattice (partially ordered by ≤).

Proof: If f,g ∈ $\mathcal{L}$(E), then $\int_E |f+g| \le \int_E |f| + \int_E |g| < \infty$ (we are using linearity and monotonicity and hence f + g ∈ $\mathcal{L}$(E) (the measurability of f + g is previously known). The rest of the proposition is trivial.

With the vector lattice structure of $\mathcal{L}$(E) it is natural to ask whether the integral is linear and monotone or not. We expect it to be true; we verify it below.

Proposition: For any f,g ∈ $\mathcal{L}$(E) and $A \subseteq E$, we have $\int_A (f+g) = \int_A f + \int_A g$ and $\int_A \alpha f = \alpha \int_A f$. Furthermore, if f ≤ g a.e. on A then $\int_A f \le \int_A g$.

Proof: The parts for monotonicity and $\int_A \alpha f = \alpha \int_A f$ are easy and left as an exercise.

So now let f,g ∈ $\mathcal{L}$(E) and $A \subseteq E$ be given, and we prove $\int_A (f+g) = \int_A f + \int_A g$. By definition of the integral, the LHS is just $\int_A (f+g)^+ - \int_A (f+g)^-$ and the RHS is $\int_A f^+ - \int_A f^- + \int_A g^+ \int_A g^-$, all terms being finite. So it suffices to show

(6) $\int_A (f+g)^+ + \int_A f^- + \int_A g^- = \int_A (f+g)^- + \int_A f^+ + \int_A g^+$,

which will be true if we can show

(7) $(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+$

a.e. on A because we can then use linearity of Section 3 to conclude that (6) is true. But (7) is clearly true a.e., because $(f + g)^+ - (f + g)^- = f + g = f^+ - f^- + g^+ - g^-$ a.e., all terms being finite a.e. This completes our proof.

Finally we prove the important Generalized Lebesgue Dominated Convergence Theorem.

Theorem 3: If $\{f_n\}$, $\{g_n\}$ are sequences of measurable functions defined on E, $|f_n| \le g_n$, $f = \lim_{n\to\infty} f_n$, $g = \liminf_{n\to\infty} g_n$ and $\lim_{n\to\infty} \int_E g_n = \int_E g < \infty$, then $\lim_{n\to\infty} \int_E f_n$ exists and is equal to $\int_E f$.

Proof: Since $|f_n| \le g_n$ implies $g_n \pm f_n$ are non-negative measurable, we see that

$$\int_E g + \int_E f = \int_E \liminf_{n\to\infty} (g_n + f_n) \le \liminf_{n\to\infty} \int_E (g_n + f_n) = \int_E g + \liminf_{n\to\infty} \int_E f_n$$

and similarly

$$\int_E g - \int_E f = \int_E \liminf_{n\to\infty} (g_n - f_n) \le \liminf_{n\to\infty} \int_E (g_n - f_n) = \int_E g + \liminf_{n\to\infty} \int_E f_n$$

So $\int_E f \le \liminf_{n\to\infty} \int_E f_n \le \limsup_{n\to\infty} \int_E f$ (note here we used the assumption that $\int_E g < \infty$) and the desired conclusion follows.

Corollary: Lebesgue Dominated Convergence Theorem

Suppose a sequence of measurable functions $\{f_n\}$ defined on E converges pointwisely a.e. on E to f. If $|f_n| \le g$ on E for some integrable function g, then $\int_E f_n$ converges to $\int_E f$.

A final word of remark: The idea of this section extends readily to complex-valued functions, and the readers who are familar with general measure theory should find that the results in the whole unit is valid on a general measure space without needing the slightest modification.

Self Assessment

Fill in the blanks:

1. For a non-negative measurable function $f : E \to [0, \infty]$ (where E is a set which may be of finite or infinite measure), we define ....
2. For non-negative .... vanishing outside a set of finite measure, this definition agrees with the old one. Also note that we allow the functions to take infinite value here.
3. Suppose $\{f_n\}$ is a sequence of non-negative measurable functions defined on E and $\{f_n\}$ converges (pointwisely) to a .... f a.e. on E. Then $\int_E f \le \liminf_{n\to\infty} \int_E f_n$.
4. If $\{f_n\}$ is an .... of non-negative measurable functions defined on E (increasing in the sense that $f_n \le f_{n+1}$ for all n on E) and $f_n \to f$ a.e. on E, then $\int_E f_n \uparrow \int_E f$ by which it means $\{\int_E f_n\}$ is an increasing sequence with limit $\int_E f$.
5. A .... f defined on E is integrable if and only if $\int_E |f| < \infty$ so.
6. For any $f,g \in \mathcal{L}(E)$ and $A \subseteq E$, we have $\int_A (f+g) = \int_A f + \int_A g$ and $\int_A \alpha f = \alpha \int_A f$. Furthermore, if $f \le g$ a.e. on A then ....

