Real Analysis

M.A. (Previous)

Directorate of Distance Education
Maharshi Dayanand University
ROHTAK – 124 001

Maharshi Dayanand University
ROHTAK – 124 001

Developed & Produced by EXCEL BOOKS PVT. LTD., A-45 Naraina, Phase 1, New Delhi-110028

Contents

CONTENTS

1

SEQUENCES AND SERIES OF FUNCTIONS

1.1. The object of this chapter is to consider sequences whose terms are functions rather than real numbers. There sequences are useful in obtaining approximations to a given function. We shall study two different notations of convergence for a sequence of functions: Pointwise convergence and uniform convergence

Pointwise and Uniform Convergence of Sequences of functions

Definition. Let $A \subseteq \mathbf{R}$ and suppose that for each $n\in\mathbf{N}$ there is a function $f_n : A\to\mathbf{R}$. Then $<f_n>$ is called a **sequence of functions** on A. For each $x \in A$, this sequence gives rise to a sequence of real numbers, namely the sequence $< f_n(x) >$.

Definition. Let $A \subseteq \mathbf{R}$ and let $< f_n >$ be a sequence of functions on A. Let $A_0 \subseteq A$ and suppose $f : A_0 \to \mathbf{R}$. Then the sequence $<f_n>$ is said to converge on A_0 to f if for each $x\in A_0$, the sequence $<f_n(x)>$ converges to $f(x)$ in **R.**

In such a case f is called the limit function on A_0 of the sequence $<f_n>$.

When such a function f exists, we say that the sequence $<f_n>$ is convergent on A_0 or that $<f_n>$ c**onverges pointwise on** A_0 to f and we write $f(x) = \lim\limits_{n\to\infty} f_n(x)$, $x\in A_0$. Similarly, it $\Sigma f_n(x)$ converges for every $x\in A_0$, and if

$$f(x) = \sum_{n=1}^{\infty} f_n(x), \quad x\in A_0,$$

the function f is called the sum of the series Σf_n.

The question arises : If each function of a sequence $<f_n>$ has certain property, such as continuity, differentiability or integrality, then to what extent is this property transferred to the limit function? For example, if each function f_n is continuous at a point x_0, is the limit function f also continuous at x_0? In general, it is not true. Thus pointwise convergence is not so strong concept which transfers above mentioned property to the limit function. Therefore some stronger methods of convergence are needed. One of these method is the notion of uniform convergence. We know that f_n is continuous at x_0 if

$$\lim_{x\to x_0} f_n(x) = f_n(x_0)$$

On the other hand, f is continuous at x_0 if

(1.1.1) $$\lim_{x\to x_0} f(x) = f(x_0)$$

But (1.1.1) can be written as

(1.1.2) $$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \lim_{x\to x_0} f_n(x)$$

Thus our question of continuity reduces to "can we interchange the limit symbols in (1.1.2)?" or "Is the order in which limit processes are carried out immaterial". The following examples show that the limit symbols cannot in general be interchanged.

Example. A sequence of continuous functions whose limit function is discontinuous: Let

$$f_n(x) = \frac{x^{2n}}{1+x^{2n}}, \; x\in\mathbf{R}, \; n = 1, 2,\ldots$$

we note that

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} 0 & \text{if } |x|<1 \\ \frac{1}{2} & \text{if } |x|=1 \\ 1 & \text{if } |x|>1 \end{cases}$$

Each f_n is continuous on **R** but the limit function f is discontinuous at x = 1 and x = −1.

Example. A double sequence in which limit process cannot be interchanged. For m = 1, 2,…,n = 1, 2, 3,…, let us consider the double sequence

$$S_{mn} = \frac{m}{m+n}$$

For every fixed n, we have

$$\lim_{n\to\infty} S_{mn} = 1$$

and so

$$\lim_{n\to\infty} \lim_{m\to\infty} S_{mn} = 1$$

On the other hand, for ever fixed m, we have

$$\lim_{n\to\infty} S_{mn} = \lim_{n\to\infty} \frac{1}{1+\frac{n}{m}} = 0$$

and so

$$\lim_{m\to\infty} \lim_{n\to\infty} S_{mn} = 0$$

Hence
$$\lim_{n\to\infty} \lim_{m\to\infty} S_{mn} \neq \lim_{m\to\infty} \lim_{n\to\infty} S_{mn}$$

Example. A sequence of functions for which limit of the integral is not equal to integral of the limit: Let

$$f_n(x) = n^2 x\,(1-x)^n, \; x\in\mathbf{R}, \; n = 1, 2, \ldots.$$

If $0 \le x \le 1$, then

$$f(x) = \lim_{n\to\infty} f_n(x) = 0$$

and so

$$\int_0^1 f(x)\,dx = 0$$

But
$$\int_0^1 f_n(x)\,dx = n^2 \int_0^1 x(1-x)^n\,dx$$

$$= \frac{n^2}{n+1} - \frac{n^2}{n+2}$$

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

and so

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = 1$$

Hence

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,dx \neq \int_0^1 (\lim_{n\to\infty} f_n(x))\,dx.$$

Example. A sequence of differentiable functions $[f_n]$ with limit 0 for which $[f_n']$ diverges: Let

$$f_n(x) = \frac{\sin nx}{\sqrt{n}} \text{ if } x\in\mathbf{R},\ n = 1, 2,$$

Then $$\lim_{n\to\infty} f_n(x) = 0\ \forall\ x.$$

But $$f_n'(x) = \sqrt{n} \cos nx$$

and so $$\lim_{n\to\infty} f_n'(x) \text{ does not exist for any } x.$$

Definition. A sequence of functions $\{f_n\}$ is said to converge uniformly to a function f on a set E if for every $\in > 0$ there exists an integer N (depending only on $\in$) such that $n > N$ implies

(1.1.3) $$|f_n(x) - f(x)| < \in \text{ for all } x\varepsilon E.$$

If each term of the sequence $<f_n>$ is real-valued, then the expression (1.1.3) can be written as

$$f(x) - \in < f_n(x) < f(x) + \in$$

for all $n > N$ and for all $x\in E$. This shows that the entire graph of f_n lies between a "band" of height $2\in$ situated symmetrically about the graph of f.

Definition. A series $\Sigma f_n(x)$ is said to converge uniformly on E if the sequence $\{S_n\}$ of partial sums defined by

$$S_n(x) = \sum_{i=1}^{n} f_i(x)$$

converges uniformly on E.

Examples. (1) Consider the sequence $<S_n>$ defined by $S_n(x) = \frac{1}{x+n}$ in any interval [a, b], a > 0. Then

$$S(x) = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \frac{1}{x+n} = 0$$

For convergence we must have

(1.1.4) $$|S_n(x) - S(x)| < \in,\quad n > n_0$$

or $$\left|\frac{1}{x+n} - 0\right| < \in,\ n > n_0$$

or $$\frac{1}{x+n} < \in$$

or $$x + n > \frac{1}{\in}$$

or $$n > \frac{1}{\in} - x$$

If we select n_0 as integer next higher to $\frac{1}{\in}$, then (1.1.4) is satisfied for m(integer) greater than $\frac{1}{\in}$ which does not depend on $x\in[a, b]$. Hence the sequence $<S_n>$ is uniformly convergent to S(x) in [a, b].

2. Consider the sequence $<f_n>$ defined by

$$f_n(x) = \frac{x}{1+nx}, \ x \geq 0$$

Then

$$f(x) = \lim_{n\to\infty} \frac{x}{1+nx} = 0 \text{ for all } x \geq 0.$$

Then $<f_n>$ converges pointwise to 0 for all $x \geq 0$. Let $\in >0$, then for convergence we must have

$$|f_n(x) - f(x)| < \in, \ n > n_0$$

or $$\left| \frac{x}{1+nx} - 0 \right| < \in, \ n > n_0$$

$$\frac{x}{1+nx} < \in$$

$$x < \in + nx \in$$

or $$n x \in > x - \in$$

or $$n > \frac{x - \in}{x \in}$$

or $$n > \frac{x}{x \in} = \frac{1}{\in}$$

If n_0 is taken as integer greater than $\frac{1}{\in}$, then

$$|f_n(x) - f(x)) < \in \ \forall \ n > n_0 \text{ and } \forall \ x \in [0, \infty)$$

Hence $<f_n>$ converges uniformly to f on $[0, \infty)$.

3. Consider the sequence $<f_n>$ defined by

$$f_n(x) = x^n, \ 0 \leq x \leq 1.$$

Then

$$f(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x = 1 \end{cases}$$

Let $\in > 0$ be given. Then for convergence we must have

$$|f_n(x) - f(x)| < \in, \ n > n_0$$

or $$x^n < \in$$

or $$\left(\frac{1}{x}\right)^n > \frac{1}{\in}$$

or $$n > \frac{\log \frac{1}{\in}}{\log \frac{1}{x}}$$

Thus we should take n_0 to be an integer next higher to log $1/\in \Big/ \log\frac{1}{x}$. If we take x = 1, then m does not exist. Thus the sequence in question is not uniformly convergent to f in the interval which contains 1.

4. Consider the sequence $< f_n >$ defined by

$$f_n(x) = \frac{nx}{1+n^2x^2}, \; 0 \le x \le a.$$

Then if x = 0, then $f_n(x) = 0$

and so $f(x) = \lim_{n\to\infty} f_n(x) = 0$

If $x \neq 0$, then

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{nx}{1+n^2x^2} = 0$$

Thus f is continuous at x = 0. For convergence we must have

$$|f_n(x) - f(x)| < \in, \; n > n_0$$

or $$\frac{nx}{1+n^2x^2} < \in$$

or $$1 + n^2 x^2 - \frac{nx}{\in} > 0$$

or $$nx > \frac{1}{2\in} + \frac{1}{2}\sqrt{\frac{1}{\in^2} - 4}$$

Thus we can find an upper bound for n in any interval $0 < a \le x \le b$, but the upper bound is infinite if the interval includes 0. Hence the given sequence in non-uniformly convergent in any interval which includes the origin. So 0 is the point of non-uniform convergence for this sequence.

5. Consider the sequence $<f_n>$ defined by

$$f_n(x) = \tan^{-1} nx, \; 0 \le x \le a$$

Then

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} \frac{\pi}{2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

This the function f is discontinuous at x = 0.

For convergence, we must have for $\in > 0$,

$$|f_n(x) - f(x)| < \in, \; n > n_0$$

or $\pi/2 - \tan^{-1} nx < \in$

or $\cot^{-1} nx < \in$ $\quad \because \tan^{-1} nx + \cot^{-1} nx = \pi/2$

or $$nx > \frac{1}{\tan \in}$$

or $$n > \frac{1}{\tan \in}\left(\frac{1}{x}\right)$$

Thus no upper bound can be found for the function on the right if 0 is an end point of the interval. Hence the convergence is non-uniform in any interval which includes 0. So, here 0 is the point of non-uniform convergence.

Definition. A sequence $\{f_n\}$ is said to be uniformly bounded on E if there exists a constant M > 0 such that $|f_n(x)| \le M$ for all x in E and all n.

The number M is called a uniform bound for $\{f_n\}$.

For example, the sequence $<f_n>$ defined by

$$f_n(x) = \sin nx \text{ , } x \,\varepsilon\, \mathbf{R}$$

is uniformly bounded. Infact,

$$|f_n(x)| = |\sin nx| \le 1 \text{ for all } x \,\varepsilon\, \mathbf{R} \text{ and all } n \in N.$$

If each individual function is bounded and it $f_n \to f$ uniformly on E, than it can be shown that $\{f_n\}$ is uniformly bounded on E. This result generally helps us to conclude that a sequence is not uniformly convergent.

We now find necessary and sufficient condition for uniform convergence of a sequence of functions.

Theorem 1. (Cauchy criterion for uniform convergence). The sequence of functions $\{f_n\}$, defined on E, converges uniformly if and only if for every $\in >0$ there exists an integer N such that $m \ge N$, $n \ge N$, $x \in E$ imply

$$|f_n(x) - f_m(x)| < \in$$

Proof. Suppose first that $<f_n>$ converges uniformly on E to f. Then to each $\in > 0$ there exists an integer N such that n > N implies

$$|f_n(x) - f(x)| < \in/2 \text{ for all } x \,\varepsilon\, E$$

Similarly for m > N implies

$$|f_m(x) - f(x)| < \in/2 \text{ for all } x \,\varepsilon\, E$$

Hence, for n > N, m > N, we have

$$\begin{aligned} |f_n(x) - f_m(x)| &= |f_n(x) - f(x) + f(x) - f_m(x)| \\ &\le |f_n(x) - f(x)| + |f_m(x) - f(x)| \\ &< \in/2 + \in/2 = \in \text{ for all } x \in E \end{aligned}$$

Hence the condition is necessary.

Conversely, suppose that Cauchy condition is satisfied, that is,

(1.1.5) $$|f_n(x) - f_m(x)| < \in, \qquad n, m > N \text{ and } x \,\varepsilon\, E.$$

This implies that $<f_n(x)>$ is a Cauchy sequence of real numbers and so is convergent. Let $f(x) = \lim\limits_{n\to\infty} f_n(x)$, $x \varepsilon E$. We shall show that $f_n \to f$ uniformly on E. Let $\in >0$ be given. We can choose N such that (1.1.5) is satisfied. Fix n, and let $m\to\infty$ in (1.1.5). Since $f_m(x)\to f(x)$ as $m\to\infty$, this yields

$$|f_n(x) - f(x)| < \in, \; n > N, \; x \in E$$

Hence $f_n \to f$ uniformly on E.

1.2. Tests for Uniform Convergence.

Theorem 2. Suppose $\lim\limits_{n\to\infty} f_n(x) = f(x)$, x ε E and let $M_n = \operatorname{lub}\limits_{x\in E} |f_n(x) - f(x)|$. Then $f_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n\to\infty$. (This result is known as **M_n – Test for uniform convergence**)

Proof. We have

$$\operatorname{lub}_{x\in E} |f_n(x) - f(x)| = M_n \to 0 \text{ as } n\to\infty.$$

Hence
$$\lim_{n\to\infty} |f_n(x) - f(x)| = 0 \ \forall\ x \in E$$

Hence to each $\in>0$, there exists an integer N such that $n>N$, x ε E imply

$$|f_n(x) - f(x)| < \in$$

Hence $f_n \to f$ uniformly on E.

Weierstrass contributed a very convenient test for he uniform convergence of infinite series of functions.

Theorem 3. (Weierstrass M-test). Let $<f_n>$ be a sequence of functions defined on E and suppose

$$|f_n(x)| \le M_n \ (x\varepsilon E, n = 1, 2, 3,\ldots),$$

where M_n is independent of x. Then Σf_n converges uniformly as well as absolutely on E if ΣM_n converges.

Proof. Absolute convergence follows immediately from comparison test.

To prove uniform convergence, we note that

$$|S_m(x) - S_n(x)| = \left|\sum_{i=1}^{m} f_n(x) - \sum_{i=1}^{m} f_n\right|$$

$$= |f_{n+1}(x) + f_{n+2}(x) + \ldots + f_m(x)|$$

$$\le M_{n+1} + M_{n+2} + \ldots + M_m$$

But since ΣM_n is convergent, given $\in > 0$, there exists N (independent of x) such that

$$|M_{n+1} + M_{n+2} + \ldots + M_n| < \in, \ n > N$$

Hence

$$|S_m(x) - S_n(x)| < \in, \ n > N, \ x \in E$$

and so $\Sigma f_n(x)$ converges uniformly by Cauchy criterion for uniform convergence.

Lemma (Abel's Lemma). If $v_1, v_2,\ldots, v_n$ be positive and decreasing, the sum

$$u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

lies between $A v_1$ and $B v_1$, where A and B are the greatest and least of the quantities

$$u_1, u_1 + u_2, u_1 + u_2 + u_3, \ldots, u_1 + u_2 + \ldots + un$$

Proof. Write

$$S_n = u_1 + u_2 + \ldots + u_n$$

Therefore

$$u_1 = S_1,\ u_2 = S_2 - S_1, \ldots, u_n = S_n - S_{n-1}$$

Hence

$$\sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_1 v_2 + \ldots + u_n v_n$$

$$= S_1 v_1 + (S_2 - S_1) v_2 + (S_3 - S_2) v_3 + \ldots + (S_n - S_{n-1}) v_n$$

$$= S_1 (v_1 - v_2) + S_2(v_2 - v_3) + \ldots + S_{n-1}(v_{n-1} - v_n) + S_n v_n$$
$$= A [v_1 - v_2 + v_2 - v_3 + \ldots + v_{n-1} - v_n + v_n]$$
$$= A v_1$$

Similarly, we can show that

$$\sum_{i=1}^{n} u_i v_i > B v_1$$

Hence the result follows.

Theorem. 4. (Abel's Test) The series $\sum_{n=1}^{n} u_n(x) v_n(x)$ converges uniformly on E if

(i) $\{v_n(x)\}$ is a positive decreasing sequence for all values of $x \varepsilon E$

(ii) $\Sigma u_n(x)$ is uniformly convergent

(iii) $v_1(x)$ is bounded for all $x\varepsilon E$, i.e., $v_1(x) < M$.

Proof. Consider the series $\Sigma u_n(x) v_n(x)$, where $\{v_n(x)\}$ is a positive decreasing sequence for each $x\varepsilon E$. By Abel's Lemma

$$|u_n(x) v_n(x) + u_{n+1}(x) v_{n+1}(x) + \ldots + u_m(x) u_m(x)| < Av_n(x)$$

where A is greatest of the magnitudes

$$|u_n(x)| , |u_n(x) + u_{n+1}(x)|, \ldots, |u_n(x) + u_{n+1}(x) + \ldots + u_m(x)|$$

clearly A is a function of x.

Since $\Sigma u_n(x)$ is uniformly convergent, it follows that

$$|u_n(x) + u_{n+1}(x) + \ldots + u_m(x)| < \frac{\in}{M} \text{ for all } n > N, x \in E$$

and so $A < \frac{\in}{M}$ for all $n > N$ (independent of x) and for all $x \varepsilon E$. Also, since $\{v_n(x)\}$ in decreasing, $v_n(x) < v_1(x) < M$ since $v_1(x)$ is bounded for all $x \varepsilon E$

Hence

$$| u_n(x) v_n(x) + u_{n+1}(x) v_{n+1}(x) + \ldots + u_m(x) v_m(x)| < \in$$

for $n > N$ and all $x \in E$ and so $\sum_{n=1}^{n} u_n(x) v_n(x)$ is uniformly convergent.

Theorem. (Dirichlet's test). The series $\sum_{n=1}^{n} u_n(x) v_n(x)$ converges uniformly on E if

(i) $\{v_n(x)\}$ is a positive decreasing sequence for all values of $x \in E$, which tends to zero uniformly on E

(ii) $\Sigma u_n(x)$ oscillates or converges in such a way that the moduli of its limits of oscillation remains less than a fixed number M for all $x \varepsilon E$.

Proof. Consider the series $\sum_{n=1}^{\infty} u_n(x) v_n(x)$, where $\{v_n(x)\}$ is a positive decreasing sequence tending to zero uniformly on E. By Abel's Lemma

$$|u_n(x) v_n(x) + u_{n+1}(x) v_{n+1}(x) + \ldots + u_m(x) v_m (x) | < A v_n(x),$$

where A is greatest of the magnitudes

$$|u_n(x)|, |u_n(x) + v_{n+1}(x)|, \ldots, |u_n(x) + u_{n+1}(x) + \ldots + u_m(x)|$$

and is a function of x.

Since $\Sigma\, u_n(x)$ converges or oscillates finitely in such a way that $\left|\sum\limits_{r}^{s} u_n(x)\right| < M$ for all $x \varepsilon E$, therefore. A is less than M. Furthermore since $v_n(x)\to 0$ uniformly as $n\to\infty$, to each $\in>0$ there exists an integer N such that

$$v_n(x) < \frac{\in}{M} \text{ for all } n > N \text{ and all } x\varepsilon E$$

Hence

$$|u_n x)\, u_n(x) + v_{n+1}(x)\, v_{n+1}(x) + \ldots + u_m(x)\, v_m(x)| < \frac{\in}{M} \,.\, M = \in$$

for all $n > N$ and $x\varepsilon E$ and so $\sum\limits_{n=1}^{\infty} u_n(x)\, v_n(x)$ is uniformly convergent on E.

Examples. 1. Consider the series $\sum\limits_{n=1}^{\infty} \frac{\cos n\theta}{n^p}$. We observe that

$$\left|\frac{\cos n\theta}{n^p}\right| \le \frac{1}{n^p}$$

Also, we know that

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

is convergent if $p > 1$. Hence, by Weierstrass M-Test, the series $\Sigma \frac{\cos n\theta}{n^p}$ converges absolutely and uniformly for all real values of θ if $p > 1$.

Similarly, the series $\sum\limits_{n=1}^{\infty} \frac{\sin n\theta}{n^p}$ converges absolutely and uniformly by Weierstrass's M-Test.