7. If $\{f_n\}$, $\{g_n\}$ are sequences of measurable functions defined on E, $|f_n| \leq g_n$, $f = \lim_{n\to\infty} f_n$, $g = \liminf_{n\to\infty} g_n$ and ...., then $\lim_{n\to\infty} \int_E f_n$ exists and is equal to $\int_E f$.

8. Suppose a sequence of measurable functions $\{f_n\}$ defined on E converges pointwisely a.e. on E to f. If $|f_n| \leq g$ on E for some integrable function g, then $\int_E f_n$ .... $\int_E f$.

14.3 Summary

- For a non-negative measurable function $f : E \to [0, \infty]$ (where E is a set which may be of finite or infinite measure), we define

  $$\int_A f = \sup\left\{\int_A \varphi : \varphi \leq f \text{ on } A,\ \varphi \in B_0(E)\right\}$$

  for any $A \subseteq E$.

  Note that for non-negative bounded measurable functions vanishing outside a set of finite measure, this definition agrees with the old one. Also note that we allow the functions to take infinite value here.

- Suppose $\{f_n\}$ is a sequence of non-negative measurable functions defined on E and $\{f_n\}$ converges (pointwisely) to a non-negative function f a.e. on E. Then

  $$\int_E f \leq \liminf_{n\to\infty} \int_E f_n$$

- If $\{f_n\}$ is an increasing sequence of non-negative measurable functions defined on E (increasing in the sense that $f_n \leq f_{n+1}$ for all n on E) and $f_n \to f$ a.e. on E, then

  $$\int_E f_n \uparrow \int_E f$$

  by which it means $\{\int_E f_n\}$ is an increasing sequence with limit $\int_E f$.

- If $\{f_n\}$ is a sequence of non-negative measurable functions on E, then $\int_E \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_E f_n$. The proof is easy and left as an exercise.

  The following proposition is concerned with the absolute continuity of the integral.

- Suppose f is a non-negative measurable function defined on E such that $\int_E f < \infty$. Then for all $\varepsilon > 0$, there is a $\delta > 0$ such that

  $$\int_E f < \varepsilon$$

  whenever $A \subseteq E$ with $m(A) < \delta$.

- Suppose a sequence of measurable functions $\{f_n\}$ defined on E converges pointwisely a.e. on E to f. If $|f_n| \leq g$ on E for some integrable function g, then $\int_E f_n$ converges to $\int_E f$.

14.4 Keywords

Fatou's Lemma: Suppose $\{f_n\}$ is a sequence of non-negative measurable functions defined on E and $\{f_n\}$ converges (pointwisely) to a non-negative function f a.e. on E. Then $\int_E f \leq \liminf_{n\to\infty} \int_E f_n$.

Monotone Convergence Theorem: If $\{f_n\}$ is an increasing sequence of non-negative measurable functions defined on E (increasing in the sense that $f_n \leq f_{n+1}$ for all n on E) and $f_n \to f$ a.e. on E, then $\int_E f_n \uparrow \int_E f$ by which it means $\{\int_E f_n\}$ is an increasing sequence with limit $\int_E f$.

Lebsegue Dominated Convergence Theorem: Suppose a sequence of measurable functions $\{f_n\}$ defined on E converges pointwisely a.e. on E to f. If $|f_n| \le g$ on E for some integrable function g, then $\int_E f_n$ converges to $\int_E f$.