2. Taking $M_n = r^n$, $0 < r < 1$, it can be shown by Weierstrass's M-Test that the series

$\Sigma\, r^n \cos n\theta$, $\Sigma\, r^n \sin n\theta$, $\Sigma\, r^n \cos^2 n\theta$, $\Sigma\, r^n \sin^2 n\theta$ converge uniformly and absolutely

3. Consider $\sum\limits_{n=1}^{\infty} \frac{x}{n(1+nx^2)}$, $x\in\mathbf{R}$.

We assume that x is +ve, for if x is negative, we can change signs of all the terms. We have

$$f_n(x) = \frac{x}{n(1+nx^2)}$$

and $$f_n'(x) = 0$$

implies $nx^2 = 1$. Thus maximum value of $f_n(x)$ is $\frac{1}{2n^{3/2}}$

Hence $\qquad f_n(x) \le \dfrac{1}{2n^{3/2}}$

Since $\Sigma \dfrac{1}{n^{3/2}}$ is convergent, Weierstarss' M-Test implies that $\sum_{n=1}^{\infty} \dfrac{x}{n(1+nx^2)}$ is uniformly convergent for all x ε **R**.

4. Consider the series $\sum_{n=1}^{\infty} \dfrac{x}{(n+x^2)^2}$, x ε **R**. We have

$$f_n'(x) = \frac{x}{(n+x^2)^2}$$

and so
$$f_n(x) = \frac{(n+x^2)^2 - 2x(n+x^2)2x}{(n+x^2)^4}$$

Thus $f_n'(x) = 0$ gives

$$x^4 + x^2 + 2nx^2 - 4nx^2 - 4x^4 = 0$$

$$-3x^4 - 2nx^2 + n^2 = 0$$

or
$$3x^4 + 2nx^2 - n^{2+} = 0$$

or
$$x^2 = \frac{n}{3} \text{ or } x = \sqrt{\frac{n}{3}}$$

Also $f_n''(x)$ is –ve. Hence maximum value of $f_n(x)$ is $\dfrac{3\sqrt{3}}{16n^2}$. Since $\Sigma \dfrac{1}{n^2}$ is convergent, it follows by Weierstrass's M-Test that the given series is uniformly convergent.

5. The series $\qquad \sum_{n=1}^{\infty} \dfrac{a_n x^n}{1+x^{2n}}$ and $\sum_{n=1}^{\infty} \dfrac{a_n x^{2n}}{1+x^{2n}}$

converge uniformly for all real values of x is Σa_n is absolutely convergent. The solutions follow the same line as for example 4.

6. Consider the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^p} \cdot \frac{x^{2n}}{1+x^{2n}}$$

We note that if p > 1, then $\Sigma \dfrac{(-1)^n}{n^p}$ is absolutely convergent and is independent of x. Hence, by Weierstrass's M-Test, the given series is uniformly convergent for all xε**R**.

If $0 \le p \le 1$, the series $\Sigma \dfrac{(-1)^n}{n^p}$ is convergent but not absolutely. Let

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

Then $\langle v_n(x) \rangle$ is monotonically decreasing sequence for |x| , 1 because

$$v_n(x) - v_{n+1}(x) = \frac{x^{2n}}{1+x^{2n}} - \frac{x^{2n+2}}{1+x^{2n+2}}$$

$$= \frac{x^{2n}(1-x^2)}{(1+x^{2n})(1+x^{2n+2})} \qquad \text{(+ve)}$$

Also $\qquad v_1(x) = \dfrac{x^2}{1+x^2} < 1.$

Hence, by Abel's Test, the series $\sum_{n=1}^{\infty} \dfrac{(-1)^n}{n^p} \cdot \dfrac{x^{2n}}{1+x^{2n}}$ in uniformly convergent for $0 < p \le 1$ and $|x| < 1$.

6. Consider the series

$$\Sigma\, a_n \,.\, \frac{x^n}{1+x^{2n}},$$

under the condition that Σa_n is convergent. Let

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

Then

$$\frac{v_n(x)}{v_{n+1}(x)} = \frac{1+x^{2n+2}}{x(1+x^{2n})}$$

and so

$$\frac{v_n(x)}{v_{n+1}(x)} - 1 = \frac{(1-x)(1-x^{2n\times 2})}{x(1+x^{2n})}$$

which is positive if $0 < x < 1$. Hence

$$v_n > v_{n+1}$$

and so $\langle v_n(x) \rangle$ is monotonically decreasing and positive. Also $v_1(x) = \dfrac{x}{1+x^2}$ is bounded. Hence, by Abel's test, the series $\Sigma\, a_n. \dfrac{x^n}{1+x^{2n}}$ is uniformly convergent in (0, 1) if $\Sigma\, a_n$ is convergent.

7. Consider the series $\Sigma\, a_n \dfrac{nx^{n-1}(1-x)}{1-x^n}$ under the condition that $\Sigma\, a_n$ is convergent. We have

$$v_n(x) = \frac{nx^{n-1}(1-x)}{1-x^n}$$

Then

$$\frac{v_n(x)}{v_{n+1}(x)} = \frac{n}{(n+1)x} \cdot \frac{1-x^{n+1}}{1-x^n}$$

Since $\frac{n}{n+1} \to 0$ as $n\to\infty$, taking n sufficient large

$$\frac{v_n(x)}{v_{n+1}(x)} > \frac{1-x^{n+1}}{1-x^n} > 1 \text{ if } 0 < x < 1.$$

Hence $< u_n(x)>$ is monotonically decreasing and positive. Hence, by Abel's Test, the given series converges uniformly in (0, 1).

1.3. Uniform Convergence and Continuity.

We know that if f and g are continuous functions, then f + g is also continuous and this result holds for the sum of finite number of functions. The question arises "Is the sum of infinite number of continuous function a continuous function?". The answer is not necessary. The aim of this section is to obtain sufficient condition for the sum function of an infinite series of continuous functions to be continuous.

Theorem. 6. Let $<f_n>$ be a sequence of continuous functions on a set $E \subseteq \mathbf{R}$ and suppose that $<f_n>$ converges uniformly on E to a function $f : E\to\mathbf{R}$. Then the limit function f is continuous.

Proof. Let $c\varepsilon E$ be an arbitrary point. If c is an isolated point of E, then f is automatically continuous at C. So suppose that c is an accumulation point of E. We shall show that f is continuous at c. Since $f_n\to f$ uniformly, for every $\in > 0$ there is an integer N such that $n \geq N$ implies

$$|f_n(x) - f(x)| < \frac{\in}{3} \text{ for all } x\in E.$$

Since f_M is continuous at c, there is a neighbourhood $S_\delta(c)$ such that $x \,\varepsilon\, S_\delta(c) \cap E$ (since c is limit point) implies

$$|f_M(x) - f_M(c)| < \in/3.$$

By triangle inequality, we have

$$\begin{aligned}|f(x) - f(c)| &= |f(x) - f_M(x) + f_M(x) - f_M(c) + f_M(c)- f(c)| \\ &\leq |f(x) - f_M(x)| + |f_M(x) - f_M(c)| + |f_M(c)- f(c)| \\ &< \frac{\in}{3}+\frac{\in}{3}+\frac{\in}{3} = \in\end{aligned}$$

Hence

$$|f(x) - f(c)| < \in, \; x \,\varepsilon\, S_\delta(\in) \cap E.$$

which proves the continuity of f at arbitrary point c ε E.

Remark. Uniform convergence of $<f_n>$ in the above theorem is sufficient but not necessary to transmit continuity from the individual terms to the limit function. For example, let $f_n : [0, 1] \to\mathbf{R}$ be defined for $n \geq 2$ by

$$f_n(x) = \begin{cases} n^2 x & \text{for } 0 \leq x \leq \frac{1}{n} \\ -n^2\left(x - \frac{2}{n}\right) & \text{for } \frac{1}{n} \leq x \leq \frac{2}{n} \\ 0 & \text{for } \frac{2}{n} \leq x \leq 1 \end{cases}$$

Each of the function f_n is continuous on [0, 1]. Also $f_n(x) \to 0$ as $n\to\infty$ for all x ε [0, 1]. Hence the limit function f vanishes identically and is continuous. But the convergence $f_n\to f$ is non-uniform.

The series version of Theorem 6 is the following:

Theorem. 7. If the series $\Sigma f_n(x)$ of continuous functions is uniformly convergent to a function f on [a, b] , then the sum function f is also continuous on [a, b].

Proof. Let $S_n(x) = \sum_{i=1}^{n} f_n(x)$, $n\in N$ and let $\in>0$. Since Σf_n converges uniformly to f on [a, b], there exists a positive integer N such that

(1.3.1.) $|S_n(x) - f(x)| < \frac{\in}{3}$ for all $n \geq N$ and $x \varepsilon [a, b]$.

Let c be any point of [a, b], then (1.3.1) implies

(1.3.2) $|S_n(c) - f(c)| < \frac{\in}{3}$ for all $n \geq N$.

Since f_n is continuous on [a, b] for each n, the partial sum

$$S_n(x) = f_1(x) + f_2(x) + \ldots + f_n(x)$$

is also continuous on [a, b] for all n. Hence to each $\in > 0$ then exist a $\delta > 0$ such that

(1.3.3) $|S_n(x) - S_n(c)| < \frac{\in}{3}$ whenever $|x-c| < \delta$

Now, by triangle inequality, and using (1.3.1), (1.3.2) and (1.3.3), we have

$$\begin{aligned} |f(x) - f(c)| &= |f(x) - S_n(x) + S_n(x) - S_n(c) + S_n(c) - f(c)| \\ &\leq |f(x) - S_n(x)| + |S_n(x) - S_n(c)| + |S_n(c) - f(c)| \\ &< \in/3 + \in/3 + \in/3 = \in, \text{ whenever } |x - c| < \delta \end{aligned}$$

Hence f is continuous at c. Since c is arbitrary point of [a, b], f is continuous on [a, b].

However, the converse of Theorem 6 in true with some additional condition on the sequence $<f_n>$ of continuous functions. The required result goes as follows.

Theorem. 8. Let E be compact and let $\{f_n]$ be a sequence of functions continuous on E which converges of a continuous function on E. If $f_n(x) \geq f_{n+1}(x)$ for n = 1, 2, 3,..., and for every $x \varepsilon E$, than $f_n \to f$ uniformly on E.

Proof. Take

$$g_n(x) = f_n(x)\ f(x).$$

Being the difference of two continuous functions, $g_n(x)$ is continuous. Also $g_n \to 0$ and $g_n \geq g_{n+1}$. We shall show that $g_n \to 0$ uniformly on E.

Let $\in > 0$ be given. Since $g_n \to 0$, there exists an integer $n \geq N_x$ such that

$$|g_n(x) - 0| < \in/2$$

In particular

$$|g_{N_x}(x) - 0| < \in/2$$

i.e. $0 \leq g_N(x) < \in/2$

The continuity and monotonicity of the sequence $\{g_n\}$ imply that there exists an open set J(x) containing x such that

$$0 \leq g_n(t) < \in$$

if $t \varepsilon J(x)$ and $x \geq N_x$.

Since E is compact, there exists a finite set of points $x_1, x_2, x_3 \ldots, x_m$ such that

$$E \subseteq J(x_1) \cup \ldots \cup J(x_m)$$

Taking

$$N = \max (Nx_1, Nx_2, \ldots, Nx_m)$$

it follows that

$$0 \leq g_n(t) \leq \in$$

for all $t \varepsilon E$ and $n \geq N$. Hence $g_n \to 0$ uniformly on E and so $f_n \to f$ uniformly on E.

1.4. Uniform convergence and Integrability.

We know that if f and g are integrable, then

$$\int (f + g) = \int f + \int g$$

and this result holds for the sum of a finite number of functions. The aim of this section is to find sufficient condition to extend this result to an infinite number of functions.

Theorem. 9. Let α be monotonically increasing on [a, b]. Suppose that each term of the sequence $\{f_n\}$ is a real valued function such that $f_n \varepsilon R(\alpha)$ on [a, b] for n = 1, 2, 3,… and suppose $f_n \to f$ uniformly on [a, b]. Then $f \varepsilon R(\alpha)$ on [a, b] and

$$\int_a^b f\, d\alpha = \lim_{n\to\infty} \int_a^b f_n\, d\alpha,$$

that is, $\int_a^b \lim_{n\to\infty} f_n(x)\, d\alpha(x) = \lim_{n\to\infty} \int_a^b f_n(x)\, d\alpha(x)$

(Thus limit and integral can be interchanged in this case. This property is generally described by saying that **a uniformly convergent sequence can be integrated term by term)**.

Proof. Let $\in$ be a positive number. Choose $\eta > 0$ such that

(1. 4. 1) $$\eta[\alpha(b) - \alpha(a)] \le \frac{\in}{3}$$

This is possible since α is monotonically increasing. Since $f_n \to f$ uniformly on [a, b], to each $\eta>0$ there exists an integer n such that

(1.4.2) $$|f_n(x) - f(x)| \le \eta,\ x\ \varepsilon\ [a, b]$$

Since $f_n\ \varepsilon\ R(\alpha)$, we choose a partition P of [a, b] such that

(1.4.3) $$U(P, f_n, \alpha) - L(P, f_n, \alpha) < \frac{\in}{3}$$

The expression (1.4.2) implies

$$f_n(x) - \eta \le f(x) \le f_n(x) + \eta$$

Now $f(x) \le f_n(x) + \eta$ implies, by (1.4.1) that

(1.4.4) $$U(P, f, \alpha)\ ,\ U(P, f_n, \alpha) + \frac{\in}{3}$$

Similarly, $f(x) \ge f_n(x) - \eta$ implies

(1.4.5) $$L(P, f, \alpha) \ge L(P, f_n, \alpha) - \frac{\in}{3}$$

Combining (1.4.3), (1.4.4) and (1.4.5), we get

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

Hence $f\ \varepsilon\ R(\alpha)$ on [a, b].

Further uniform convergence implies that to each $\in>0$, there exists an integer N such that for $n \ge N$

$$|f_n(x) - f(x)| < \frac{\in}{[\alpha(b)-\alpha(a)]},\ x \in [a, b]$$

Then for $n > N$,

$$\left|\int_a^b f d\alpha - \int_a^b f_n d\alpha\right| = \left|\int_a^b (f - f_n)\, d\alpha\right| \le \int_a^b |f - f_n|\, d\alpha$$

$$< \frac{\in}{[\alpha(b)-\alpha(a)]}\int_a^b d\alpha(x)\, dx$$

$$= \frac{\in[\alpha(b)-\alpha(a)]}{\alpha(b)-\alpha(a)}$$

$$= \in.$$

Hence

$$\int_a^b f\, d\alpha = \lim_{n\to\infty} \int f_n\, d\alpha$$

and the result follows.

The series version of Theorem 9 is :

Theorem. 10. Let $f_n\ \varepsilon\ \mathbf{R}$, n = 1, 2,… If Σf_n converges uniformly to f on [a, b], then $f\ \varepsilon\ \mathbf{R}$ and

$$\int_a^b f(x)\, d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n(x)\, d\alpha,$$

i.e. **the series Σf_n is integrable term by term.**

Proof. Let <S_n> denote the sequence of partial sums of Σf_n. Since Σf_n converges uniformly to f on [a, b], the sequence <S_n> converges uniformly to f. Then S_n being the sum of n integrable function is integrable for each n. Therefore, by theorem 9, f is also integrable in Riemann sense and

$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b S_n(x)\,dx$$

But

$$\int_a^b S_n(x)\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx + \ldots + \int_a^b f_m(x)\,dx$$

$$= \sum_{i=1}^{n} \int_a^b f_i(x)\,dx$$

Hence

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{\infty} \int_a^b f_i(x)\,dx$$

$$= \sum_{i=1}^{\infty} \int_a^b f_i(x)\,d\alpha,$$

and the proof of the theorem is complete.

Example. 1. Consider the sequence <f_n> for which $f_n(x) = nx\,e^{-nx^2}$, n ε N, x ε [0, 1]. We note that

$$f(x) = \lim_{n\to\infty} f_n(x)$$

$$= \lim_{n\to\infty} \frac{nx}{\frac{1+nx^2}{1} + \frac{n^2x^2}{\sqrt{2}} + \ldots} = 0, \quad x\ \varepsilon\ (0, 1]$$

Then

$$\int_0^1 f(dx) = 0$$

and

$$\int_0^1 f_n(x)\,dx = \int_0^1 nx\,e^{-nx^2}\,dx$$

$$= \frac{1}{2}\int_0^n e^{-t}\,dt\ ,\ t = nx^2$$

$$= \frac{1}{2}[1 - e^{-n}]$$

Therefore

$$\lim_{n\to\infty} \int f_n(x)\,dx = \lim_{n\to\infty} \frac{1}{2}[1-e^{-n}]$$

$$= \frac{1}{2}$$

If < f_n > were uniformly convergent, then $\int_0^1 f(x)\,dx$ should have been equal to $\lim_{n\to\infty} \int f_n 9x)\,dx$. But it is not the case. Hence the given sequence is not uniformly convergent to f infact, x = 0 is the point of non-uniform convergence.

2. Consider the series $\sum_{n=1}^{\infty} \frac{x}{(n+x^2)^2}$. This series is uniformly convergent and so in integrable term by term. Thus

$$\int_0^1 \left(\sum_{n=1}^{\infty} \frac{x}{(n+x^2)^2} \right) dx = \lim_{m\to\infty} \sum_{n=1}^{m} \int_0^1 \frac{x}{(n+x^2)^2}$$

$$= \lim_{m\to\infty} \sum_{n=1}^{m} \int_0^1 n(n + x^2)^{-2}\, dx$$

$$= \lim_{m\to\infty} \sum_{n=1}^{m} \left[\frac{(n+x^2)^{-1}}{-2} \right]_0^1$$

$$= \lim_{m\to\infty} \sum_{n=1}^{m} \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+1}\right)$$

$$= \lim_{m\to\infty} \frac{1}{2}\left[\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\ldots+\left(\frac{1}{m}-\frac{1}{m+1}\right)\right]$$

$$= \lim_{m\to\infty} \frac{1}{2}\left(1-\frac{1}{m+1}\right) = \frac{1}{2}$$

3. Consider the series $\sum_{n=1}^{\infty} \left[\frac{nx}{1+n^2x^2} - \frac{(n-1)x}{1+(n-1)^2x^2} \right]$, $a \le x \le 1$.

Let $S_n(x)$ denote the partial sum of the series. Then

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

and so $\qquad f(x) = \lim_{n\to\infty} S_n(x) = 0$ for all $x \in [0, 1]$

As we know that 0 is a point of non-uniform convergence of the sequence $<S_n(x)>$, the given series is not uniformly convergent on [0, 1]. But

$$\int_0^1 f(x)\, dx = \int_0^1 0\, dx = 0$$

and

$$\int_0^1 S_n(x)\, dx = \int_0^1 \frac{nx}{1+n^2x^2} dx$$

$$= \frac{1}{2n}\int_0^1 \frac{2n^2x}{1+n^2x^2} dx$$

$$= \frac{1}{2n} \left| \log(1+n^2x^2) \right|_0^1$$

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

Hence

$$\lim_{n\to\infty} \int_0^1 S_n(x) dx = \lim_{n\to\infty} \frac{1}{2n} \log(1+n^2) \qquad \left(\frac{\infty}{\infty} \text{form}\right)$$

$$= \lim \frac{n}{1+n^2} \qquad \left(\frac{\infty}{\infty} \text{ form}\right)$$

$$= \lim_{n\to\infty} \frac{1}{2n} = 0$$

Thus

$$\int_0^1 f(dx)\, dx = \lim_{n\to\infty} \int_0^1 S_n(x)\, dx,$$

and so the series is integrable term by term although 0 is a point of non-uniform convergence.

Theorem. 11. Let $\{g_n\}$ be a sequence of function of bounded variation on [a, b] such that $g_n(a) = 0$, and suppose that there is a function g such that

$$\lim_{n\to\infty} V(g - g_n) = 0$$

and g(a) = 0. Then for every continuous function f on [a, b], we have

$$\lim_{n\to\infty} \int_a^b f\,dg_n = \int_a^b f\,dg.$$

and $g_n \to g$ uniformly on [a, b].

Proof. If V denotes the total variation on [a, b], then

$$V(g) \le V(g_n) + V(g-g_n)$$

Since g_n is of bounded variation and $\lim_{n\to\infty} V(g - g_n) = 0$ it follows that total variation of g is finite and so g is of bounded variation on [a, b]. Thus the integrals in the assertion of the theorem exist. Suppose $|f(x)| \le M$ on [a, b]. Then

$$\left|\int_a^b f dg - \int_a^b f\,dg_n\right| = \left|\int_a^b f d\,(g-g_n)\right|$$
$$\le M\,V(g - g_n)$$

Since $V(g-g_n) \to 0$ as $n\to\infty$, it follows that

$$\int_\infty^b f dg = \lim_{n\to\infty} \int_\infty^b f\,dg_n$$

Furthermore,

$$|g(x) - g_n(x)| \le V(g-g_n), \qquad a \le x \le b$$

Therefore, as $n\to\infty$, we have

$$g_n \to f \text{ uniformly.}$$

1.5. Uniform Convergence and Differentiation

If f and g are derivable, then

$$\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx} f(x) + \frac{d}{dx} g(x)$$

and that this can be the extended to finite number of derivable function. In this section, we shall extend this phenomenon under some suitable condition to infinite number of functions.

Theorem. 12. Suppose $\{f_n\}$ is a sequence of functions, differentiable on [a, b] and such that $[f_n(x_0)\}$ converges for some point x_0 on [a, b]. If $\{f_n'\}$ converges uniformly on [a, b], then $\{f_n\}$ converges uniformly on [a, b], to a function f, and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad (a \le x \le b).$$

Proof. Let $\in>0$ be give. Choose N such that $n \ge N$, $m \ge N$ implies

(1.5.1) $$|f_n(x_0) - f_m(x_0)| < \frac{\in}{2}$$

and

(1.5.2) $$|f_n'(t) - f_m'(t)| < \frac{\in}{2(b-a)} \qquad (a \le t < b)$$

Application of Mean Value Theorem to the function $f_n - f_m$, (1.5.2) yields

(1.5.3) $$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \le \frac{|x-t|\in}{2(b-a)} \le \frac{\in}{2}$$

for any x and t on [a, b] if $n \ge N$, $m \ge$ **N**. Since

$$|f_x(x) - f_m(x)| \le f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)|$$
$$+ |f_n(x_0) - f_m(x_0)|,$$

the relation (1.5.1) and (1.5.3) imply for $n \ge N$, $m \ge N$,

$$|f_x(x) - f_m(x)| < \in/2 + \in/2 = \in \quad (a \le x \le b)$$

Hence, by Cauchy criterion for uniform convergence, it follows that $\{f_n\}$ converges uniformly on [a, b]. Let

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad (a \le x < b).$$

For a fixed point $x \in [a, b]$, let us define

$$(1.5.4) \qquad \phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}, \qquad \phi(t) = \frac{f(t) - f(x)}{t - x}$$

for $a \le t \le b$, $t \ne x$. Then

$$(1.5.5) \qquad \lim_{t\to x} \phi_n(t) = \lim_{t\to x} \frac{f_n(t) - f_n(x)}{t - x} = f_n'(x) \qquad (n = 1, 2, \ldots)$$

Further, (1.5.3) implies

$$|\phi_n(t) - \phi_m(t)| \le \frac{\in}{2(b-a)} \quad (n \ge N, m \ge N).$$

Hence $\{\phi_n\}$ converges uniformly for $t \ne x$. We have proved just now that $\{f_n\}$ converges to f uniformly on [a, b]. Therefore (1.5.4) implies that

$$(1.5.6) \qquad \lim_{n\to\infty} \phi_n(t) = \phi(t)$$

uniformly for $a \le t \le b$, $t \ne x$. Therefore using uniform convergence of $<\phi_n>$ and (1.5.5) we have

$$\begin{aligned}\lim_{t\to x} \phi(t) &= \lim_{t\to x} \lim_{n\to\infty} \phi_n(t)\\ &= \lim_{n\to\infty} \lim_{t\to x} \phi_n(t)\\ &= \lim_{n\to\infty} f_n'(x)\end{aligned}$$

But $\lim_{t\to x} \phi(t) = f'(x)$. Hence

$$f'(x) = \lim_{n\to\infty} f_n'(x).$$

Remark. If in addition to the above hypothesis, each f_n' is continuous, then the proof becomes simpler. Infact, we have then

Theorem 13. Let $<f_n>$ be a sequence of functions such that

(i) Each f_n is differentiable on [a, b]
(ii) Each f_n' is continuous on [a, b]
(iii) $<f_n>$ converges to f on [a, b]
(iv) $<f_n'>$ converges uniformly to g on [a, b], then f is differentiable and $f'(x) = g(x)$ for all $x \,\varepsilon\, [a, b]$.

Proof. Since each f_n' is continuous on [a, b] and $<f_n'>$ converges uniformly to g on [a, b], the application of theorem 6 of this chapter implies that g is continuous and hence Riemann-integrable. Therefore, Theorem 9 implies

$$\int_a^t g(x)\, dx = \lim_{n\to\infty} \int_a^t f_n'(x)\, dx$$

But, by Fundamental Theorem of Integral calculus,

$$\int_a^t f_x'(x)\, dx = f_n(t) - f_n(a)$$

Hence

$$\int_a^t g(x)\, dx = \lim_{n\to\infty} [f_n(t) - f_n(a)]$$

Since $<f_n>$ converges to f on [a, b], we have

$$\lim_{n\to\infty} f_n(t) = f(t) \text{ and } \lim_{n\to\infty} f_n(a) = f(a)$$

Hence

$$\int_a^t g(x)\, dx = f(t) - f(a)$$

and so

$$\frac{d}{dt}\left(\int_a^t g(x)\,dx\right) = f'(t)$$

or $$g(t) = f'(t),\ t \in [a, b]$$

This completes the proof of the theorem.

The series version of theorem 13 is

Theorem. 14. If a series Σf_n converges to f on [a, b] and

(i) each f_n is differentiable on [a, b]

(ii) each f_n' is continuous on [a, b]

(iii) the series $\Sigma f_n'$ converges uniformly to g on [a, b]

then f is differentiable on [a, b] and $f'(x) = g(x)$ for all $x \in [a, b]$.

Proof. Let $<S_n>$ be the sequence of partial sums of the series $\sum_{n=1}^{\infty} f_n$. Since Σf_n converges to f on [a, b], the sequence $<S_n>$ converges to f on [a, b]. Further, since $\Sigma f_n'$ converges uniformly to g on [a, b], the sequence $<S_n'>$ of partial sums converges uniformly to g on [a, b]. Hence, Theorem 13 is applicable and we have

$$f'(x) = g(x) \text{ for all } x \in [a, b].$$

Examples. 1. Consider the series $\sum_{n=1}^{\infty}\left[\frac{nx}{1+n^2x^2} - \frac{(n-1)x}{1+(n-1)^2x^2}\right]$

For this series, we have

$$S_n(x) = \frac{nx}{1+n^2x^2}, \qquad 0 \le x \le 1$$

We have seen that 0 is a point of non-uniform convergence for this sequence. We have

$$f(x) = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty}\frac{nx}{1+n^2x^2}$$
$$= 0 \text{ for } 0 \le x \le 1$$

Therefore

$$f'(0) = 0$$

$$S_n'(0) = \lim_{n\to 0}\frac{S_n(0+h) - S_n(0)}{h}$$

$$= \lim_{n\to 0}\frac{n}{1+n^2h^2} = n$$

Hence

$$\lim_{n\to\infty} S_n'(0) = \infty$$

Then

$$f'(0) \ne \lim_{n\to\infty} S_n'(0).$$

2. Consider the series $\sum_{n=1}^{\infty}\frac{\sin nx}{n^3}$, $x \in \mathbf{R}$. We have

$$f_n(x) = \frac{\sin nx}{n^3}$$

$$f_n'(x) = \frac{\cos nx}{n^2}$$

Thus

$$\Sigma f_n'(x) = \Sigma\frac{\cos nx}{n^2}$$

Since $\left|\dfrac{\cos nx}{n^2}\right| < \dfrac{1}{n^2}$ and $\Sigma \dfrac{1}{n^2}$ is convergent, therefore, by Weierstrass's M-test the series $\Sigma f_n'(x)$ is uniformly as well as absolutely convergent for all x ε **R** and so Σf_n can be differentiated term by term.

Hence $$\left(\sum_{n=1}^{\infty} f_n\right)' = \sum_{n=1}^{\infty} f_n'$$

or $$\left(\sum_{n=1}^{\infty} \frac{\sin nx}{n^3}\right)' = \sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$$

1.6. Weierstrass's Approximation Theorem.

Weierstrass proved an important result regarding approximation of continuous function which has many application in Numerical Methods and other branches of mathematics.

The following computation shall be required for the proof of Weierstrass's Approximation Theorem.

For any p, q ε **R**, we have, by Binomial Theorem

(1.6.1) $$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n, \quad n \in I,$$

where

$$\binom{n}{k} = \frac{\lfloor n}{\lfloor k \lfloor n-k}.$$

Differentiating with respect to p, we obtain

$$\sum_{k=0}^{n} \binom{n}{k} k\, p^{k-1} q^{n-k} = m\,(p+q)^{n-1},$$

which implies

(1.6.2) $$\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} p^k q^{n-k} = p(p+q)^{n-1}, \ n \in I$$

Differentiating once more, we have

$$\sum_{k=0}^{n} \frac{k^2}{n} \binom{n}{k} p^{k-1} q^{n-k} = p(n-1)(p+q)^{n-2} + (p+q)^{n-1}$$

and so

(1.6.3) $$\sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} p^k q^{n-k} = p^2\left(1-\frac{1}{n}\right)(p+q)^{n-2} + \frac{p}{n}(p+q)^{n-1}$$

Now if $x\in[0, 1]$, take p = x and q = 1 –x. Then (1.6.1), (1.6.2) and (1.6.3) yield

(1.6.4) $$\begin{cases} \displaystyle\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1 \\ \displaystyle\sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} = x \\ \displaystyle\sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} x^k (1-x)^{n-k} = x^2\left(1-\frac{1}{n}\right) + \frac{x}{n} \end{cases}$$

On expanding $\left(\dfrac{k}{n} - x\right)^2$, it follows from (1.6.4) that

(1.6.5) $$\sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 \binom{n}{k} x^k (1-x)^{n-k} = \frac{x(1-x)}{n} \qquad (0 \le x \le 1)$$

For any $f \in [0, 1]$, we define a sequence of polynomials $\{B_n\}_{n=1}^{\infty}$ as follows:

(1.6.6.) $$B_n(x) \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right), \qquad 0 \le x \le 1,\ n \,\varepsilon\, I.$$

The polynomial B_n is called the nth **Bernstein Polynomial** for f.

We are now in a position to state and prove Weierstrass's Theorem.

Theorem. 15. (Weierstrass's Approximation Theorem). Let f be a continuous function defined on [a, b]. Then given $\in > 0$, there exists a polynomial P such that

$$|P(x) - f(x)| < \in,\ a \le x \le b$$

Proof. We first show that it is sufficient to prove the theorem for the case in which [a, b] = [0, 1].

Suppose that the theorem is true for continuous functions defined on [0, 1]. If f is continuous on [a, b] and $\in>0$ we must show that there is a polynomial P such that

(1.6.7) $$|P(x) - f(x)| < \in \qquad a \le x \le b$$

Define g by $$g(x) = f(a + [b- a]x) \qquad (0 \le x \le 1)$$

Then $g(0) = f(a)$, $g(1) = f(b)$. Clearly g is continuous on [0, 1]. Since the theorem holds for continuous function defined on [0, 1], there is a polynomial Q such that

(1.6.8) $$|g(y) - Q(y)| < \in \quad (0 \le y \le 1)$$

If we put $$y = \frac{x-a}{b-a}$$

Then

$$g(y) = g\left(\frac{x-a}{b-a} = f\ (a + ((b-a)\frac{x-a}{b-a}\right)$$
$$= f(x)$$

Therefore (1.6.8) reduces to

$$|f(x) - Q\left(\frac{x-a}{b-a}\right)| < \in \qquad (a < x < b)$$

If we define P by

$$P(x) = Q\left(\frac{x-a}{b-a}\right),$$

then P is a polynomial (because Q is polynomial). Hence

$$|f(x) - P(x)| < \in,\ \ a \le x < b.$$

Thus we prove the theorem for functions f which are continuous over (0, 1]

Since f is continuous over compact set [0, 1], it is uniformly continuous on [0, 1]. Hence given $\in > 0$ there exists $\delta>0$ such that

$$|f(x) - f(y)| < \frac{\in}{2}, \qquad (|x-y| < \delta;\ x, y \,\varepsilon\, [0, 1])$$

Suppose N ε I such that

(1.6.9) $$\frac{1}{\sqrt[4]{N}} < \delta$$

and such that

(1.6.10) $$\frac{1}{\sqrt{N}} < \frac{\in}{4|f|} \qquad (|f| > 0)$$

Fix x ε [0, 1]. Multiplying the first identity in (1.6.4) by f and subtracting (1.6.6), we obtain for any n ε I,

(1.6.11) $$f(x) - B_n(x) = \sum_{k=0}^{n} \left[f(x) - f\left(\frac{k}{n}\right)\right] \binom{n}{k} x^h (1-x)^{n-k}$$
$$= \Sigma_1 + \Sigma_2, \text{ say,}$$

where Σ_1 is the sum over those values of k such that

(1.6.12) $$\left|\frac{k}{n} - x\right| < \frac{1}{\sqrt[4]{N}}\ ,$$

while Σ_2 is the sum over other values of k. If k does not satisfy (1.6.12), that is, if $\left|\frac{k}{n} - x\right| > \frac{1}{4\sqrt{n}}$,

then

$$(k - nx)^2 = n^2 \left|\frac{k}{n} - x\right|^2 \geq \sqrt{n^3}$$

Hence

$$|\Sigma_2| = |\Sigma_2[f(x) - f\left(\frac{k}{n}\right)] \binom{n}{k} x^k (1-x)^{n-k}|$$

$$\leq \Sigma_2[|f(x)| + |f\left(\frac{k}{n}\right)|] \binom{n}{k} x^k (1-x)^{n-k}$$

$$\leq 2|f(x)|\ \Sigma_2 \binom{n}{k} x^k (1-x)^{n-k}$$

$$\leq 2\ \frac{|f|}{\sqrt{n^3}}\ \Sigma_2\ (k-nx)^2 \binom{n}{k} x^k (1-x)^{n-k}$$

$$\leq \frac{2|f|}{\sqrt{n^3}} \sum_{h=0}^{n} (k - nx)^2 \binom{n}{k} x^k (1-x)^{n-k}$$

Here, by (1.6.5)

$$|\Sigma_2| \leq \frac{2|f|}{\sqrt{n^3}}\ n\ x\ (1-x)$$

$$\leq \frac{2|f|}{\sqrt{n}}$$

If $n \geq N$, it follows from (1.6.10) that $\frac{1}{\sqrt{n}} < \frac{\in}{4|f|}$ and so

$$|\Sigma_2| < \in/2.$$

Moreover, if $n \geq N$ and k satisfies (1.6.12), then by (1.6.9) and (1.6.12), $\left|\frac{k}{n} - x\right| < \delta$ and so

$$|f(x) - f\left(\frac{k}{n}\right)| < \in/2$$

Thus $$|\Sigma_1| = |\Sigma_1[f(x) - f\left(\frac{k}{n}\right)] \binom{n}{k} x^k (1-x)^{n-k}|$$

$$< \frac{\in}{2}\ \Sigma_1 \binom{n}{k} x^k (1-x)^{n-k}$$

and so by first identity of (1.6.4), we have

$$|\Sigma_1| < \frac{\in}{2}$$

Thus, (1.6.11) yields

$$|f(x) - B_n(x) \leq |\Sigma_1| + |\Sigma_2|$$

$$< \frac{\in}{2} + \frac{\in}{2} = \in.$$

Since x was arbitrary point in [0, 1] and n any integer with $n \geq N$, this shows that

$$|f(x) - B_n(x)| < \in, \qquad 0 \leq x \leq 1, n \varepsilon I.$$

This completes the proof of the theorem.

Example. If f is continuous on [0, 1] and if

$$\int_0^1 x^n f(x)\, dx = 0 \text{ for } n = 0, 1, 2, \ldots,$$

use Weierstrass's Approximation Theorem to prove that $f(x) = 0$ on [0, 1]

Solution. The given hypothesis is that the integral of the product of f with any polynomial is zero. We shall show that $\int_0^1 f^2 = 0$. We have, by Weierstrass's Approximation Theorem

$$\int_0^1 f^2 \doteq \int_0^1 P(x) f(x) = 0$$

$$\Rightarrow \qquad f = 0$$

1.7. Power Series

In this section we shall consider power series with real coefficients, and study its properties.

Definition. A series of the form $\sum\limits_{n=0}^{\infty} a_n x^n$ is called a power series

Applying Cauchy's root test, we observe that the power series $\sum\limits_{n=0}^{\infty} a_n x^n$ is convergent if

$$|x| < \frac{1}{l},$$

where

$$l = \overline{\lim}\, |a_n|^{1/n}$$

The series is divergent if $|x| > \frac{1}{l}$

Taking

$$r = \frac{1}{\overline{\lim}\, |a_n|^{1/n}}$$

We say that the power series is absolutely convergent if $|x| < r$ and divergent if $|x| > r$. If $a_0, a_1, \ldots$ are all real and if x is real, we get an interval $-r < x < r$ inside which the series is convergent.

If x is replaced by a complex number z, the power series $\Sigma\, a_n z^n$ converges absolutely at all points z inside the circle $|z| = r$ and does not converge at any point outside this circle. The circle is known as **circle of convergence** and r is called **radius of convergence**. In case of real power series the interval $(-r, r)$ is called interval of convergence.

If $\overline{\lim}\, |a_n|^{1/n} = 0$, then $r = \infty$ and the power series converges for all finite value of x(r z). The function represented by the sum of the series is then called an **Entire function** or an integral function. For example, e^z, sin z and cos z are integral functions.

If $\overline{\lim}\, |a_n|^{1/n} = \infty$, $r = 0$, the power series does not converge for any value of x except $x = 0$.

Theorem. 16. Suppose the series $\sum\limits_{n=0}^{\infty} a_n x^n$ converges for $|x| < r$ and define

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \qquad (|x| < r)$$

Then $\sum\limits_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-r + \in, r - \in]$, $\in > 0$. The function f is continuous and differentiable in $(-r, r)$ and

$$f'(x) = \sum_{n=1}^{\infty} n\, a_n x^{n-1} \qquad (|x| < r)$$

Proof. Let $\in$ be a positive number. If $|x| \le r - \in$, we have

$$|a_n x^n| \le |a_n (r - \in)^n|$$

Since every power series converges absolutely in the interior of its interval of convergence by Cauchy's root test, the series $\Sigma a_n (r - \in)^n$ converges absolutely and so, by Weierstrass's M-test, $\Sigma a_n x^n$ converges uniformly on $[-r + \in, r - \in]$. Also then the sum $f(x)$ of $\Sigma a_n x^n$ is a continuous function at all points inside the interval of convergence.

Since $(n)^{1/n} \to 1$ as $n \to \infty$, we have

$$\overline{\lim}\, (n|a_n|)^{1/n} = \overline{\lim}\, (|a_n|)^{1/n}$$

Hence the series $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} n\, a_n x^{n-1}$ have the same interval of convergence. Since $\sum_{n=0}^{\infty} n\, a_n x^{n-1}$ is a power series, it converges uniformly in $[-r + \in, r - \in]$ for every $\in > 0$. Then, by term by term differentiation (Theorem 14) yields.

$$\Sigma\, n\, a_n x^{n-1} = f'(x) \text{ if } |x| < r - \in.$$

But, given any x such that $|x| < r$, we can find an $\in > 0$ such that $|x| < r - \in$. Hence

$$\Sigma\, n\, a_n x^{n-1} = f'(x) \text{ if } |x| < r.$$

Theorem 17. Under the hypothesis of Theorem 16, f hs derivative of all orders in $(-r, r)$ which are given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\ldots(n-k+1)\, a_n x^{n-k}$$

In particular

$$f^{(k)}(0) = \underline{|k}\, a_h,\ k = 0, 1, 2, \ldots$$

Proof. Let

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

Then by the above theorem

$$f'(x) = \Sigma\, n\, a_n x^{n-1}$$

Now applying the theorem 16 to $f'(x)$, we have

$$f''(x) = \Sigma\, n(n-1)\, a_n x^{n-2}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\ldots(n-k+1)\, a_n x^{n-k}$$

Clearly $f^{(k)}(0) = k\, a_k$; the other terms vanish at $x = 0$.

Remark. If the coefficients of a power series are known, the values of the derivatives of f at the centre of the interval of convergence can be found from the relation

$$f^{(k)}(0) = \underline{|k}\, a_k.$$

Also we can find coefficients from the values at origin of t, $f', f'', \ldots$

Theorem. 18. (Uniqueness Theorem). If $\Sigma a_n x^n$ and $\Sigma b_n x^n$ converge on some interval $(-r, r)$, $r > 0$ to the same function f, then

$$a_n = b_n \text{ for all } n \in \mathbf{N}.$$

Proof. Under the given condition, the function f have derivatives of all order in $(-r, r)$ given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\ldots(n-k+1)\, a_n x^{n-k}$$

Putting $x = 0$, this yields

$$f^{(k)}(0) = \underline{|k}\, a_k \text{ and } f^{k}(0) = \underline{|k}\, b_k$$

for all k $\in$**N**. Hence

$$a_k = b_k \text{ for all } k\in\mathbf{N}.$$

This completes the proof or the theorem.

Theorem. 19. (Abel). Let $(-r, r)$ be the interval of convergence of the power series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

If the series is convergent when $x = r$, then

$$\lim_{x\to r-0} f(x) = f(r)$$

A similar result holds if the series is convergent when $x = -r$.
Putting $x = ry$, we obtain the power series

$$\sum_{n=0}^{\infty} a_n r^n y^n = \sum_{n=0}^{\infty} b_n y^n, \text{ say,}$$

whose interval of convergence in $(-1, 1)$. It is therefore sufficient to prove the theorem for $r = 1$. Hence we shall prove the following.