14.5 Review Questions

1. For a non-negative measurable function f defined on E, show that $\int_A f = \int_E f \chi_A$ for any A Í E. Also show that $\int_A f \le \int_B f$ if $A \subseteq B \subseteq E$.
2. Show that if A, B C E are disjoint and f is a non-negative measurable function defined on E, then $\int_{A \cup B} f = \int_A f + \int_B f$.
3. Show that if f is a non-negative measurable function defined on E and $\int_E f = 0$, then f = 0 a.e. on E.
4. Show that if f is a non-negative measurable function defined on E and $\int_E f < \infty$, then f is finite a.e.
5. Show that w may have strict inequality in Fatou's Lemma.

   (***Hint:*** Consider the sequence {fn} defined by fn(x) = 1 if n x < n + 1,with fn(x) = 0 otherwise.)
6. Show that the monotone convergence theorem need not hold for decreasing sequence of functions.

   (***Hint:*** Let fn(x) = 0, if x < n, fn(x) = 1 for xn.)
7. Show that if f and g are measurable and y $|f| \le |g|$ a.e., and if g is integrable, then prove that f is intergrable.

Answers: Self Assessment

1.	$\int_A f = \sup\left\{\int_A \varphi : \varphi \le f \text{ on } A, \varphi \in B_0(E)\right\}$	2.	bounded measurable functions
3.	non-negative function	4.	increasing sequence
5.	measurable function	6.	$\int_A f \le \int_A g$
7.	$\lim_{n \to \infty} \int_E g_n = \int_E g < \infty$	8.	converges to

14.6 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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

Unit 15: The General Lebesgue Integral and Convergence in Measure

CONTENTS

Objectives

After studying this unit, you will be able to:

- Explain the General Lebesgue integral of a measurable function
- Discuss the Properties of Lebesgue integral
- Discuss Lebesgue convergence theorem
- Explain Generalization of Lebesgue convergence theorem
- Describe convergence in measure of a sequence of measurable functions

Introduction

In this unit, you are going to study about the general Lebesgue integral, some of its properties, convergence in measure and theorems related to them.

15.1 The General Lebesgue Integral

Definition: The positive part of a function f is $f^+ = f \vee 0$ i.e $f^+(x) = \max\{f(x), 0\}$

The negative part of a function is $f^- = f \wedge 0$. i.e $f^-(x) = \min\{f(x), 0\}$

Hence $f = f^+ - f^-$.

And $|f| = f^+ + f^-$

Definition: A measurable function f is said to be integrable over E if f^+ and f^- are both integrable over E.

Then the integral of f is defined as

$$\int_E f = \int_E f^+ - \int_E f^-$$

Theorem 1: Let f and g are integrable over E. Then

(i) The function cf is integrable over E, and $\int_E cf = c\int_E f$.

(ii) The function f + g is integrable over E, and $\int_E f + g = \int_E f + \int_E g$.

(iii) If f g a.e., then $\int_E f \;\; \int_E g$

(iv) If A and B are disjoint measurable sets contained in E, then $\int_{A\cup B} f = \int_A f + \int_B f$

Proof:

(i) Since f is integrable over E, both f^+ and f^- are integrable over E and the integral of f is given by

$$\int_E f = \int_E f^+ - \int_E f^-$$

Hence,

both cf^+ and cf^- are integrable over E, and hence, $cf = cf^+ - cf^-$ are integrable over E and

$$\begin{aligned}\int_E cf &= \int_E cf^+ - \int_E cf^- \\ &= c\int_E f^+ - c\int_E f^- \\ &= c[\int_E f^+ - \int_E f^-] \\ &= c\int_E f.\end{aligned}$$

Hence (i) is proved.

(ii) Suppose if f_1 and f_2 are nonnegative integrable functions with $f = f_1 - f_2$,

Then $f^+ - f^- = f_1 - f_2$.

Hence,

$$f^+ + f_2 = f^- + f_1.$$

As you know

$$f^+ + f_2 = f^- - f_1.$$

Therefore,

$$f = f^+ - f^-$$

$$= f_1 - f_2.$$

Since f and g are measurable,

f^+, f^-, g^+, g^- are measurable.

Hence,

$f^+ + g^+, f^- + g^-$ are also measurable.

And $f + g = (f^+ + g^+) - (f^- + g^-)$.

Hence by(1),

$$(f + g) = (f^+ + g^+) - (f^- + g^-)$$

$= f^+ + g^+ - f^- - g^-$

$= (f^+ - f^-) + (g^+ - g^-)$

$= f + g.$

Hence (ii) is proved.