Theorem. 19. (Abel). Let $(-1, 1)$ be interval of convergence for the power series $\Sigma\, a_n x^n$. if $\sum\limits_{n=0}^{\infty} a_n = S$, than

$$\lim_{x\to 1-0} \sum_{n=0}^{\infty} a_n x^n = S.$$

Proof. Let $S_n = a_0 + a_1 + \ldots + a_n$, $S_{-1} = 0$. Then

$$\begin{aligned}
\sum_{n=0}^{m} a_n x^n &= \sum_{n=0}^{m} (S_n - S_{n-1})\, x^n \\
&= \sum_{n=0}^{m} S_n x^n - \sum_{n=0}^{m} S_{n-1} x^n \\
&= \sum_{n=0}^{m-1} S_n x^n + S_m x^m - \sum_{n=0}^{m} S_{n-1} x^n \\
&= \sum_{n=0}^{m-1} S_n x^n - x \sum_{n=0}^{m} S_{n-1} x^{n-1} + S_m x^m \\
&= (1-x) \sum_{n=0}^{m-1} S_n x^n + S_m x^m
\end{aligned}$$

For $|x| < 1$, let $m\to\infty$ and obtain

(1.7.1) $$f(x) = (1-x) \sum_{n=0}^{\infty} S_n x^n$$

Since $\Sigma\, a_n = S$, $S_n \to S$ as $n\to\infty$. So to each $\in>0$, there exists an integer N such that $n>N$ implies

$$|S - S_n| < \in/2$$

Also we know that

$$(1-x) \sum_{n=0}^{\infty} x^n = 1 \quad (|x| <1)$$

or

(1.7.2) $$S = (1-x) \sum_{n=0}^{\infty} Sx^n \qquad (|x| <1)$$

Then (1.7.1) and (1.7.2) yield

$$| f(x) - S| = |(1-x) \sum_{n=0}^{\infty} (S_n - S) x^n|$$

$$\leq (1-x) \sum_{n=0}^{N} |S_n-S| |x|^n + \sum_{n=N+1}^{\infty} |S_n-S| |x|^n$$

$$\leq (1-x) \sum_{n=0}^{N} |S_n - S| |x|^n + \frac{\in}{2}$$

But for a fixed N, $(1-x) \sum_{n=0}^{N} |S_n -S| |x|^n$ is a positive continuous function of x having zero value at x = 1. Therefore there exists $\delta>0$ such that for $1 - \delta < x < 1$, $(1-x) \sum_{n=0}^{N} |S_n -S| |x|^n$ is less than $\frac{\in}{2}$. Hence

$$| f(x) -S| < \frac{\in}{2} + \frac{\in}{2} = \in, \; 1-\delta < x < 1$$

and so

$$\lim_{x\to 1-} f(x) = S = \sum_{n=0}^{\infty} a_n$$

Tauber's Theorem. The converse of Abel's theorem proved above is false in general. If *f* is given by

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad -r < x < r$$

the limit $f(r-)$ may exist but yet the series $\Sigma\, a_n r^n$ may fail to converge. For example, if $a_n = (-1)^n$, then

$$f(x) = \frac{1}{1+x}, \; -1 < x < 1$$

and $f(x) \to \frac{1}{2}$ as $x\to 1-$. However $\Sigma\, (-1)^n$ is not convergent. Tauber showed that the converse of Abel's theorem can be obtained by imposing additional condition on the coefficients a_n. A large number of such results are known now a days as Tauberian Theorems. We present here only Tauber's first theorem.

Theorem. 20. (Tauber). Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $-1 < x < 1$ and suppose that $\lim_{n\to\infty} n\, a_n = 0$. If $f(x) \to S$ as $x\to 1-$, then $\sum_{n=0}^{\infty} a_n$ converges and has the sum S.

Proof. Let $n\, \sigma_n = \sum_{h=0}^{n} k\, |a_k|$. Then $\sigma_n\to 0$ as $n\to\infty$. Also, $\lim_{n\to\infty} f(x_n) = S$ if $x_n = 1 - \frac{1}{n}$. Therefore to each $\in >0$, we can choose an integer N such that $n \geq N$ implies

$$| (fx_n) - S| < \frac{\in}{3}, \quad \sigma_n < \frac{\in}{3}, \quad n\, |a_n| < \frac{\in}{3}.$$

Let $S_n = \sum_{h=0}^{n} a_k$. Then for $-1 < x < 1$, we have

$$S_n - S = f(x) - S + \sum_{k=0}^{n} a_k (1-x^k) - \sum_{h=n+1}^{\infty} a_k x^k$$

Let $x \in (0, 1)$. Then

$$(1-x^k) = (1-x) (1+x + -+ x^{k-1}) \leq k(r-x)$$

for each k. Therefore, if $n \geq N$ and $0 < x < 1$, we have

$$|S_n - S| \le |f(x) - S| + (1-x)\sum_{h=0}^{n} k\,|a_k| + \frac{\in}{3n(1-x)}$$

Putting $x = x_n = 1-\frac{1}{n}$, we find that

$$|S_n - S| < \frac{\in}{3} + \frac{\in}{3} + \frac{\in}{3} = \in_1$$

which completes the proof.

2

FUNCTIONS OF SEVERAL VARIABLES

2.1. In this chapter, we shall study derivatives and partial derivatives of functions of several variables alongwith their properties.

2.2. Linear Transformations

Definition. A mapping $\boldsymbol{f}$ of a vector space X into a vector space Y is said to be a linear transformation if

$$\boldsymbol{f}(\mathbf{x}_1 + \mathbf{x}_2) = \boldsymbol{f}(\mathbf{x}_1) + \boldsymbol{f}(\mathbf{x}_2),$$
$$\boldsymbol{f}(c\mathbf{x}) = c\boldsymbol{f}(\mathbf{x})$$

for all $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2 \in X$ and all scalars c.

Clearly, if f is linear transformation, then $\boldsymbol{f}(0) = 0$.

A linear transformation of a vector space X into X is called linear operator on X.

If a linear operator T on a vector space X is one-to-one and onto, then T is invertible and its inverse is denoted by $\mathbf{T}^{-1}$. Clearly $\mathbf{T}^{-1}(T\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in X$. Also, if T is linear, then $\mathbf{T}^{-1}$ is also linear.

Theorem 1. A linear operator T on a finite dimensional vector space X is one-to-one if and only if the range of T is whose of X.

Proof. Let R(T) denote range of T. Let $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ be a basis of X. Since T is linear the set $(T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n)$ spans R(T). The range of T will be whole of X if and only if $\{T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n\}$ is linearly independent

So, Suppose first that T is one-to-one. We shall prove that $\{T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n\}$ is linearly independent. Hence, let

$$c_1T\mathbf{x}_1, c_2T\mathbf{x}_2 + \ldots + c_nT\mathbf{x}_n = 0$$

Since T is linear, this yields

$$T(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_n\mathbf{x}_n) = 0$$

and so $$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_n\mathbf{x}_n = 0$$

Since $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is linearly independent, we have

$$c_1 = c_2 = \ldots c_n = 0$$

Thus $\{T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n\}$ is linearly independent and so R(T) = X if T is one-to-one.

Conversely, suppose $\{T\mathbf{x}_1, T\mathbf{x}_2, \ldots, T\mathbf{x}_n\}$ is linearly independent and so

(2.2.1) $$c_1T\mathbf{x}_1 + c_2T\mathbf{x}_2 + \ldots + c_nT\mathbf{x}_n = 0$$

implies $c_1 = c_2 = \ldots = c_n = 0$. Since T is linear, (2.2.1) implies

$$T(c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n) = 0$$
$$\Rightarrow \quad c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n = 0$$

Thus T(x) = 0 only if x = 0. Now

$$T(\mathbf{x}) = T(\mathbf{y}) \quad \Rightarrow T(\mathbf{x}-\mathbf{y}) = 0 \qquad \Rightarrow \mathbf{x} - \mathbf{y} = 0 \quad \Rightarrow \mathbf{x} = \mathbf{y}$$

and so T is one-to-one. This completes the proof of the theorem

Definition. Let L(X, Y) be the set of all linear transformations of the vector space X into the vector space Y. If $T_1, T_2 \in L(X, Y)$ and if c_1, c_2 are scalars, then $c_1 T_1 + c_2 T_2$ is defined by

$$(c_1T + c_2T_2)(\mathbf{x}) = c_1T_1\mathbf{x} + c_2T_2\mathbf{x}, \mathbf{x} \in X.$$

It can be shown that $c_1 T_1 + c_2 T_2 \in L(X, Y)$.

Definition. Let X, Y, Z be vector spaces over the same field. If T, S $\in$ L(X, Y) we define their product ST by

$$(ST)(\mathbf{x}) = S(T\mathbf{x}), \ x \in X$$

Also $$ST \in L(X, Y).$$

Definition. Let R^n denote n-dimensional Euclidean space and let $T \in L(\mathbf{R}^n, \mathbf{R}^m)$. Then

$$\text{lub}\{|T\mathbf{x}| : \mathbf{x} \in \mathbf{R}^n, |x| \le 1\}$$

is called Norm of T and is denoted by $\|T\|$. The inequality

$$|T\mathbf{x}| \le \|T\|\,|\mathbf{x}|$$

holds for all $\mathbf{x} \in \mathbf{R}^n$. Also if λ is such that

$$|T\mathbf{x}| \le \lambda|\mathbf{x}|,\ \mathbf{x} \in \mathbf{R}^n, \text{ then } \|T\| \le \lambda.$$

We are now in a position to prove the following theorem.

Theorem. 2. Let T, S$\in$L ($\mathbf{R}^n$, R^m) and c be a scalar. Then

(a) $\|T\| < \infty$ and T is uniformly continuous mapping of R^n into R^m.

(b) $\|S + T\| \le \|T\| + \|S\|$

$\|CT\| = |C|\,\|T\|$.

(c) If $d(T, S) = \|T-S\|$, than d is a metric

Proof. (a) Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be the standard basis in $\mathbf{R}^n$ and let $\mathbf{x} \in \mathbf{R}^n$. Then $\mathbf{x} = \sum_{i=1}^{n} c_i\, e_i$. Suppose $|\mathbf{x}| < 1$ so that $|c_i| \le 1$ for i = 1, 2, …n. Then

$$|T\mathbf{x}| = |\Sigma\, c_i T\, \mathbf{e}_i| \le \Sigma\, |c_i|\, |T\mathbf{e}_i|$$
$$\le \Sigma\, |T\mathbf{e}_i|$$

Taking lub over $\mathbf{x}\in\mathbf{R}^n$, $|\mathbf{x}| \le 1$

$$\|T\mathbf{x}\| \le \sum_{i=1}^{n} |T\mathbf{e}_i| < \infty.$$

Further

$$|T\mathbf{x} - T\mathbf{y}| = |T(\mathbf{x}-\mathbf{y})| \le \|T\|\,|\mathbf{x}-\mathbf{y}|\ ;\ \mathbf{x}, \mathbf{y}, \in\mathbf{R}^n$$

so if $|\mathbf{x}-\mathbf{y}| < \dfrac{\in}{\|t\|}$, then

$$|T\mathbf{x} - T\mathbf{y}| < \in,\ \mathbf{x}, \mathbf{y}, \in\mathbf{R}^n.$$

Hence T is uniformly continuous.

(b) We have

$$|(T + S)\,\mathbf{x}\,| = |T\mathbf{x} + S\mathbf{x}|$$
$$\le |T\mathbf{x}| + |S\mathbf{x}|$$
$$\le \|T\|\,|\mathbf{x}| + \|S\|\,|\mathbf{x}|$$
$$= (\|T\|\ \ \|S\|)\,|\mathbf{x}|$$

Taking lub over $\mathbf{x} \in\mathbf{R}^n$, $|\mathbf{x}| \le 1$, we have

$$\|T + S\| \le \|T\| + \|S\|$$

Similarly, it can be shown that

$$\|cT\| = |c|\,\|T\|.$$

(c) We have $\quad d(T, S) = \|T-S\| \ge 0$ and $d(T, S) = \|T - S\| = 0 \Leftrightarrow T = S$.

Also $\quad d(T, S) = \|T-S\| = \|S-T\| = d(S, T)$

Further, if S, T, U $\in$ L ($\mathbf{R}^n$, $\mathbf{R}^m$), then

$$\|S - U\| = \|S - T + T - U\|$$
$$\le \|S - T\| + \|T - U\|$$

Hence d is a metric.

Theorem. 3. If T $\in$ L(R^n, R^m) and S $\in$L (R^m, R^k), then

$$\|S\,T\| \le \|S\|\,\|T\|$$

Proof. We have

$$|(ST)\,\mathbf{x}| = |\,s(T\mathbf{x})| \le \|S\|\,|\,T\mathbf{x}|$$
$$\le \|S\|\,\|T\|\,|\mathbf{x}|$$

Taking sup. over $\mathbf{x}$, $|\mathbf{x}| \le 1$, we get

$$\|ST\| \le \|S\|\ \|T\|.$$

In Theorem 2, we have seen that the set of linear transformation form a metric space. Hence the concepts of convergence, continuity, open sets, etc make sense in $\mathbf{R}^n$.

Theorem. 4. Let *C* be the collection of all invertible linear operator on $\mathbf{R}^n$.

(a) If $T \in C$, $\|T^{-1}\| = \dfrac{1}{\alpha}$, $S \in L(\mathbf{R}^n, \mathbf{R}^n)$ and $\|S - T\| = \beta < \alpha$, then $S \in C$.

(b) C is an open subset of $L(\mathbf{R}^n, \mathbf{R}^n)$ and the mapping $T \to T^{-1}$ is continuous on C.

Proof. We note that

$$|\mathbf{x}| = |T^{-1} T \mathbf{x}| \le \|T^{-1}\| \, |T\mathbf{x}|$$
$$= \frac{1}{\alpha} |T\mathbf{x}| \text{ for all } \mathbf{x} \in \mathbf{R}^n$$

and so

(2.2.2)
$$\begin{aligned}(\alpha - \beta) |\mathbf{x}| &= \alpha|\mathbf{x}| - \beta|\mathbf{x}| \\ &\le |T\mathbf{x}| - \beta |\mathbf{x}| \\ &\le |T\mathbf{x}| - |(S - T) \mathbf{x}| \\ &\le |S\mathbf{x}| \text{ for all } \mathbf{x} \in \mathbf{R}^n.\end{aligned}$$

Thus kernel of S consists of 0 only. Hence S is one-to-one. Then theorem 1 implies that T is also onto. Hence S is invertible and so $S \in C$. But this holds for all S satisfying $\|S-T\| < \alpha$. Hence every point of C is an interior point and so C is open.

Replacing $\mathbf{x}$ by $S^{-1} \mathbf{y}$ in (2.2.2), we have

$$(\alpha - \beta) |S^{-1} \mathbf{y}| \le |SS^{-1} \mathbf{y}| = |\mathbf{y}|$$

or
$$|S^{-1} y| \le \frac{|\mathbf{y}|}{\alpha - \beta}$$

and so
$$\|S^{-1}\| \le \frac{1}{\alpha - \beta}$$

Since
$$S^{-1} - T^{-1} = S^{-1} (T - S) T^{-1}.$$

We have

(2.2.3)
$$\begin{aligned}\|S^{-1} - T^{-1}\| &\le \|S^{-1}\| \, \| T-S\| \, \|T^{-1}\| \\ &\le \frac{\beta}{\alpha(\alpha - \beta)}\end{aligned}$$

Thus if f is the mapping which maps $T \to T^{-1}$, then (2.2.3) implies

$$\|f(S) - f(T)\| \le \frac{\| S - T \|}{\alpha(\alpha - \beta)}$$

Hence, if $\|S - T\| \to 0$ then $f(S) \to f(T)$ and so f in continuous. This completes the proof of the theorem.

2.3. Total derivative of *f* defined on a subset E of $\mathbf{R}^n$.

In one-dimensional case, a function f with a derivative at c can be approximated by a linear polynomial. In fact if $f'(c)$ exists, let r(h) denote the difference

(2.3.1)
$$r(h) = \frac{f(x + h) - f(x)}{h} - f'(x) \text{ if } h \ne 0$$

and let $r(0) = 0$. Then we have

(2.3.2)
$$f(x + h) = f(x) + h f'(x) + h\, r(h),$$

an equation which holds also for $h = 0$. The equation (2.3.2) is called the **First Order Taylor Formula for approximating** $f(x + h) - f(x)$ by $h f'(x)$. The error committed in this approximation is $h\, r(h)$. From (2.3.1), we observe that $r(h) \to 0$ as $h \to 0$. The error $h\, r(h)$ is said to be of smaller order than h as $h \to 0$. We also note that $hf'(x)$ is a linear function of h. Thus, if we write $Ah = hf'(x)$, then

$$A(a h_1 + b h_2) = aAh_1 + bAh_2$$

The aim of this section is to study total derivative of a function $\boldsymbol{f}$ from $\mathbf{R}^n$ to $\mathbf{R}^m$ in such a way that the above said properties of $h f'(x)$ and $h\, r(h)$ are preserved.

Definition. Suppose E is an open set in $\mathbf{R}^n$ and let $\boldsymbol{f} : E \to \mathbf{R}^n$ be a function defined on a set E in $\mathbf{R}^n$ with values in $\mathbf{R}^m$. Let $\mathbf{x} \in E$ and $\mathbf{h}$ be a point in $\mathbf{R}^n$ such that $|\mathbf{h}| < r$ and $\mathbf{x} + \mathbf{h} \in B(\mathbf{x}, r)$. Then $\boldsymbol{f}$ is said to be differentiable at $\mathbf{x}$ if there exists a linear transformation A of $\mathbf{R}^n$ into $\mathbf{R}^m$ such that

(2.3.3) $$\boldsymbol{f}(\mathbf{x} + \mathbf{h}) = \boldsymbol{f}(\mathbf{x}) + A\mathbf{h} + \mathbf{r}(\mathbf{h}),$$

where the remainder $\mathbf{r}(\mathbf{h})$ is small in the sense that

$$\lim_{h\to 0} \frac{|r(\mathbf{h})|}{|\mathbf{h}|} = 0.$$

and we write $f'(\mathbf{x}) = A$.

The equation (2.3.3) is called a **First Order Taylor Formula**.

The equation (2.2.4) can bw written as

(2.3.4) $$\lim_{h\to 0} \frac{|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A\mathbf{h}|}{|\mathbf{h}|} = 0$$

The equation (2.3.4) thus can be interpreted as "For fixed $\mathbf{x}$ and small $\mathbf{n}$, $\boldsymbol{f}(\mathbf{x} + \mathbf{h}) - \boldsymbol{f}(\mathbf{x})$ is approximately equal to $\boldsymbol{f}'(\mathbf{x})\mathbf{h}$, that is, the value of a linear function applied to $\mathbf{h}$."

Also (2.3.3) shows that $\boldsymbol{f}$ is continuous at any point at which $\boldsymbol{f}$ is differentiable

The derivatives A$\mathbf{h}$ derived by (2.3.3) or (2.3.4) is called **total derivative** of $\boldsymbol{f}$ at $\mathbf{x}$ or the **differential of $\boldsymbol{f}$ at x**.

In particular, let f be a real valued function of three variables x,y, z say. Then f is differentiable at the point (x, y, z) if it possesses a determinate value in the neighbourhood of this point and if $\Delta f = f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z) = A\Delta x + B\Delta y + C\Delta z + \in\rho$, where $\rho = |\Delta x| + |\Delta x| + |\Delta z|$, $\in \to 0$ as $\rho \to 0$ and A, B, C are independent of x, y, z. In this case $A\Delta x + B\Delta y + C\Delta z$ is called differential of f at (x, y, z).

Theorem 5. (**Uniqueness of Derivative of a function**). Let E be an open set in $\mathbf{R}^n$ and $\boldsymbol{f}$ maps E in $\mathbf{R}^m$ and $\mathbf{x} \in E$. Suppose $\mathbf{h} \in \mathbf{R}^n$ is small enough such that $\mathbf{x} + \mathbf{h} \in E$. Then $\boldsymbol{f}$ has a unique derivative.

Proof. If possible, let there the two derivatives A_1 and A_2. Therefore

$$\lim_{\mathbf{h}\to 0} \frac{|f(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_1\mathbf{h}|}{|\mathbf{h}|} = 0$$

and

$$\lim_{\mathbf{h}\to 0} \frac{|f(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_2\mathbf{h}|}{|\mathbf{h}|} = 0$$

Consider $B = A_1 - A_2$. Then

$$\begin{aligned} B\mathbf{h} &= A_1\mathbf{h} - A_2\mathbf{h} \\ &= f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) + f(\mathbf{x}) - f(\mathbf{x}+\mathbf{h}) + A_1\mathbf{h} - A_2\mathbf{h} \\ &= \boldsymbol{f}(x+\mathbf{h}) - \boldsymbol{f}(\mathbf{x}) - A_2\mathbf{h} + f(\mathbf{x}) - f(\mathbf{x}+\mathbf{h}) + A_1\mathbf{h} \end{aligned}$$

and so $$|B\mathbf{h}| < |\boldsymbol{f}(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A_1\mathbf{h}| + |f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A_2\mathbf{h}|$$

which implies

$$\lim_{\mathbf{h}\to 0} \frac{|B\mathbf{h}|}{\mathbf{h}} \le \lim_{h\to 0} \frac{|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_2\mathbf{h}|}{|\mathbf{h}|} + \frac{|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - A_2\mathbf{h}|}{|\mathbf{h}|}$$

$$= 0$$

For fixed $\mathbf{h} \neq 0$, it follows that

(2.3.5) $$\frac{|B(t\mathbf{h})|}{|t\mathbf{h}|} \to 0 \text{ as } t\to 0$$

The linearity of B shows that L.H.S. of (2.3.5) is independent of t. Thus $B\mathbf{h} = 0$ for all $\mathbf{h} \in \mathbf{R}^n$. Hence B = 0, that is, $A_1 = A_2$, which proves uniqueness of the derivative.

The following theorem, known as chain rule, tells us how to compute the total derivatives of the composition of two functions.

Theorem. 6. (**Chain rule**). Suppose E is an open set in $\mathbf{R}^n$, $\boldsymbol{f}$ maps E into $\mathbf{R}^m$, $\boldsymbol{f}$ is differentiable at $\mathbf{x}_0$ will total derivative $\boldsymbol{f}'(x_0)$, $\mathbf{g}$ maps on open set containing $\boldsymbol{f}$ (E) into $\mathbf{R}^k$ and $\mathbf{g}$ is differentiable at $\boldsymbol{f}(\mathbf{x}_0)$ with total derivative $\mathbf{g}'$

$(\boldsymbol{f}(\mathbf{x}_0))$. Then the composition map $F = \boldsymbol{f} \circ \mathbf{g}$ mapping E into $\mathbf{R}^k$ and defined by $\mathbf{F}(\mathbf{x}) = \mathbf{g}(\boldsymbol{f}(\mathbf{x}))$ is differentiable at $\mathbf{x}_0$ and has the derivative

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(f(x_0))\, f'(x_0)$$

Proof. Take

$$\mathbf{y}_0 = \boldsymbol{f}(\mathbf{x}_0),\ A = \boldsymbol{f}'(x_0),\ B = \mathbf{g}'(y_0)$$

and define

$$\mathbf{r}_1(\mathbf{x}) = \boldsymbol{f}(\mathbf{x}) - \boldsymbol{f}(\mathbf{x}_0) - A(\mathbf{x} - \mathbf{x}_0)$$
$$\mathbf{r}_2(\mathbf{y}) = \mathbf{g}(\mathbf{y}) - \mathbf{g}(\mathbf{y}_0) - B(\mathbf{y} - \mathbf{y}_0)$$
$$\mathbf{r}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{x}_0) - BA(\mathbf{x} - \mathbf{x}_0)$$

To prove the theorem it is sufficient to show that

$$\mathbf{F}'(\mathbf{x}_0) = BA,$$

That is,

(2.3.6) $$\frac{\mathbf{r}(\mathbf{x})}{|\mathbf{x} - \mathbf{x_0}|} \to 0 \text{ as } \mathbf{x} - \mathbf{x}_0$$

But, in term of definition of $\mathbf{F}(\mathbf{x})$, we have

$$\mathbf{r}(\mathbf{x}) = \mathbf{g}(\boldsymbol{f}(\mathbf{x})) - \mathbf{g}(y_0) - B(\boldsymbol{f}(\mathbf{x}) - \boldsymbol{f}(\mathbf{x}_0) - A(\mathbf{x} - \mathbf{x}_0))$$

so that

(2.3.7) $$\mathbf{r}(\mathbf{x}) = r_2(\boldsymbol{f}(\mathbf{x})) + B\,\mathbf{r}_1(\mathbf{x})$$

If $\in > 0$, it follows from the definitions of A and B that there exist $\eta > 0$ and $\delta > 0$ such that

$$\frac{|r_2(\mathbf{y})|}{|\mathbf{y} - \mathbf{y}_0|} \le \in \text{ if } |\mathbf{y} - \mathbf{y}_0| < \eta$$

or $\quad |\mathbf{r}_2(\mathbf{y})| \le \in |\mathbf{y} - \mathbf{y}_0|$ if $|\mathbf{y} - \mathbf{y}_0| < \eta$, i.e. if $|\boldsymbol{f}(\mathbf{x}) - \boldsymbol{f}(\mathbf{x})| < \eta$

and $\quad |\mathbf{r}_1(\mathbf{x})| \le \in |\mathbf{x} - \mathbf{x}_0|$ if $|\mathbf{x} - \mathbf{x}_0| < \delta$.

Hence

(2.3.8) $$\begin{aligned} |\mathbf{r}_2(\boldsymbol{f}(\mathbf{x}))| &\le \in |\boldsymbol{f}(\mathbf{x}) - \boldsymbol{f}(\mathbf{x}_0)| \\ &= \in |\mathbf{r}_1(\mathbf{x}) + A(\mathbf{x} - \mathbf{x}_0)| \\ &\le \in^2 |\mathbf{x} - \mathbf{x}_0| + \in \|A\|(\mathbf{x} - \mathbf{x}_0) \end{aligned}$$

and

(2.3.9) $$\begin{aligned} |B\,\mathbf{r}_1(\mathbf{x})| &\le \|B\|\,|\mathbf{r}_1(\mathbf{x})| \\ &\le \in \|B\|\,|\mathbf{x} - \mathbf{x}_0| \quad \text{if } |\mathbf{x} - \mathbf{x}_0| < \delta. \end{aligned}$$

Using (2.3.8) and (2.3.9), the expression (2.3.7) yields

$$|\mathbf{r}(\mathbf{x})| < \in^2 |\mathbf{x} - \mathbf{x}_0| + \in \|A\|\,|\mathbf{x} - \mathbf{x}_0| + \in \|B\|\,|\mathbf{x} - \mathbf{x}_0|$$

and so

$$\begin{aligned} \frac{|r(x)|}{|\mathbf{x} - \mathbf{x}_0|} &\le \in^2 + \in \|A\| + \in \|B\| \\ &= \in [\in + \|A\| + \|B\|] \quad \text{if } |\mathbf{x} - \mathbf{x}_0| < \delta \end{aligned}$$

Hence

$$\frac{|r(x)|}{|\mathbf{x} - \mathbf{x}_0|} \to 0 \text{ as } \mathbf{x} \to \mathbf{x}_0$$

which in turn implies

$$\mathbf{F}'(x_0) = BA = \mathbf{g}'(\boldsymbol{f}(\mathbf{x}_0))\, \boldsymbol{f}'(x_0)$$

2.4. Partial Derivatives. Let $\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}\}$ be the standard basis of $\mathbf{R}^n$. Suppose $\boldsymbol{f}$ maps on open set $E \subset \mathbf{R}^n$ into $\mathbf{R}^m$ and let $f_1, f_2, \ldots, f_m$ be components of $\boldsymbol{f}$. Define $D_k f_i$ on E by

(2.4.1) $$(D_k f_i)(\mathbf{x}) = \lim_{t\to 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_k) - f_i(\mathbf{x})}{t}$$

provided the limit exists.

Writing $f_i(x_1, x_2, \ldots, x_n)$ in place of $f_i(\mathbf{x})$, we observe that $D_k f_i$ is the derivative of f_i with respect to x_k, keeping the other variable fixed. That is why, we use $\dfrac{\partial f_i}{\partial x_k}$ oftenly in place of $D_k f_i$.

Since $f = (f_1, f_2, \ldots, f_n)$, we have

$$D_k\, \boldsymbol{f}(\mathbf{x}) = (D_k\, f_1(\mathbf{x}), D_k\, f_2(\mathbf{x}),\ldots, D_k\, f_n(\mathbf{x}))$$

which is partial derivative of $\boldsymbol{f}$ with respect to x^k.

Furthermore, if $\boldsymbol{f}$ is differentiable at $\mathbf{x}$, then the definition of $\boldsymbol{f}'(\mathbf{x})$ shown that

(2.4.2) $$\lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x} + t\mathbf{h}_k) - \mathbf{f}(\mathbf{x})}{t} = \boldsymbol{f}'(\mathbf{x})\, \mathbf{h}_k$$

If we take $\mathbf{h}_k = \mathbf{e}_k$, taking components of vector is (2.4.2), it follows that

"If $\boldsymbol{f}$ is differentiable at x, then all partial derivatives $(D_k\, f_i)$ (x) exist."

In particular, if f in real valued (m = 1), then (2.4.1) takes the form

$$(D_k\, f)\, (\mathbf{x}) = \lim_{t\to 0} \frac{f(\mathbf{x} + t) - f(\mathbf{x})}{t}$$

For example, if f is a function of three variables x, y, and z, then

$$Df(x) = \lim_{\Delta x\to 0} \frac{f(x+\Delta x, y, z) - f(x, y, z)}{\Delta x}$$

$$Df(y) = \lim_{\Delta y\to 0} \frac{f(x, y+\Delta y, z) - f(x, y, z)}{\Delta y}$$

and $$Df(z) = \lim_{\Delta z\to 0} \frac{f(x, y, z+\Delta z) - f(x, y, z)}{\Delta z}$$

and are known respectively as partial derivatives of f with respect to x, y, z.

The next theorem shown that $A\mathbf{h} = \boldsymbol{f}'(\mathbf{x})\,(\mathbf{h})$ is a linear combination of partial derivatives of $\boldsymbol{f}$

Theorem. 7. Let $E \subseteq \mathbf{R}^n$ and let $\boldsymbol{f} : E\to\mathbf{R}^n$ be differentiable at $\mathbf{x}$ (interior point of open set E). If $\mathbf{h} = c_1\, \mathbf{e}_1 + c_2\, \mathbf{e}_2 +\ldots+ c_n\, \mathbf{e}_n$, where $\{\mathbf{e}_1, \mathbf{e}_2,\ldots, \mathbf{e}_n\}$ is a standard basis for $\mathbf{R}^n$, then

$$\boldsymbol{f}'(\mathbf{x})\, (\mathbf{h}) = \sum_{k=1}^{n} c_k\, D_k\, \boldsymbol{f}(\mathbf{x})$$

Proof. Using the linearity of $f'(\mathbf{x})$, we have

$$\boldsymbol{f}'(\mathbf{x})\, (\mathbf{h}) = \sum_{k=1}^{n} \boldsymbol{f}'(x)\, (\mathbf{c}_k\, e_h)$$

$$= \sum_{k=1}^{n} c_h\, \boldsymbol{f}'(\mathbf{x})\, \mathbf{e}_h$$

But, by (2.4.2),

$$\boldsymbol{f}'(\mathbf{x})\, \mathbf{e}_h = (D_h\, \boldsymbol{f})\, (\mathbf{x})$$

Hence $$\boldsymbol{f}'(\mathbf{x})\, (\mathbf{h}) = \sum_{k=1}^{n} c_k\, D_k\, (\boldsymbol{f})\, (\mathbf{x})$$

If f is real valued (m = 1), we have

$$\boldsymbol{f}'(\mathbf{x})\, (\mathbf{h}) = (D_1\boldsymbol{f}(\mathbf{x}), D_2\boldsymbol{f}(\mathbf{x}),\ldots, D_n\boldsymbol{f}(\mathbf{x})).\ \mathbf{h}.$$

Definition. A differentiable mapping $\boldsymbol{f}$ of an open set $E \subset \mathbf{R}^n$ into $\mathbf{R}^m$ is said to be continuously differentiable in E if $\boldsymbol{f}'$ is continuous mapping of E into $L(\mathbf{R}^n, \mathbf{R}^m)$.

Thus to every $\in > 0$ and every $\mathbf{x} \in E$ there exists a $\delta>0$ such that

$$\|\boldsymbol{f}(\mathbf{y}) - \boldsymbol{f}'(\mathbf{x})\| < \in \text{ if } \mathbf{y} \in E \text{ and } |\mathbf{y} - \mathbf{x}| < \delta.$$

In this case we say that $\boldsymbol{f}$ is a C'-mapping in E or that $\boldsymbol{f}\in C'(E)$.

Theorem. 8. Suppose $\boldsymbol{f}$ maps an open set $E \subset \mathbf{R}^n$ into $\mathbf{R}^m$. Then $\boldsymbol{f}$ is continuously differentiable if and only if the partial derivatives $D_j\, f_i$ exist and are continuous on E for $1 \le i \le m$, $1 \le j \le n$.

Proof. Suppose first that $\boldsymbol{f}$ is continuously differentiable in E. Therefore to each $\mathbf{x} \in E$ and $\in>0$, there exist a $\delta>0$ such that

$$\|\boldsymbol{f}'(\mathbf{y}) - \boldsymbol{f}'(\mathbf{x})\| < \in \text{ if } \mathbf{y} \in E \text{ and } |\mathbf{y}-\mathbf{x}| < \delta.$$

We have then

$$(2.4.3)\qquad \begin{aligned}|\boldsymbol{f}(\mathbf{y})\,\mathbf{e}_j - f(\mathbf{x})\,e_j| &= |(\boldsymbol{f}'(\mathbf{y}) - \boldsymbol{f}'(\mathbf{x}))\,(\mathbf{e}_j)|\\ &\le \|\boldsymbol{f}'(\mathbf{y}) - \boldsymbol{f}'(\mathbf{x})\|\,\|\mathbf{e}_j\|\\ &= \|\boldsymbol{f}'(\mathbf{y}) - \boldsymbol{f}'(\mathbf{x})\|\\ &< \in \text{ if } \mathbf{y} \in E \text{ and } |\mathbf{y}-\mathbf{x}| < \delta.\end{aligned}$$

Since $\boldsymbol{f}$ is differentiable, partial derivatives $D_j f_i$ exist. Taking components of vectors in (2.4.3), it follows that

$$|(D_j f_i)\,(\mathbf{y}) - D_j f_i(\mathbf{x})| < \in, \text{ if } \mathbf{y} \in E \text{ and } |\mathbf{y} - \mathbf{x}| < \delta.$$

Hence $D_j f_i$ are continuous on E for $1 \le i \le m$, $1 \le j \le n$.

Conversely, suppose that $D_j f_i$ are continuous on E for $1 \le i \le m$, $1 \le j \le n$. It is sufficient to consider one-dimensional case, i.e., the case m = 1. Fix $\mathbf{x} \in E$ and $\in > 0$. Since E is open, $\mathbf{x}$ is an interior point of E and so there is an open ball B $\subset$ E with centre at $\mathbf{x}$ and radius r. The continuity of $D_j f$ implies that r can be chosen so that

$$(2.4.4)\qquad |(D_j \boldsymbol{f})\,(\mathbf{y}) - (D_j \boldsymbol{f})\,(\mathbf{x})| < \frac{\in}{n},\ \mathbf{y} \in B,\ 1 \le j \le n.$$

Suppose $\mathbf{h} = \Sigma\, \mathbf{h}_j\, \mathbf{e}_j$, $|\mathbf{h}| < r$, and take $\mathbf{v}_0 = 0$

and $\qquad \mathbf{v}_k = h_1\, \mathbf{e}_1 + h_2\, \mathbf{e}_2 + \ldots + h_k\, \mathbf{e}_k$ for $1 \le k \le n$.

Then

$$(2.4.5)\qquad \boldsymbol{f}(\mathbf{x} + \mathbf{h}) - \boldsymbol{f}(\mathbf{x}) = \sum_{j=1}^{n} [f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + \mathbf{v}_{j-1})]$$

Since $|\mathbf{v}_h| < r$ for $1 \le k \le n$ and since s is converse, the end points $\mathbf{x} + \mathbf{v}_{j-1}$ and $\mathbf{x} + \mathbf{v}_j$ lie in s. Further, since

$$\mathbf{v}_j = \mathbf{v}_{j-1} + h_j\, \mathbf{e}_j$$

Mean Value Theorem implies

$$(2.4.6)\qquad \begin{aligned}\boldsymbol{f}(\mathbf{x} + \mathbf{v}_j) - \boldsymbol{f}(x + v_{j-1}) &= \boldsymbol{f}(\mathbf{x} + \mathbf{v}_{j-1} + h_j\, \mathbf{e}_j) - f(\mathbf{x} + \mathbf{v}_{j-1})\\ &= h_j\, \mathbf{e_j}\, (D_i\, \boldsymbol{f})\, (\mathbf{x} + \mathbf{v}_{j-1} + v_j\, h_j\, \mathbf{e}_j)\end{aligned}$$

for some $\theta \in (0, 1)$ and by (2.4.4) this differ from $h_j\, (D_i\, \boldsymbol{f})\, (\mathbf{x})$ by less than $|h_j|\frac{\in}{n}$. Hence (2.4.5) gives

$$\begin{aligned}\left|\, \boldsymbol{f}(\mathbf{x} + \mathbf{h}) - \boldsymbol{f}(x) - \sum_{j=1}^{n} h_j\, (D_i\, \boldsymbol{f})\, (\mathbf{x})\right| &\le \frac{1}{n} \sum_{j=1}^{n} |h_j| \in\\ &= |\mathbf{h}| \in\end{aligned}$$

for all $\mathbf{h}$ satisfying $|\mathbf{h}| < r$.

Hence $\boldsymbol{f}$ is differentiable at $\mathbf{x}$ and $\boldsymbol{f}'(x)$ is the linear function which assigns the number $\Sigma\, h_j(D_i\, \boldsymbol{f})\, (\mathbf{x})$ to the vector $\mathbf{h} = \Sigma\, h_j\, \mathbf{e}_j$. The matrix $[\boldsymbol{f}'(x)]$ consists of the row $(D_i\, f)\, (\mathbf{x}), \ldots, (D_n\, f)\, (\mathbf{x})$. Since $D_1\, \boldsymbol{f}$, $D_2\, \boldsymbol{f}, \ldots, D_n\, \boldsymbol{f}$ are continuous functions on E, it follows that $\boldsymbol{f}'$ is continuous and hence $\boldsymbol{f} \in C'(E)$.

2.5. Classical Theory for Functions of more than one Variable

Consider a variable u connected with the three independent variables x, y and z by the functional relation

$$u = u\,(x, y, z)$$

If arbitrary increments Δx, Δy, Δz are given to the independent variables, the corresponding increment Δu of the dependent variable of course depends upon the three increments assigned to x, y and z.

Definition. A function u = u(x, y, z) is said to be **differentiable** at the point (x, y, z) if it possesses a determinate value in the neighbourhood of this point and if.

$$\Delta u = A\, \Delta x + B\Delta y + C\Delta z + \in\rho,$$

where $\rho = |\Delta x| + |\Delta y| + |\Delta z|$, $\in \to 0$, as $\rho \to 0$ and A, B, C are independent of Δx, Δy, Δz.

In the above definition ρ may always be replaced by η, where

$$\eta = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$$

Definition. If the increment ratio

$$\frac{u(x + \Delta x, y, z) - u(x, y, z)}{\Delta x}$$

tends to a unique limit as Δx tends to zero, this limit is called the **partial derivative** of u with respect to x and is written $\frac{\partial u}{\partial x}$ or u_x.

Similarly $\frac{\partial u}{dy}$ and $\frac{\partial u}{\partial z}$ can be defined.

The differential coefficients: If in the equation

$$\Delta u = A\Delta x + B\Delta y, + C\Delta z + \in\rho$$

we suppose that $\Delta y = \Delta z = 0$, then, on the assumption that u is differentiable at the point (x, y, z),

$$\Delta u = u(x + \Delta x, y, z) - u\,(x, y, z)$$

$$= A\Delta x + \in|\Delta x|$$

and dividing by Δx,

$$\frac{u(x+\Delta x, y, z) - u(x, y, z)}{\Delta x} = A \pm \in$$

and by taking the limit as $\Delta x \to 0$, since $\in \to 0$ as $\Delta x \to 0$, we get $\frac{\partial u}{\partial x} = A$

Similarly $\frac{\partial u}{\partial y} = B$ and $\frac{\partial u}{\partial z} = C.$

Hence, when the function u = u(x, y, z) is differentiable, the partial derivatives $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}$ are respectively the differential coefficients A, B, C and so

$$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \frac{\partial u}{\partial z}\Delta z + \in\rho$$

The differential of the dependent variable du is defined to be the principal part of Δu so that the above expression may be written as

$$\Delta u = du + \in\rho.$$

Now as in the case of functions of one variable, the differentials of the independent variables are identical with the arbitrary increments of these variables. It we write u = x, u = y, u = z. respectively, it follows that

$$dx = \Delta x,\ dy = \Delta y,\ dz = \Delta z$$

Therefore, expression for du reduces to

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz$$

The distinction between derivatives and differential coefficients

We know that the necessary and sufficient condition that the function y = f(x) should be differentiable at the point x is that it possesses a finite definite derivative at that point. Thus for functions of one variable, the existence of the derivative $f'(x)$ implies the differentiability of f(x) at any given point.

For functions of more than one variable this is not true. If the function u = u (x, y, z) is differentiable at the point (x, y, z), the partial derivatives of u with respect to x, y and z certainly exist and are finite at this point, for then they are identical with differential coefficients A, B and C respectively. The partial derivatives, however, may exist at a point when the function is not differentiable at that point. In other words, the partial derivatives need not always be differential coefficients.

Example.1. Let f be a function defined by $f(x, y) = \frac{x^3 - y^3}{x^2 + y^2}$, where x and y are not simultaneously zero, $f(0, 0) = 0$.

If this function is differentiable at the origin, then, by definition,

(2.5.1) $$f(h, k) - f(0, 0) = Ah + Bk + \in\eta, \quad (1)$$

where $\eta = \sqrt{h^2 + k^2}$ and $\in \to 0$ as $\eta \to 0$.

Putting h = η cosθ, k = η sin θ in (2.5.1) and dividing through by η, we get

$$\cos^3\theta - \sin^3\theta = A\cos\theta + B\sin\theta + \in.$$

Since ∈→0 as η→0, we get, by taking the limit as η→0

$$\cos^3\theta - \sin^3\theta = A\cos\theta + B\sin\theta$$

which is impossible, since θ is arbitrary.