(iii) Since f g a.e., f^+ - f^- $g^+ - g^-$ a.e.,

Hence, $f^+ + g - g^+ + f^-$ a.e,

$(f^+ + g^-)$ $(g^+ + f^-)$.

Hence

$f^+ + g - g^+ + f^-$.

Hence,

$f^+ - f - g^+ - g^-$

Hence,

f g.

Hence (iii) is proved.

(iv) Consider

$$\begin{aligned}\int_{A\cup B} f &= \int f\cdot\chi_{A\cup B}\\ &= \int f\cdot(\chi_A + \chi_B)\\ &= \int f\cdot\chi_A + \int f\cdot\chi_B\\ &= \int_A f + \int_B f\end{aligned}$$

15.2 Lebesgue Convergence Theorem

Theorem 2: Let g be integrable over E and let $\{f_n\}$ be a sequence of measurable functions such that $|f_n|$ g on E and for almost all x in E we have f(x) = lim fn(x). Then

$$\int_E f = \lim \int_E f_n$$

Proof: Since $|f_n|$ g on E, g – f_n is nonnegative and hence by Fatou's Lemma,

$$\int_E (g - f)\underline{\lim}\int_E (g - f_n) \qquad ...(1)$$

Since f(x) = lim f_n(x) a.e. on E and

$|f_n|$ g on E,

$|f|$ g on E.

Hence since g is integrable,

f is also integrable.

$$\int_E (g - f) = \int_E g - \int_E f \qquad ...(2)$$

Also,

$$\underline{\lim}\int_E (g - f_n) = \int_E g - \overline{\lim}\int_E f_n \qquad ...(3)$$

Substituting (2) and (3) in (1), we get

$$\int_E g - \int_E f \quad \int_E g - \overline{\lim} \int_E f_n$$

Hence

$$\int_E f \quad \overline{\lim} \int_E f_n \qquad \text{...(4)}$$

Similarly by considering g + f_n, we get

$$\int_E f \quad \underline{\lim} \int_E f_n \qquad \text{...(5)}$$

From (4) and (5), we get

$$\overline{\lim} \int_E f_n \quad \int_E f \quad \underline{\lim} \int_E f_n \qquad \text{...(6)}$$

But it is always true that

$$\underline{\lim} \int_E f_n \quad \overline{\lim} \int_E f_n \qquad \text{...(7)}$$

From (6) and (7)

$$\int_E f = \lim \int_E f_n .$$

Hence the theorem.

> *Notes* If we replace g by g_n's, we get the following generalization of the Lebesgue Convergence theorem.

Theorem 3: Let $\{g_n\}$ be a sequence of integrable functions which converges a.e to an integrable function g. Let $\{f_n\}$ be a sequence of measurable functions such that $|f_n| \; g_n$ and $\{f_n\}$ converges to f a.e.

If $\int g = \lim \int g_n$,

then $\int f = \lim \int f_n$.

15.3 Convergence in Measure

Definition: A sequence $\{f_n\}$ of measurable functions is said to converge to f in measure if, given $\varepsilon > 0$, there is an N such that for all n N we have

$$m\{x / |f(x) - f_n(x)| \; \varepsilon\} < \varepsilon.$$

Remark: From this definition and littlewood's third principle, it is clear that,

If $\{f_n\}$ is a sequence of measurable functions defined on a measurable set E of finite measure and $f_n > f$ a.e, then $\{f_n\}$ converges to f in measure.

Example: Construct the sequence $\{f_n\}$ as follows:

Let $\quad n = k + 2^v, \; 0 \; k < 2^v$, and

Set $\quad f_n(x) = 1 \;$ if $x \in [k2^{-v}, (k + 1)\, 2^{-v}]$

Notes

And $f_n(x) = 0$ otherwise.

Then $m\{x / |fn(x)| > \varepsilon\} = 2^{-\nu}\ \ 2/n$ [since $2^{\nu}\ \ n < 2^{\nu} + 1$]

Hence $f_n > 0$ in measure.

> *Notes* That the sequence $\{f_n(x)\}$ has the value 1 for arbitrarily large values of n.
>
> Hence $\{f_n(x)\}$ does not converge for any x in [0, 1].