The function is therefore not differentiable at (0, 0). But the partial derivative exist however, for

$$f_x(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{h-0}{h} = 1$$

$$f_y(0,0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{-k}{+k} = -1.$$

Example. 2. $f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } x^2y^2 \neq 0 \\ 0 & \text{if } x = y = 0 \end{cases}$

Then $f_x(0,0) = 0 = f_y(0,0)$

and so partial derivatives exist. If it is different, then

df = f (h, k) − f(0, 0) = Ah + Bh + ∈η, where A = $f_x(0, 0)$ B = fy (0, 0),

This yields

$$\frac{h\,h}{\sqrt{h^2+k^2}} = \in\sqrt{h^2+h^2}\ ,\ \eta = \sqrt{h^2+k^2}$$

or $h_k = h^2 + k^2$

Putting k = mh we get

$$mh^2 = \in h^2\,(1+m^2)$$

or $$\frac{m}{1+m^2} = \in$$

Hence $$\lim_{k\to 0}\frac{m}{1+m^2} = 0\ ,$$

which is impossible. Hence the function is not differentiable at the origin.

Remarks:

1. Thus the information given by the existence of the two first partial derivatives is limited. The values of $f_x(x, y)$ and of $f_y(x, y)$ depend only on the values of $f(x, y)$ along two lines through the point (x, y) respectively parallel to the axes of x and y. This information is incomplete, and tells us nothing at all about the behaviour of the function $f(x, y)$ as the point (x, y) is approached along a line which is inclined to the axis of x at any given angle θ which is not equal to 0 or π/2.

2. Partial derivatives are also in general functions of x, y and z which may posses partial derivatives with respect to each of the three independent variables, we have the definition

(i) $$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) = \lim_{\Delta x\to 0}\frac{u_x(x+\Delta x,y,z)-u_x(x,y,z)}{\Delta x}$$

(ii) $$\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right) = \lim_{\Delta y\to 0}\frac{u_x(x+y+\Delta y,z)-u_x(x,y,z)}{\Delta y}$$

(iii) $$\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial x}\right) = \lim_{\Delta z\to 0}\frac{u_x(x,y,z+\Delta z)-u_x(x,y,z)}{\Delta z}$$

Provided that each of these limits exist. We shall denote the second order partial derivatives by $\frac{\partial^2 u}{\partial x^2}$ or u_{xx}, $\frac{\partial^2 u}{\partial y \partial x}$ or u_{yx} and $\frac{\partial^2 u}{\partial z \partial x}$ or u_{zx}.

Similarly we may define higher order partial derivatives of $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial z}$.

The following example shows that certain second partial derivatives of a function may exist at a point at which the function is not continuous.

Example. 3. Let $\phi(x, y) = \frac{x^3 + y^3}{x - y}$ when $x \neq y$

$\phi(x, y) = 0$ when $x = y$.

This function is discontinuous at the origin. To show this it suffices to prove that if the origin is approached along different path, $\phi(x, y)$ does not tend to the same definite limit. For, if $\phi(x, y)$ were continuous at (0, 0), $\phi(x, y)$ would tend to zero (the value of the function at the origin) by what ever path the origin were approached.

Let the origin be approached along the three curves

(i) $y = x - x^2$, (ii) $y = x - x^3$, (iii) $y = x - x^4$;

then we have

(i) $\phi(x, y) = \frac{2x^3 + 0(x^4)}{x^2} \to 0$ as $x \to 0$

(ii) $\phi(x, y) = \frac{2x^3 + 0(x^4)}{x^3} \to z$ as $x \to 0$

(iii) $\phi(x, y) = \frac{2x^3 + 0(x^4)}{x^4} \to \infty$ as $x \to 0$

Certain partial derivatives, however, exist at (0, 0), for if $\phi_{x,x}$ denote $\frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial x}\right)$ we have, for example,

$$\phi_x(0, 0) = \lim_{h \to 0} \frac{\phi(h,0) - \phi(0,0)}{h} = \lim_{h \to 0} \frac{h^2}{h} = 0,$$

$$\phi_{xx}(0, 0) = \lim_{h \to 0} \frac{\phi_x(h,0) - \phi_x(0,0)}{h} = \lim_{h \to 0} \frac{2h}{h} = 2,$$

since $\phi(x, 0) = x^2$, $\phi_x(x, 0) = 2x$ when $x \neq 0$.

The following example shows that u_{xy} is not always equal to u_{yx}.

Example. 4. Let

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$$

$f(0, 0) = 0$.

When the point (x, y) is not the origin, then

(2.5.2) $$\frac{\partial f}{\partial x} = y\left[\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^2}{(x^2 + y^2)^2}\right]$$

(2.5.3) $$\frac{\partial f}{\partial y} = x\left[\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^2}{(x^2 + y^2)^2}\right]$$

while at the origin,

(2.5.4) $f_x(0, 0) = \lim\limits_{h\to 0} \dfrac{f(h,0) - f(,0)}{h} = 0$

and similarly $f_y(0, 0) = 0$.

From (2.5.2) and (2.5.3) we see that

(2.5.5) $f_x(0, y) = -y \ (y \neq 0)$

(2.5.6) $f_y(x, 0) = x \ (x \neq 0)$

Now we have, using (2.5.4), (2.5.5) and (2.5.8)

$$f_{xy}(0, 0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = \lim \frac{h}{h} = 1$$

$$f_{yx}(0, 0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = \lim \frac{-k}{k} = -1$$

and so $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

Example. 5. Prove that the function

$$f(x, y) = (|xy|)^{1/2}$$

is not differentiable at the point (0, 0), but that $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ both exist at the origin and have the value zero.

Hence deduce that these two partial derivatives are continuous except at the origin.

Solution. We have

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0$$

$$f_y(0, 0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = 0$$

If $f(x, y)$ is differentiable at (0, 0), then we must have

$$f(h, k) = 0.h + 0.k + \in \sqrt{h^2 + k^2}$$

where $\in \to 0$ as $\sqrt{h^2 + k^2} \to 0$

Now $\in = \dfrac{|hk|^{1/2}}{\sqrt{h^2 + k^2}}$

Putting $h = \rho \cos\theta$, $k = \rho \sin\theta$, we get

$$\in = \sqrt{|\sin\theta\cos\theta|}$$

$\therefore \lim\limits_{\rho\to 0} = \in \sqrt{|\sin\theta\cos\theta|} \quad \Rightarrow \quad \sqrt{|\cos\theta\sin|} = 0$ which is impossible for arbitrary 0.

Hence, f is not differentiable.

Now suppose that $(x, y) \neq (0, 0)$. Then

$$\frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$

$$= \lim_{h\to 0} \frac{\sqrt{|(x+h)y|} - \sqrt{|xy|}}{h}$$

$$= \lim_{h\to 0} \frac{|(x+h)y| - |xy|}{h(\sqrt{x+h)y} + \sqrt{|xy|}} = \lim_{h\to 0} |y| \frac{|x+h| - |x|}{h[\sqrt{x+h} + \sqrt{|x|}]}$$

Now, we can take h so small that x + h and x have the same sign. Hence the limit is $\frac{|y|}{2\sqrt{|xy|}}$ or $\frac{1}{2}\sqrt{\frac{|y|}{|x|}}$.

Similarly $\frac{\partial f}{\partial y} = \frac{|x|}{2\sqrt{|xy|}}$ or $\frac{1}{2}\sqrt{\frac{|x|}{|y|}}$ Both of these are continuous except at (0, 0).

We now prove two theorems, the object of which is to set out precisely under what conditions it is allowable to assume that

$$f_{xy}(a, b) = f_{yx}(a, b)$$

Theorem. 9. (Young) If (i) f_x and f_y exist in the neighbourhood of the point (a, b) and (ii) f_x and f_y are differentiable at (a, b); then

$$f_{xy} = f_{yx}$$

Proof. We shall prove this theorem by taking equal increments h both for x and y and calculating $\Delta^2 f$ in two different ways, where

$$\Delta^2 f = f(a + h, b + h) - f(a + h, b) - f(a, b+h) + f(a, b).$$

Let

$$H(x) = f(x, b+h) - f(x, b)$$

Then, we have

$$\Delta^2 f = H(a + h) - H(a)$$

Since f_x exists in the neighbourhood of (a, b), the function H(x) is derivable in (a, a+h). Applying Mean Value Theorem to H(x) for $0 < \theta < 1$, we obtain

$$H(a + h) - H(a) = h\, H'(a + \theta h).$$

Therefore

(2.5.7) $$\Delta^2 f = hH'(a + \theta h)$$
$$= h[f_x(a + \theta h, b + h) - f_x(a + \theta h, b)]$$

By hypothesis (ii) of the theorem, $f_x(x, y)$ is differentiable at (a, b) so that

$$f_x(a + \theta h, b+h) - f_x(a, b) = \theta h f_{xx}(a, b) + h\, f_{yx}(a, b) + \in' h$$

and

$$f_x(a + \theta h, b) - f_x(a, b) = \theta h\, f_{xx} + \in'' h,$$

where $\in'$ and $\in''$ tend to zero as $h \to 0$. Thus, we get (on subtracting)

$$f_x(a + \theta h, b+h) - f_n(a + \theta h, b) = h f_{yx}(a, b) + h(\in' - \in'')$$

Putting this value in (1), we obtain

(2.5.8) $$\Delta^2 f = h^2 f_{yx} + \in_1 h^2,$$

where $\in_1 = \in' - \in''$,

so that $\in_1$ tends to zero with h .

Similarly, if we take

$$K(y) = f(a + h, y) - f(a, y),$$

Then we can show that

(2.5.9) $$\Delta^2 f = h^2 f_{xy} + \in_2 h^2$$

where $\in_2 \to 0$ with h.

From (2.5.8) and (2.5.9), we have

$$\frac{\Delta^2 f}{h^2} = f_{yx}(a, b) \in_1 = f_{xy}(a, b) + \in_2$$

Taking limit as $h \to 0$, we have

$$\lim_{\lambda \to 0} \frac{\Delta^2 f}{h^2} = f_{yx}(a, b) = f_{xy}(a, b)$$

which establishes the theorem.

Theorem. 10. (Schwarz). If (i) f_x, f_y, f_{yx} all exist in the neighbourhood of the point (a, b) and (ii) f_{yx} is continuous at (a, b), then f_{xy} also exists at (a, b) and $f_{xy} = f_{yx}$.

Proof. Let (a + h, b + k) be point in the neighbourhood of (a, b). Let (as in the above theorem)

$$\Delta^2 f = f(a+h, b+k) - f(a+h, h) - f(a, b+k) + f(a, b)$$

and

$$H(x) = f(x, b+k) - f(x, b),$$

so that we have

$$\Delta^2 f = H(a+h) - H(a).$$

Since f_x exists in the neighbourhood of (a, b), H(x) is derivable in (a, a+h).
Applying Mean Value Theorem for 0 < 0 < 1, we have

$$H(a+h) - H(a) = hH'(a + \theta h)$$

and therefore

$$\Delta^2 f = hH'(a+\theta h) = h\,[f_x(a+\theta h, b+k) - f_x(a+\theta h, h)].$$

Now, since $f_{yx}(x)$ exists in the neighbourhood of (a, b), the function f_x is derivable with respect to y in (b, b+k). Therefore by Mean Value Theorem, we have

$$\Delta^2 f = hk\, f_{yx}(a+\theta h, b+\theta' k)\ ,\ 0 < \theta' < 1$$

That is

$$\frac{1}{h}\left[\frac{f(a+h, b+k) - f(a+h, b)}{k} - \frac{f(a, b+k) - f(a, b)}{k}\right] = f_{yx}(a+\theta h, b+\theta' k)$$

Taking limit as k tends to zero, we obtain

$$\frac{1}{h}[(a+h, b) - f_y(a, b)] = \lim\,[f_{yx}(a+\theta h, b+\theta' k)] \tag{2.5.10}$$

$$= f_{yx}(a+\theta h, b)$$

Since f_{yx} is given to be continuous at (a, b), we have

$$f_{yx}(a+\theta h, b) = f_{yx}(a, b) + \in,$$

where $\in \to 0$ as $h \to 0$.
Hence taking limit as $h \to 0$ in (2.5.10), we have

$$\lim_{h\to 0} \frac{f_y(a+h, b) - f_y(a, b)}{h} = \lim_{h\to 0}\,[f_{yx}(a, b) + \in]$$

that is, $f_{xy}(a, b) = f_{yx}(a, b)$

This completes the proof of the theorem

Remark. The condition of Young or Schwarz's Theorem are sufficient for $f_{xy} = f_{yx}$ but they are not necessary. For example, consider the function

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

We have $$f_x(0, 0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0$$

$$f_y(0, 0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = 0$$

Also, for $(x, y) \neq (0, 0)$, we have

$$f_x(x, y) = \frac{(x^2+y^2)2xy^2 - x^2y^2.2x}{(x^2+y^2)^2} = \frac{2xy^4}{(x^2+y^2)^2}$$

$$f_y(x, y) = \frac{2x^4y}{(x^2y^2)^2}$$

Again $$f_{yx}(0, 0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = 0 \text{ and } f_{xy}(0, 0) = 0$$

So that $f_{yx}(0, 0) = f_{xy}(0, 0)$

For $(x, y) \neq (0, 0)$, we have $f_{yx}(x, y) = \dfrac{8xy^3(x^2+y^2)^2 - 2xy^4.4y(x^2+y^2)}{(x^2+y^2)^4}$

$$= \frac{8x^3y^3}{(x^2+y^2)^2}$$

Putting y = mx, we can show that

$$\lim_{(x,y)\to(0,0)} f_{yx} \neq 0 = f_{yx}(0, 0)$$

so that f_{xy} is not continuous at (0, 0). Thus the condition of Schwarz's theorem is not satisfied.

To see that conditions of Young's Theorem are also not satisfied, we notice that

$$f_{xx}(0,0) = \lim_{h\to 0} \frac{f_x(h,0) - f_x(0,0)}{h} = 0$$

If f_x is differentiable at (0, 0) we should have

$$f_x(h, k) - f_x(0, 0) = f_{xx}(0, 0).\, h + f_{yx}(0, 0).\, k + \in\eta$$

$$\frac{2hk^4}{(h^2+k^2)^2} = \in \eta, \text{ where } \in\to 0 \text{ as } \eta\to 0.$$

Put $h = \rho \cos\theta$, $k = \rho \sin\theta$, then $\eta = \sqrt{h^2+k^2} = \rho$
so we have

$$\frac{2\rho\cos\theta.\rho^4\sin^4\theta}{\rho^4} = \in\rho$$

i.e. $2\cos\theta \sin^4\theta = \in$
Taking limit as $\rho\to 0$, we have

$$2\cos\theta\sin^4\theta = 0$$

which is impossible for arbitrary θ

Euler's theorem on homogeneous functions.

Definition. A function $f(x, y, z,\ldots)$ is a homogeneous function of degree n if it has the property

$$f(tx, ty, tz,\ldots) = t^n f(x, y, z,\ldots) \qquad (1)$$

(1) To prove Euler's theorem, write $x' = t_x$, $y' = t_y,\ldots$, then

$$f(x', y', z',\ldots) = t^n f(x, y, z,\ldots) \qquad (2)$$

and if we take partial derivatives with respect to t, we get

$$\frac{\partial f}{\partial x}.\frac{\partial x'}{\partial t} + \frac{\partial f}{\partial y'}.\frac{\partial y'}{\partial t} + \ldots = n\, t^{n-1} f(x, y, z),$$

that is,

$$x\frac{\partial f}{\partial x'} + y\frac{\partial f}{\partial y'} + \ldots = nt^{n-1} f(x, y, z)$$

Now put t = 1, so that $x' = x$, $y' = y,\ldots$ and we get

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + \ldots = nf(x, y,) \qquad (3)$$

which is known as **Euler's theorem.**

(2) Differentiate the equation (2) m times; since t is the only independent variable, we have

$$d^m f = n(n-1)\ldots(n-m+1)\, t^{n-m} f(x, y, z,\ldots)\, dt^m$$

Now $\qquad d^m f = \dfrac{\partial^m f}{\partial x'^m} dx'^m + \dfrac{\partial^m f}{\partial y'^m} dy'^m + \ldots,$

and since $x'^m = t^m x^m$, $y'^m = t^m y^m,\ldots$, when t = 1 we get

$$\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+...\right)^m f(x, y, z) = n(n-1)\ldots(n-m+1)\, f(x, y, z,\ldots)$$

which is the generalization of Euler's theorem.

2.6. In view of Taylor's theorem for functions of one variable, it is not unnatural to expect the possibility of expanding a function of more than one variable $f(x + h, y + k, z + l,)$ in a series of ascending powers of h, k, l, To fix the ideas, consider a function of two variables only; the reasoning in the general case is precisely the same.

Taylor's Theorem. If $f(x, y)$ and all its partial derivatives of order n are finite and continuous for all point (x, y) in the domain $a \le x \le a + h$, $b \le y \le b + k$, then

$$f(a + h, b+k) = f(a, b) + df(a, b) + \frac{1}{\lfloor 2}d^2 f(a, b) +\ldots+ \frac{1}{\lfloor n-1}d^{n-1} f(a, b) + R_n$$

where $$R_n = \frac{1}{\lfloor n}d^n f(a\, \theta h, b + \theta k),\ 0 < 0 < 1.$$

Proof. Consider a circular domain of centre (a, b) and radius large enough for the point (a + h, b + k) to be also with in the domain. Suppose that $f(x, y)$ is a function such that all the partial derivatives of order n of $f(x, y)$ are continuous in the domain. Write

$$x = a + ht,\ y = b+kt,$$

so that, as t ranges from 0 to 1, the point (x, y) moves along the line joining the point (a, b) to be point (a +h, b+k); then

$$f(x, y) = f(a + ht, b + kt) = \phi(t).$$

Now, $\phi'(t) = \frac{\partial f}{\partial x}.\frac{dx}{dt}+\frac{\partial f}{\partial y}.\frac{dy}{dt}=h\frac{\partial f}{\partial x}+k\frac{\partial f}{\partial y}=df,$

and similarly

$$\phi''(t) = d^2 f,\ldots,\ \phi^{(n)}(t) = d^n(f)..$$

We thus see that $\phi(t)$ and its n derivatives are continuous functions of t in the interval $0 \le t \le 1$, and so, by Maclaurin's theorem

$$\phi(t) = \phi(0) + t\phi'(0) + \frac{t^2}{\lfloor 2}\phi''(0) +\ldots+ \frac{t^n}{\lfloor n}\phi^{(n)}(\theta t),$$

where $0 < \theta < 1$. Now put t = 1 and observe that

$$\phi(1) = f(a + h, b+k),\ \phi(0) = f(a, b),\ \phi'(0) = d\, f(a, b)$$
$$\phi''(0) = d^2 f(a, b),\ldots.,\ \phi^{(n)}(\theta\, t) = d^n f(a + \theta h, h + \theta k).$$

It follows immediately that

(2.5.11) $$f(a + h, b + k) = f(a, b) + d\,f(a, b) + \frac{1}{\lfloor 2}d^2 f(a, b) +\ldots + \frac{1}{\lfloor n-1}d^{n-1} f(a, b) + R_n$$

where $R_n = \frac{1}{\lfloor n}d^n f(a + \theta h, b + \theta k),\ 0 < \theta < 1.$

We have assumed here that all the partial derivatives of order n are continuous in the domain in question. Taylor expansion does not necessarily hold if these derivatives are not continuous.

Maclaurin's theorem. If we put a = b = 0, h = x, k = y, we get at once, from the equation

$$f(a + h, b + k) = f(a, b) + df(a, b) + \frac{1}{\lfloor 2}d^2 f(a, b) + + \frac{1}{\lfloor n-1}d^{n-1} f(a, b) + R_n$$

where $R_n = \frac{1}{\lfloor n} d^n f(a + \theta h, b + \theta k), 0 < 0 < 1,$

that $f(x, y) = f(0, 0) + df(0, 0) + \frac{1}{\lfloor 2} d^2 f(0, 0) + \ldots + \frac{1}{\lfloor n-1} d^{n-1} f(0, 0) + R_n$

where $R_n = \frac{1}{\lfloor n} d^n f(0x, \theta h), 0 < \theta < 1.$

The theorem easily extend to any number of variables.

Example. 6. If $f(x, y) = (|x\ y|)^{1/2}$, prove that Taylor's expansion about the point (x, x) is not valid in any domain which includes the origin. Give reasons.

Solution. If a Taylor expansion were possible (n = 1)

$$f(x + h, x + h) = f(x, x) + h\ \{f_x(\xi, \xi) + f_y(\xi, \xi)\}$$

where $x < \xi < x + h$. This is not valid for all x, h for it implies that

$$|x + \lambda| = |x| \pm h, \xi \neq 0$$
$$= |x|, \xi = 0$$

(The reason is that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial z}$ are not continuous at the origin).

2.7. Implicit functions

Let $F(x_1, x_2, \ldots, x_n, u) = 0$ (1)

be a functional relation between the n + 1 variables $x_1, \ldots, x_n$, u and let $x = a_1, x_2 = a_2, \ldots, x_n = a_n$ be a set of values such that the equation

$$F(a_1, \ldots, a_n, u) = 0 \quad (2)$$

is satisfied for at least one value of u, that is the equation (2) in u has at least one root. We may consider u as a function of the x's : $u = \phi(x_1, x_2, \ldots, x_n)$ defined in a certain domain, where $\phi(x_1, x_2, \ldots, x_n)$ has assigned to it at any point $(x_1, x_2, \ldots, x_n)$ the roots u of the equation (1) at this point. We say that u is the implicit function defined by (1). It is, in general, a many valued function.

More generally, consider the set of equations

$$F_p\ (x_1, \ldots, x_n, u_1, \ldots, u_m) = 0\ (p = 1, 2, \ldots, m) \quad (3)$$

between the n +m variables $x_1, \ldots, x_n, u_1, \ldots, u_m$ and suppose that the set of equations (3) are such that there are points $(x_1, x_2, \ldots, x_n)$ for which these m equations are satisfied for at least one set of values $u_1, u_2, \ldots, u_m$. We may consider the u's as functions of the x's,

$$u_p = \phi_p\ (x_1, x_2, \ldots, x_n)\ (p = 1, 2, \ldots, m)$$

where the functions ϕ have assigned to them at the point $(x_1, x_2, \ldots, x_n)$ the values of the roots $u_1, u_2, \ldots, u_m$ at this point. We say that $u_1, u_2, \ldots, u_m$ constitute a system of implicit functions defined by the set of equation (3). These functions are in general many valued.

Theorem. 12 (Existence Theorem). Let F(u, x, y) be a continuous function of the variables u, x, y. Suppose that

(i) $F(u_0, a, b) = 0$

(ii) F(u, x, y) is differentiable at (u_0, a, b)

(iii) The partial derivative $\frac{\partial F}{\partial u}(u_0, a, b) \neq 0$

Then there exists at least one function u = u (x, y) reducing to u_0 at the point (a, b) and which, in the neighbourhood of this point, satisfies the equation F (u, x, y) = 0 identically.

Also, every function u which possesses these two properties is continuous and differentiable at the point (a, b).

Proof. Since $F(u_0, a, b) = 0$ and $\frac{\partial F}{\partial u}(u_0, a, b) \neq 0$, the function F is either an increasing or decreasing function of u when $u = u_0$. Thus there exists a positive number δ such that $F(u_0 - \delta, a, b)$ and $F(u_0 + \delta, a,b)$ have opposite signs. Since F is given to be continuous, a positive number η can be found so that the functions

$$F(u_0 - \delta, x, y) \text{ and } F(u_0 + \delta, x, y)$$

the values of which may be as near as we please to

$$F(u_0 - \delta, a, b) \text{ and } F(u_0 + \delta, a, b)$$

will also have opposite signs so long as $|x - a| < \eta$ and $|y - b| < \eta$.

Let x, y be any two values satisfying the above conditions. Then $F(u, x, y)$ is a continuous function of u which changes sign between $u_0 - \delta$ and $u_0 + \delta$ and so vanishes somewhere in this interval. Thus for these x and y there is a u in $[u_0 - \delta, u_0 + \delta]$ for which $F(u, x, y) = 0$. This u is a function of x and y, say $u(x, y)$ which reduces to u_0 at the point (a, b).

Suppose that Δu, Δx, Δy are the increments of such function u and of the vanishes x and y measured from the point (a, b). Since F is differentiable at (u_0, a, b), we have

$$\Delta F = [F_u(u_0, a, b) + \in] \Delta u + [F_x(u_0, a, b) + \in'] \Delta x + [F_y(u_0, a, b) + \in''] \Delta y = 0$$

since $\Delta F = 0$ because of $F = 0$. The numbers $\in, \in', \in''$ tend to zero with Δu, Δx and Δy and can be made as small as we please with δ and η. Let δ and η be so small that the numbers $\in, \in', \in''$ are all less than $\frac{1}{2}|F_u(u_0, a, b)|$, which is not zero by our hypothesis. The above equation then shows that $\Delta u \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$ which means that the function $u = u(x, y)$ is continuous at (a, b).

Moreover, we have

$$\Delta u = -\frac{[F_x(u_0, a, b) + \in']\Delta x + [F_y(u_0, a, b) + \in'']\Delta y}{F_u(u_0, a, b) + \in}$$

$$= -\frac{F_x(u_0, a, b)}{F_u(u_0, a, b)}\Delta x - \frac{F_y(u_0, a, b)}{F_u(u_0, a, b)}\Delta y + \in_1 \Delta x + \in_2 \Delta y,$$

$\in_1$ and $\in_2$ tending to zero as Δx and Δy tend to zero.

Hence u is differentiable at (a, b).

Cor.1. If $\frac{\partial F}{\partial u}$ exists and is not zero in the neighbourhood of the point (u_0, a, b), the solution u of the equation $F = 0$ is unique.

Suppose that there are two solutions u_1 and u_2. Then we should have, by Mean Value Theorem, for $u_1 < u' < u_2$

$$\theta = F(u_1, x, y) - F(u_2, x, y) = (u_1 - u_2) F_u(u', x, y),$$

and so $F_u(u, x, y)$ would vanish at some point in the neighbourhood of (u_0, a, b) which is contary to our hypothesis

Cor. 2. If $F(u, x, y)$ is differentiable in the neighbourhood of (u_0, a, b), the function $u = u(x, y)$ is differentiative in the neighbourhood of the point (a, b).

This is immediate, because the preceding proof is then applicable at every point (u, x, y) in that neighbourhood.

Corollary 1 is of great importance, for by considering a function of wo variables only, $F(u, x) = 0$, and taking $F(u, x) = f(u) - x$, we can enunciate the fundamental theorem on inverse functions as follows.

Theorem. 13 (Inverse Function Theorem). If, in the neighbourhood of $u = u_0$, the function $f(u)$ is a continuous function of u, and if (i) $f(u_0) = a$, (ii) $f'(u) \neq 0$ in the neighbourhood of the point $u = u_0$, then there exists a unique continuous function $u = \phi(x)$, which is equal to u_0 when $x = a$, and which satisfied identically the equation

$$f(u) - x = 0,$$

in the neighbourhood of the point $x = a$.

The function $u = \phi(x)$ thus defined is called the inverse function of $x = f(u)$.

2.8. Extreme Values

Definition. A function $f(x, y, z)$ of several independent variables x, y, z,… is said to have an extreme value at the point (a, b, c,…) if the increment

$$\Delta f = f(a + h, b + k, c + l) - f(a, b, c)$$

preserves the same sign for all values of h, k, l, whose moduli do not exceed a sufficiently small positive number η.
If Δf is negative, then the extreme value is a **maximum** and if Δf is positive it is a **minimum**.

Now we find necessary and sufficient conditions for extreme values. We will consider a function of two independent variables.
By Taylor's theorem we have

$$f(x + h, y + k,\ldots) - f(x, y, \ldots) = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} + \ldots + \text{terms of the second and higher orders.}$$

Now by taking h, k, l, sufficiently small, the first degree terms can be made to govern the sign of the right hand side and therefore of the left side also, of the above equation, therefore by changing the sign of h, k, l, the sign of the left hand member would be changed. Hence as a first condition for the extreme value we must have

$$h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} + l\frac{\partial f}{\partial z} + \ldots = 0,$$

and since these arbitrary increments are independent of each other, we must have

$$\frac{\partial f}{\partial x} = 0,\ \frac{\partial f}{\partial y} = 0,\ \frac{\partial f}{\partial z} = 0,\ \ldots$$

which are necessary conditions for extreme points. These conditions are not sufficient for extreme points.

To find sufficient conditions we will consider only the case of two variables.

Let f be a real valued function of two variables. Let (a, b) be an interior point of the domain of f such that f admits of second order continuous partial derivatives in this neighbourhood. We suppose that $f_x(a, b) = 0 = f_y(a, b)$.

We write r, s, t for the values of $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y^2}$ respectively when x = a and y = b. That is,

$$f_{x,x}(a, b) = r,\ f_{x,y}(a, b) = S,\ f_{y,y}(a, b) = t$$

If (a + h, b + k) is any point of neighbourhood of (a, b), then by Taylor's theorem we have

$$f(a\ \ h, b + k) - f(a, b) = h f_x(a, b) + k f_y(a, b)$$
$$+ \frac{1}{2}[h^2 f_{x,x}(a, b) + 2 f_{xy}(a, b)\, hk + k^2 f_{y,y}(a, b)]$$
$$+ R_3 = \frac{1}{2}[h^2 + 25hk + tk^2] + R_3 \quad \left(\because\ \frac{\partial f}{\partial x}(a,b) = \frac{\partial f}{\partial y}(a,b) = 0\right)$$

where R_3 consists of terms of the third and higher orders of small quantities, and by taking h and k sufficiently small the second degree terms now can be made to govern the sign of the right hand side and therefore of the left hand side also. If these terms are of permanent sign for all such values of h and k, we shall have a maximum or minimum for $f(x, y,\ldots)$ according as that sign is negative or positive.

Now condition for the invariable sign of $(r\,h^2 + 25hk + tk^2)$ is that $rt - S^2$ shall be positive and the sign will be that of r. If $rt - S^2$ is positive, it is clear that r and t must have the same sign.
Thus, if $rt - S^2$ is positive we have a maximum or minimum according as r and t are both negative or both positive.
This condition was first pointed out by Lagrange and is known as **Lagrange's condition**.
If, However, $rt = S^2$, the quadratic terms

$$rh^2 + 2shk + tk^2 \text{ becomes } \frac{1}{r}(hr + ks)^2 \qquad (*)$$

and are therefore of the same sign as r or t unless

$$\frac{h}{k} = -\frac{S}{r} = B \text{ say for which * vanishes}$$

In this case we must consider terms of higher degree in the expansion $f(a + h, h + k) - f(a, b)$. The cubic term must vanish collectively when $\frac{h}{k} = \beta$; otherwise, by changing the sign of both h and k we could change the sign of $f(a + h, b + k) - f(a, b)$. And the biquadratic terms must collectively be of the same sign as r and t when $\frac{h}{k} = \beta$.

If r = 0, S ≠ 0, * changes sign with k and there is no extreme value. If r = 0 = S * does not change sign but it vanishes where h = 0 (without h = 0). This is a doubtful case.

In the case in which x, s, t are each of them zero, the quadratic terms are altogether absent, and the cubic terms would change sign with h and k and therefore all the differential coefficients of the third order must vanish separately when x = a and y = b and the biquadratic terms must be such that they retain the same sign for all sufficiently small values of h, k.

Therefore we may state that
The value $f(a, b)$ is an extreme value of $f(x, y)$ if

$$f_x(a, b) = f_y(a, b) = 0$$

and if
$$f_{xx}(a, b)\, f_{yy}(a, b) > f_{xy}^2(a, b)$$

and the value is maximum or a minimum according as f_{xx} (or f_{yy}) is negative or positive.

Example.7. Let $u = xy + \frac{a^3}{x} + \frac{a^3}{y}$,

$$\left.\begin{aligned} \frac{\partial u}{\partial x} &= y - \frac{a^3}{x^2} = 0 \\ \frac{\partial u}{\partial y} &= x - \frac{a^3}{b^2} = 0 \end{aligned}\right] \text{putting } x = a,\ y = a$$

Hence $$\frac{\partial^2 u}{\partial x^2} = \frac{2a^3}{x^3} = 2, \frac{\partial^2 u}{\partial x \partial y} = 1, \quad \frac{\partial^2 u}{\partial y^2} = \frac{2a^3}{y^3} = 2$$

Therefore r and t are +ve when x = a = y and $rt - s^2 = 2.2 - 1 = 3$ (+ve) therefore there is a minimum value of u, viz. $u = 3a^2$.

Example. 8. Let

$$f(x, y) = y^2 + x^2 y + x^4.$$

It can be verified that

$$f_x(0, 0) = 0, f_y(0, 0) = 0$$
$$f_{xx}(0, 0) = 0, f_{yy}(0, 0) = 2$$
$$f_{xy}(0, 0) = 0.$$

So, at the origin we have

$$f_{xx} f_{yy} = f_{xy}^2.$$

However, on writing

$$y^2 + x^2y + x^4 = (y + \frac{1}{2}x^2)^2 + \frac{3x^4}{4},$$

it is clear that $f(x, y)$ has a **minimum value at the origin**, since

$$\Delta f = f(h, k) - f(0, 0) = \left(k + \frac{h^2}{2}\right)^2 + \frac{3h^4}{4}$$

is greater than zero for all values of h and k.

Example. 9. Let

$$f(x, y) = 2x^4 - 3x^2 y + y^2$$

Then $$\frac{\partial f}{\partial x} = 8x^3 - 6xy \quad \Rightarrow \frac{\partial f}{\partial x}(0,0) = 0; \quad \frac{\partial f}{\partial y} = -3x^2 + 2y \quad \Rightarrow \quad \frac{\partial f}{\partial y}(0,0) = 0$$

$$r = \frac{\partial^2 f}{\partial x^2} = 24x^2 - 6y = 0 \text{ at } (0, 0),\ S = \frac{\partial^2 f}{\partial x \partial y} = -6x = 0 \text{ at } (0, 0)$$

$$t = \frac{\partial^2 f}{\partial y^2} = 2. \text{ Thus } rt - S^2 = 0. \text{ Thus it is a doubtful case}$$

However, we can write $f(x, y) = (x^2 - y)(2x^2 - y)$, $f(0, 0) = 0$

$$f(x, y) - f(0, 0) = (x^2 - y)(2x^2 - y) > 0 \text{ for } y < 0 \text{ or } x^2 > y > 0$$

$$< 0 \text{ for } y > x^2 > \frac{y}{2} > 0$$

Thus Δf does not keep the some sign mean (0, 0). Therefore it does not have maximum or minimum at(0, 0).

2.9. Lagrange's Method of Undermined Multipliers

Let $u = \phi(x_1, x_2, x_n)$ be a function of n variables which are connected by m equations

$$f_1(x_1, x_2, \ldots, x_n) = 0,\ f_2(x_1, x_2, \ldots, x_n) = 0,\ \ldots,\ f_m(x_1, x_2, \ldots, x_n) = 0,$$

so that only n–m of the variables are independent.

When u is a maximum or minimum

$$du = \frac{\partial u}{\partial x_1} dx_1 + \frac{\partial u}{\partial x_2} dx_2 + \frac{\partial u}{\partial x_3} dx_3 + \ldots + \frac{\partial u}{\partial x_n} dx_n = 0$$

Also

$$df_1 = \frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 + \frac{\partial f_1}{\partial x_3} dx_3 + \ldots + \frac{\partial f_1}{\partial x_n} dx_n = 0$$

$$df_2 = \frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 + \frac{\partial f_1}{\partial x_3} dx_3 + \ldots + \frac{\partial f_1}{\partial x_n} dx_n = 0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$df_m = \frac{\partial f_m}{\partial x_1} dx_1 + \frac{\partial f_m}{\partial x_2} dx_2 + \frac{\partial f_m}{\partial x_3} dx_3 + \ldots + \frac{\partial f_m}{\partial x_n} dx_n = 0$$

Multiplying these lines respectively by 1, $\lambda_1, \lambda_2, \ldots, \lambda_n$ and adding, we get a result which may be written

$$P_1\, dx_1 + P_2\, dx_2 + P_3\, dx_3 + \ldots + P_n\, d_{xn} = 0\ ,$$

where

$$P_r = \frac{\partial u}{\partial x_r} + \lambda_1 \frac{\partial f_1}{\partial x_r} + \lambda_2 \frac{\partial f_2}{\partial x_r} + \ldots + \lambda_m \frac{\partial f_m}{\partial x_r}$$

The m, quantities $\lambda_1, \lambda_2, \ldots, \lambda_m$ are at our choice. Let us choose them so as to satisfy the m linear equations

$$P_1 = P_2 = P_3 \ldots = P_m = 0$$

The above equation is now reduced to

$$P_{m+1}\, dx_{m+1} + P_{m+2}\, dx_{m+2} + \ldots + P_n\, dx_n = 0$$

It is indifferent which n – m of the n variables are regarded as independent. Let them be $x_{m+1}, x_{m+2}, \ldots, x_n$. Then since n–m quantities $dx_{m+1}, dx_{m+2}, \ldots, dx_n$ are all independent their coefficients must be separately zero. Thus we obtain the additional n–m equations

$$P_{m+1} = P_{m+2} = \ldots = P_n = 0$$

Thus the m + n equations

$$f_1 = f_2 = f_3 = \ldots = f_m = 0$$

and

$$P_1 = P_2 = p_3 = \ldots = P_n = 0$$

determines the m multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and values of the n variables $x_1, x_2, \ldots, x_n$ for which maxima and minima values of u are possible.

Example. 10. Find the length of the axes of the section of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ by the plane $lx + my + nz = 0$

Solution. We have to find the extreme values of the function r^2 where $r^2 = x^2 + y^2 + z^2$, subject to the two equations of condition

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0,$$

$$lx + my + nz = 0$$

Then $x\,dx + y\,dy + z\,dz = 0$ (1)

$$\frac{x}{a^2}dx + \frac{y}{b^2}dy + \frac{z}{c^2}dz = 0 \qquad (2)$$

$l\,dx + m dy + n dz = 0$ (3)

Multiplying these equation by 1, λ_1 and λ_2 and adding we get

$$x + \lambda_1 \frac{x}{a^2} + \lambda_2 l = 0 \qquad (4)$$

$$y + \lambda_1 \frac{y}{b^2} + \lambda_2 m = 0 \qquad (5)$$

$$z + \lambda_1 \frac{z}{c^2} + \lambda_2 n = 0 \qquad (6)$$

Multiplying (4), (5) and (6) by x, y, z and adding we get

$$(x^2 + y^2 + z^2) + \lambda_1 \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) + \lambda_2 (lx, my + nz) = 0$$

or $r^2 + \lambda_1 = 0 \quad \Rightarrow \lambda_1 = -r^2$

Hence from (4) (5) and (6) we have

$$x = \frac{\lambda_2 l}{\left(\frac{r^2}{a^2} - 1\right)}, y = \frac{\lambda_2 m}{\left(\frac{r^2}{b^2} - 1\right)}, z = \frac{\lambda_2 n}{\left(\frac{r^2}{c^2} - 1\right)}$$

But $lx + my + nz = 0$ $\Rightarrow \lambda_2 \left(\frac{l^2a^2}{r^2 - a^2} + \frac{m^2b^2}{r^2 - b^2} + \frac{n^2c^2}{r^2 - c^2}\right) = 0$ and since $\lambda_2 \neq 0$, the equation giving the values of r^2, which are the squares the length of the semi-axes required in the quadratic in r^2 is

$$\frac{l^2a^2}{r^2 - a^2} + \frac{m^2b^2}{r^2 - b^2} + \frac{n^2c^2}{r^2 - c^2} = 0$$

Example. 11. Investigate the maximum and minimum radii vector of the sector of "surface of elasticity" $(x^2 + y^2 + z^2)^2 = a^2 x^2 + b^2 y^2 + c^2 z^2$ made by the plane $lx + my + nz = 0$

Solution. We have

$x dx + y dy + z dz$ (1)

$a^2 x dx + b^2 y dy + c^2 z dz = 0$ (2)

and $l dx + m dy + n dz = 0$ (3)

Multiplying these equations by 1, λ_1, λ_2 respectively and adding we get

$x + a^2 x\lambda_1 + l\lambda_2 = 0$ (4)

$y + b^2 y\lambda_1 + m\lambda_2 = 0$ (5)

$z + c^2\, 2\lambda_1 + n\lambda_2 = 0$ (6)

Multiplying by x, y, z respectively and adding we get

$$(x^2 + y^2 + z^2) + (a^2 x^2 + b^2y^2 + c^2 z^2)\, \lambda_1 + (ln + my + nz)\, \lambda_2 = 0$$

$\Rightarrow$ $r^2 + \lambda_1 r^4 = 0$ $\Rightarrow$ $\lambda_1 = -\dfrac{1}{r^2}$

$$\Rightarrow \quad x = \frac{\lambda_2\, l r^2}{a^2 - r^2}, y = \frac{\lambda_2 m r^2}{b^2 - r^2}, z = \frac{\lambda_2 n r^2}{c^2 - r^2}$$

Then $lx + my + nz = 0$ $\Rightarrow$ $\dfrac{\lambda_2\, l^2 r^2}{a^2 - r^2}, y = \dfrac{\lambda_2 m^2 r^2}{b^2 - r^2}, z = \dfrac{\lambda_2 n^2 r^2}{c^2 - r^2} = 0$

$$\Rightarrow \quad \frac{l^2}{a^2 - r^2} + \frac{m^2}{b^2 - r^2} + \frac{n^2}{c^2 - r^2} = 0$$

It is a quadratic in r and gives its required values.

Example. 12. Prove that the volume of the greatest rectangular parallelopiped that can be inscribed in the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \text{ is } \frac{8abc}{3\sqrt{3}}.$$

Solution. Volume of a parallelepiped is = 8xyz. Its maximum value is to be find under the condition that it is inscribed in the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$. We have

$$u = 8xyz$$

$$f_1 = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Therefore

$$du = 8yz\,dx + 8xz\,dy + 8xy dz = 0 \qquad (1)$$

$$df_1 = \frac{x^2}{a^2}dx + \frac{y^2}{b^2}dy + \frac{z^2}{c^2}dz = 0 \qquad (2)$$

Multiplying (1) by 1 and (2) by λ and adding we get

$$yz \quad \frac{x}{a^2}\lambda = 0 \qquad (3)$$

$$zx + \frac{y^2}{b^2}\lambda = 0 \qquad (4)$$

$$zy + \frac{z}{c^2}\lambda = 0 \qquad (5)$$

From (3), (4) and (5) we get

$$\lambda = \frac{a^2yz}{x} = \frac{b^2zx}{y} = \frac{c^2xy}{z}$$

and so

$$\frac{a^2yz}{x} = \frac{b^2zx}{y} = \frac{c^2xy}{z}$$

Dividing throughout by x, y, z we get

$$\frac{a^2}{a^2}=\frac{b^2}{y^2}=\frac{c^2}{z^2}$$

Hence $\dfrac{3x^2}{a^2}$ = 1 or x = $\dfrac{a}{\sqrt{3}}$. Similarly y = $\dfrac{b}{\sqrt{3}},\ z=\dfrac{c}{\sqrt{3}}$

It follows therefore that

$$u = 8\,xyz = \frac{8abc}{3\sqrt{3}}$$

Example 13. Find the point of the circle $x^2 + y^2 + z^2 = k^2$, $lx + my + nz = 0$ at which the function $u = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$ attain its greatest and its least value.