Theorem 4: Let $\{f_n\}$ be a sequence of measurable functions that converges in measure to f.

Then there is a subsequence $\{f_{nk}\}$ that converges to f almost everywhere.

Proof: Since $\{f_n\}$ is a sequence of measurable functions that converges in measure to f,

Given ν, there is an integer n_ν such that for all n nν,

$$m\{x / |f(x) - f_n(x)|\ \ 2^{-\nu}\} < 2^{-\nu} \qquad ...(1)$$

Let $E_\nu = \{x / |f_{n\nu}(x) - f(x)|\ \ 2^{-\nu}\}$

Therefore,

if $x \notin \bigcup_{\upsilon=k}^{\infty} E_\upsilon$

then $|f_{n\nu}(x) - f(x)| < 2^{-\nu}$ for νk.

Therefore,

$F_{n\nu}(x) > f(x)$.

Hence $f_{n\nu}(x) > f(x)$ for any $x \notin A \bigcap_{k=1}^{\infty} \bigcup_{\upsilon=k}^{\infty} E_\upsilon$

But $mA\ \ m\left[\bigcup_{\upsilon=k}^{\infty} E_\upsilon\right]$

$\sum_{\upsilon=k}^{\infty} mE_\upsilon$

$= 2^{-k+1}$.

Hence $mA = 0$

Theorem 5: Let $\{f_n\}$ be a sequence of measurable functions defined on a measurable set E of finite measure.

Then $\{f_n\}$ converges to f in measure if and only if every subsequence of $\{f_n\}$ has in turn a subsequence that converges almost everywhere to f.

Theorem 6: Fatou's lemma and the monotone and Lebesgue Convergence theorem remain valid if 'convergence a.e.' is replaced by 'convergence in measure'.

Self Assessment

Notes

Fill in the blanks:

1. A .... f is said to be integrable over E if f^+ and f^- are both integrable over E.
2. Let g be integrable over E and let $\{f_n\}$ be a sequence of measurable functions such that $|f_n|$ g on E and for almost all x in E we have …………………….
3. Let $\{g_n\}$ be a sequence of ……………….. which converges a.e to an integrable function g.
4. A sequence $\{f_n\}$ of measurable functions is said to ………………......….. in measure if, given $\varepsilon > 0$, there is an N such that for all nN we have $m\{x/ |f(x) - fn(x)| \varepsilon\} < \varepsilon$.
5. Let $\{f_n\}$ be a sequence of measurable functions that converges in measure to f. Then there is a subsequence {nk f} that .... to f almost everywhere.

15.4 Summary

- Definition of General Lebesgue integral of a measurable function
- Properties of Lebesgue integral
- Lebesgue convergence theorem
- Generalization of Lebesgue convergence theorem
- Definition of convergence in measure of a sequence of measurable functions and
- Every sequence of measurable sequence that converges in measure contains a subsequence that converges almost everywhere.

15.5 Keywords

Convergence in Measure: A sequence $\{f_n\}$ of measurable functions is said to converge to f in measure if, given $\varepsilon > 0$, there is an N such that for all n N we have $m\{x/ |f(x) - f_n(x)| \varepsilon\} < \varepsilon$.

Lebesgue Convergence Theorem: Let g be integrable over E and let $\{f_n\}$ be a sequence of measurable functions such that $|f_n|$ g on E and for almost all x in E we have $f(x) = \lim f_n(x)$. Then

$$\int_E f = \lim \int_E f_n.$$

15.6 Review Questions

1. Show that if f is integrable over E, then so is $|f|$ and $\left|\int_E f\right| \le \int_E |f|$. Does the integrability of $|f|$ imply that of f?.
2. Let $\{f_n\}$ be a sequence of integrable functions such that $f_n > f$ a.e with f integrable. Then $\int |f_n - f| \to 0$ if and only if $\int |f_n| \to \int |f|$.
3. Show that if f is integrable over E, then $|f|$ is also integrable over E. further $\left|\int_E f\right| \le \int_E |f|$ is the converse true?

Answers: Self Assessment

1. measurable function
2. f(x) = lim fn(x).
3. integrable functions
4. converge to f
5. converges

15.7 Further Readings

Books

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

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

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

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

S.C. Malik: Mathematical Analysis.

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