Solution. We have

$$u = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$$

$$f_1 = lx + my + nz = 0$$

$$f_2 = x^2 + y^2 + z^2 = k^2$$

Then
$$ax\,dx + by\,dy + czdz + fy\,dz + fz\,dy + gz\,dx + gx\,dz + hx\,dy + hydx = 0$$

$$l\,dx + mdy + ndz = 0$$

$$x\,dx + ydy + zdz = 0$$

Multiplying by 1, λ_1, λ_2 respectively and adding

$$ax + ky,\ gz + \lambda_1 l + \lambda_2 x = 0$$

$$by + hx + fz + \lambda_1 m + \lambda_2 y = 0$$

$$cz + gx + fy + \lambda_1 n + \lambda_2 z = 0.$$

Multiplying by x, y, z respectively and adding we get

$$u + \lambda_2 = 0 \qquad \Rightarrow \lambda_2 = -u$$

Putting this value in the above equation we have

$$x(a-u) + hy + gz + l\lambda_1 = 0$$

$$hx + y(b-u) + fz + m\lambda_1 = 0$$

$$gx + fy + z(c-u) + n\,\lambda_1 = 0$$

$$lx + my + zx + 0 = 0$$

Eliminating x, y, z and λ_1 we get

$$\begin{vmatrix} a-u & h & g & l \\ h & b-u & f & m \\ g & f & c-u & n \\ l & m & n & o \end{vmatrix} = 0 \text{ Ans.}$$

Example. 14. If a, b, c are positive and

$$u = (a^2x^2 + b^2y^2 + c^2z^2)/x^2y^2z^2,\ ax^2 + by^2 + cz^2 = 1,$$

show that a stationary value of u is given by

$$x^2 = \frac{\mu}{2a(\mu+a)},\ y^2=\frac{\mu}{2b(\mu+b)},\ z^2=\frac{\mu}{2c(\mu+c)},$$

where μ is the +ve root of the cubic

$$\mu^3 - (bc + ca + ab)\,\mu - 2abc = 0$$

Solution. We have

$$u = \frac{a^2x^2 + b^2y^2 + c^2z^2}{x^2y^2z^2} \quad (1)$$

$$ax^2 + by^2 + cz^2 = 1 \quad (2)$$

Differentiating (1), we get

$$\Sigma \frac{1}{x^3}\left(\frac{b^2}{z^2} + \frac{c^2}{y^2}\right) dx = 0$$

which on multiplication with $x^2\, y^2\, z^2$ yields

$$\Sigma \frac{1}{x}(b^2y^2 + c^2z^2)\, dz = 0 \quad (3)$$

Differentiating (2) we have

$$\Sigma\, ax\, dx = 0 \quad (4)$$

Using Lagrange's multiplier we obtain

$$\frac{1}{x}(b^2\, y^2 + c^2\, z^2) = \mu\, ax$$

i.e. $$b^2y^2 + c^2z^2 = \mu\, ax^2 \quad (5)$$

Similarly $$c^2z^2 + a^2x^2 = \mu\, by^2 \quad (6)$$

$$a^2x^2 + b^2y^2 = \mu\, c\, z^{2+} \quad (7)$$

Then (6) + (7) – (5) yields

$$2a^2\, x^2 = \mu(by^2 + cz^2 - ax^2)$$

$$= \mu(a - 2ax^2) \text{ by (2)}$$

Therefore

$$\therefore \quad 2a\,(a + \mu)\, x^2 = \mu$$

$$\Rightarrow \quad x^2 = \frac{\mu}{2a(a+\mu)}$$

Similarly $y^2 = \dfrac{\mu}{2b(b+\mu)}$ and $z^2 = \dfrac{\mu}{2c(\mu+c)}$

Substituting these values of x^2, y^2 and z^2 in (2) we obtain

$$\frac{\mu}{2(a+\mu)} + \frac{\mu}{2(b+\mu)} + \frac{\mu}{2(c+\mu)} = 1$$

which is equal to

$$\mu^3 - (bc + ca + ab)\,\mu - 2\,abc = 0 \quad (8)$$

Since a, b, c are +ve, any one of (5), (6), (7) shows that μ must be +ve. Hence μ is the +ve root (8)

2.10. Jacobians

If $u_1, u_2, \ldots, u_n$ be n functions of the n variables $x_1, x_2, x_3, \ldots, x_n$ the determinant

$$\begin{vmatrix} \frac{\partial u_1}{\partial x_1}, \frac{\partial u_1}{\partial x_2},\ldots, \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1}, \frac{\partial u_2}{\partial x_2},\ldots, \frac{\partial u_2}{\partial x_n} \\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots \\ \frac{\partial u_n}{\partial x_1}, \frac{\partial u_n}{\partial x_2},\ldots, \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

is called the **Jacobian** of $u_1, u_2,\ldots, u_n$ with regard to $x_1, x_2,\ldots, x_n$. This determinant is often denoted by

$$\frac{\partial(u_1, u_2,\ldots, u_n)}{\partial(x_1, x_2,\ldots, x_n)}, \qquad J\,(u_1, u_2,\ldots, u_n)$$

or shortly J, when there can be no doubt as to the variables referred to

Theorem. 14. If $u_1, u_2,\ldots, u_n$ are n differentiable functions of the n independent variables $x_1, x_2,\ldots,x_n$, and there exists an identical differentiable functional relation $\phi(u_1, u_2, \ldots, u_n) = 0$ which does not involve the x's explicity, then the Jacobian

$$\frac{\partial(u_1, u_2,\ldots, u_n)}{\partial(x_1, x_2,\ldots, x_n)}$$

vanishes identically provided that ϕ, as a function of the u's has no stationary values in the domain considered.

Proof. Since

$$\phi(u_1, u_2,\ldots, u_n) = 0,$$

we have

$$\frac{\partial \phi}{\partial u_1} du_1 + \frac{\partial \phi}{\partial u_2} du_2 + \ldots + \frac{\partial \phi}{\partial u_n} du_n = 0 \tag{2.10.1}$$

But

$$\begin{cases} du_1 = \frac{\partial u_1}{\partial x_1} dx_1 + \frac{\partial u_1}{\partial x_2} dx_2 + \ldots + \frac{\partial u_1}{\partial x_n} dx_n \\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots \\ du_n = \frac{\partial u_n}{\partial x_1} dx_1 + \frac{\partial u_n}{\partial x_2} dx_2 + \ldots + \frac{\partial u_n}{\partial x_n} dx_n \end{cases} \tag{2.10.2}$$

and on substituting these values in (2.10.1) we get an equation of the form

$$A_1\, dx_1 + A_2\, dx_2 + \ldots + A_n\, dx_n = 0 \tag{2.10.3}$$

and since $dx_1, dx_2,\ldots, dx_n$ are the arbitrary differentials of the independent variables, it follows that

$$A_1 = 0, A_2 = 0,\ldots, A_n = 0$$

In other words,

(2.10.4)
$$\begin{cases} \dfrac{\partial \phi}{\partial u_1}.\dfrac{\partial u_1}{\partial x_1}+\dfrac{\partial \phi}{\partial u_2}.\dfrac{\partial u_2}{\partial x_1}+...+\dfrac{\partial \phi}{\partial u_n}\dfrac{\partial u_n}{\partial x_1}=0 \\ \dfrac{\partial \phi}{\partial u_1}.\dfrac{\partial u_1}{\partial x_2}+\dfrac{\partial \phi}{\partial u_2}.\dfrac{\partial u_2}{\partial x_2}+...+\dfrac{\partial \phi}{\partial u_n}\dfrac{\partial u_n}{\partial x_n}=0 \\ .... \\ \dfrac{\partial \phi}{\partial u_1}.\dfrac{\partial u_1}{\partial x_n}+\dfrac{\partial \phi}{\partial u_2}.\dfrac{\partial u_2}{\partial x_n}+...+\dfrac{\partial \phi}{\partial u_n}\dfrac{\partial u_n}{\partial x_n}=0 \end{cases}$$

and since, by hypothesis, we cannot have

$$\frac{\partial \phi}{\partial u_1}=\frac{\partial \phi}{\partial u_2}=...=\frac{\partial \phi}{\partial u_n}=0$$

on eliminating the partial derivatives of ϕ from the set of equation (2.10.4) we get

$$\frac{\partial(u_1,u_2,...,u_n)}{\partial(x_1,x_2,...,x_n)}=0$$

which establishes the theorem.

Theorem. 15. If $u_1, u_2,..., u_n$ are n functions of the n variables $x_1, x_2,..., x_n$ say $u_m = f_m(x_1, x_2,..., x_m)$, (m = 1, 2,...n), and if $\dfrac{\partial(u_1,u_2,...,u_n)}{\partial(x_1,x_2,...,x_n)}=0$, then if all the differential coefficients concerned are continuous, there exists a functional relation connecting some or all of the variables $u_1, u_2,..., u_n$ which is independent of $x_1, x_2,..., x_n$

Proof. First we prove the theorem when n = 2. We have $u = f(x, y)$, $v = g(x, y)$ and

$$\begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = 0$$

If v does not depend on y, then $\dfrac{\partial v}{\partial y}$ = 0 and so either $\dfrac{\partial u}{\partial y}$ = 0 or else $\dfrac{\partial v}{\partial x}=0$. In the former case u and v are functions of x only, and the functional relation sought is obtained from

$$u = f(x),\ v = g(x).$$

by regarding x as a function of v and substituting in $u = f(x)$. In the latter case v is a constant, and the functional relation is

$$v = a$$

If v does depend on y, since $\dfrac{\partial v}{\partial y} \neq 0$ the equation $v = g(x, y)$ defines y as a function of x and v, say

$$y = \psi(x, v),$$

and on substituting in the other equation we get an equation of the form

$$u = F(x, v).$$

(The *f*n. F [x, g (x, y)] is the same function of x and y as $f(x, y)$)

Then

$$0 = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial x} \end{vmatrix} = \begin{vmatrix} \frac{\partial F}{\partial x} + \frac{\partial F}{\partial v}.\frac{\partial v}{\partial x} & \frac{\partial F}{\partial v}\frac{\partial v}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{\partial F}{\partial x} & 0 \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

(obtained on multiplying the second now by $\frac{\partial F}{\partial u}$ and subtracting from the first) and so, either $\frac{\partial v}{\partial y}$ = 0, which is contrary to hypothesis or else $\frac{\partial F}{\partial x}$ =0, so that F is a function of v only ; hence the functional relation is

$$u = F(v)$$

Now assume that the theorem holds for n−1.

Now u_n must involve one of the variables at least, for if not there is a functional relation $u_n = a$. Let one such variable be called x_n Since $\frac{\partial u_n}{\partial x_n} \neq 0$ we can solve the equation

$$u_n = f_n(x_1, x_2, \ldots, x_n)$$

for x_n in terms of $x_1, x_2, \ldots, x_{n-1}$ and u_n, and on substituting this value in each of the other equations we get n−1 equations of the form

(2.10.5) $$u_r = g_r(x_1, x_2, \ldots, x_{n-1}, u_n),\ (r = 1, 2, \ldots, n-1) \qquad (2)$$

If now we substitute $f_n(x_1, x_2, \ldots, x_n)$ for u_n the functions $g_r(x_1, x_2, \ldots, x_{n-1}, u_n)$ become

$$f_r(x_1, x_2, \ldots, x_{n-1}, x_n),\ (r = 1, 2, \ldots, n-1)$$

Then

$$0 = \begin{vmatrix} \frac{\partial f_1}{\partial x_1}\frac{\partial f_1}{\partial x_2}, \ldots, \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1}\frac{\partial f_2}{\partial x_2}, \ldots, \frac{\partial f_2}{\partial x_n} \\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots \\ \frac{\partial f_n}{\partial x_1}\frac{\partial f_n}{\partial x_2}, \ldots, \frac{\partial f_n}{\partial x_n} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial g_1}{\partial x_1} + \frac{\partial g_1}{\partial u_n}.\frac{\partial u_n}{\partial x_1}, \ldots, \frac{\partial g_1}{\partial x_{n-1}} + \frac{\partial g_1}{\partial u_n}.\frac{\partial u_n}{\partial x_{n-1}}, & \frac{\partial g_1}{\partial u_n}.\frac{\partial u_n}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} + \frac{\partial g_2}{\partial u_n}.\frac{\partial u_n}{\partial x_1}, \ldots, \frac{\partial g_2}{\partial x_{n-1}} + \frac{\partial g_2}{\partial u_n}.\frac{\partial u_n}{\partial x_{n-1}}, & \frac{\partial g_2}{\partial u_n}.\frac{\partial u_n}{\partial x_n} \\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ldots\ldots & \ldots \\ \frac{\partial u_n}{\partial x_1}, \quad \ldots\ldots\ldots, \quad \frac{\partial u_n}{\partial x_{n-1}} \quad , & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial g_1}{\partial x_1},\dots,\frac{\partial g_1}{\partial x_{n-1}}, & 0 \\ \frac{\partial g_2}{\partial x_1},\dots,\frac{\partial g_2}{\partial x_{n-1}}, & 0 \\ \dots\dots\dots\dots\dots\dots & \dots\dots \\ \frac{\partial u_n}{\partial x_1},\dots,\frac{\partial u_n}{\partial x_{n-1}}, & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

by subtracting the elements of the last row multiplied by

$$\frac{\partial g_1}{\partial u_n},\frac{\partial g_2}{\partial u_n},\dots,\frac{\partial g_{n-1}}{\partial u_n}$$

from each of the others. Hence

$$\frac{\partial u_n}{\partial x_n}.\frac{\partial(g_1,g_2,\dots,g_{n-1})}{\partial(x_1,x_2,\dots,x_{n-1})}=0.$$

Since $\frac{\partial u_n}{\partial x_n} \neq 0$ we must have $\frac{\partial(g_1,g_2,\dots,g_{n-1})}{\partial(x_1,x_2,\dots,x_{n-1})}=0$, and so by hypothesis there is a functional relation between $g_1, g_2,\dots, g_{n-1}$, that is between $u_1, u_2,\dots, u_{n-1}$ into which u_n may enter, because u_n may occur in set of equation (2.10.5) as an auxiliary variable. We have therefore proved by induction that there is a relation between $u_1, u_2,\dots, u_n$.

Properties of Jacobian

Lemma. If U, V are functions of u and v, where u and v are themselves functions of x and y, we shall have

$$\frac{\partial(U,V)}{\partial(x,y)}=\frac{\partial(U,V)}{\partial(u,v)}.\frac{\partial(u,v)}{\partial(x,y)}$$

Proof. Let

$$U = f(u, v),\ V = F(u, v)$$

$$u = \phi(x, y),\ v = \psi(x, y)$$

Then

$$\frac{\partial U}{\partial x}=\frac{\partial U}{\partial u}.\frac{\partial u}{\partial x}+\frac{\partial U}{\partial v}\frac{\partial v}{\partial x}$$

$$\frac{\partial U}{\partial y}=\frac{\partial U}{\partial v}.\frac{\partial u}{\partial y}+\frac{\partial U}{\partial v}\frac{\partial v}{\partial y}$$

$$\frac{\partial V}{\partial x}=\frac{\partial V}{\partial u}.\frac{\partial u}{\partial x}+\frac{\partial V}{\partial v}\frac{\partial v}{\partial x}$$

$$\frac{\partial V}{\partial y}=\frac{\partial V}{\partial u}.\frac{\partial u}{\partial y}+\frac{\partial V}{\partial v}\frac{\partial v}{\partial y}$$

and

$$\frac{\partial(U,V)}{\partial(u,v)}.\frac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix} \frac{\partial U}{\partial u} & \frac{\partial U}{\partial v} \\ \frac{\partial V}{\partial u} & \frac{\partial V}{\partial v} \end{vmatrix} \times \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial U}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial U}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial U}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial U}{\partial v}\frac{\partial v}{\partial y} \\ \frac{\partial V}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial V}{\partial v}\frac{\partial u}{\partial x} & \frac{\partial V}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial V}{\partial v}\frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \end{vmatrix} = \frac{\partial(U, V)}{\partial(x, y)}$$

The same method of proof applies if there are several functions and the same number of variables.

Lemma. If J is he Jacobian of system u, v with regard to x, y and J′ the Jacobian of x, y with regard to u, v, then J J′ = 1.

Proof. Let $u = f(x, y)$ and $v = F(x, y)$, and suppose that these are solved for x and y giving

$$x = \phi(u, v) \text{ and } y = \psi(u, v),$$

we then have an differentiating $u = f(x, y)$ w.r.t u and v; = F(x, y) w.r.t u and v

$$\left.\begin{aligned} 1 &= \frac{\partial u}{\partial v}\frac{\partial x}{\partial u}+\frac{\partial u}{\partial y}.\frac{\partial y}{\partial u} \\ 0 &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial u}{\partial y}.\frac{\partial y}{\partial v} \end{aligned}\right\} \text{obtained from } u = f(x, y)$$

$$\left.\begin{aligned} 0 &= \frac{\partial v}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}.\frac{\partial y}{\partial u} \\ 1 &= \frac{\partial v}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial v}{\partial y}.\frac{\partial y}{\partial u} \end{aligned}\right\} \text{obtained from } v = F(x, y)$$

Also
$$J\,J' = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \times \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial u}{\partial y}.\frac{\partial y}{\partial u}, & \frac{\partial u}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial u}{\partial y}.\frac{\partial y}{\partial v} \\ \frac{\partial v}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial v}{\partial y}.\frac{\partial y}{\partial u}, & \frac{\partial v}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial v}{\partial y}.\frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Example15. If $u = x + 2y + z$, $v = x - 2y + 3z$
$w = 2xy - xz + 4yz - 2z^2$,

prove that $\dfrac{\partial(u, v, w)}{\partial(x, y, z)} = 0$, and find a relation between u, v, w.

Solution. We have

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 1 \\ 1 & -2 & 3 \\ 2y-z & 2x+4z & -x+4y-4z \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 1 & -4 & 2 \\ 2y-z & 2x+6z-4y & -x+2y-3z \end{vmatrix}$$ Performing c_2-2c_1 and c_3-c_1

$$= \begin{vmatrix} -4 & 2 \\ 2x+6y-4y & -x+2y-3z \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ 0 & -x+2y-3z \end{vmatrix}$$ Performing c_1+2c_2

$= 0$

Hence a relation between u, v and w exists

Now,

$u = v = 2x + 4z$

$u - v = 4y - 2z$

$w = x(2y-z) + 2z(2y-3)$

$= (x+2z)(2y-z)$

$\Rightarrow \quad 4w = (u+v)(u-v)$

$\Rightarrow \quad 4w = u^2 - v^2$

which is the required relation.

Example. 16. Find the condition that the expressions $px + qy + rz$, $p'x + q'y + r'z$ are connected with the expression $ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$. By a functional relation.

Solution. Let

$u = px + qy + rz$

$v = p' + q'y + r'z$

$w = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$

We know that the required condition is

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0$$

Therefore

$$\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} = 0$$

But

$$\frac{\partial u}{\partial x} = p, \quad \frac{\partial u}{\partial y} = q, \quad \frac{\partial u}{\partial z} = r$$

$$\frac{\partial v}{\partial x} = p' \quad \frac{\partial y}{\partial y} = q', \quad \frac{\partial v}{\partial z} = r'$$

$$\frac{\partial w}{\partial x} = 2ax + 2hy + 2gz$$

$$\frac{\partial w}{\partial y} = 2hx + 2by + 2fz$$

$$\frac{\partial w}{\partial z} = 2gx + 2fy + 2cz$$

Therefore

$$\begin{vmatrix} p & q & r \\ p' & q' & r' \\ 2ax + 2hy + 2gz & 2hx + 2by + 2fz & 2gx + 2fy + 2cz \end{vmatrix} = 0$$

$$\Rightarrow \quad \begin{vmatrix} p & q & r \\ p' & q' & r' \\ a & h & g \end{vmatrix} = 0, \begin{vmatrix} p & q & r \\ p' & q' & r' \\ h & b & f \end{vmatrix} = 0, \begin{vmatrix} p & q & r \\ p' & q' & r' \\ g & f & c \end{vmatrix} = 0$$

which is the required condition.

Example. 17. Prove that if $f(0) = 0$, $f(x) = \frac{1}{1+x^2}$, then

$$f(x) + f(y) = f\left(\frac{x+y}{1-xy}\right)$$

Solution. Suppose that

$$u = f(x) + f(y)$$

$$v = \frac{x+y}{1-xy}$$

Now $$J(u, v) = \begin{vmatrix} \frac{\partial u}{\partial x}, & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{1}{1+x^2} & \frac{1}{1+y^2} \\ \frac{1+y^2}{(1-xy)^2} & \frac{1+x^2}{(1+xy)^2} \end{vmatrix} = 0$$

Therefore u and v are connected by a functional relation

Let $u = \phi(v)$, that is,

$$f(x) + f(y) = \phi\left(\frac{x+y}{1-xy}\right)$$

Putting $y = 0$, we get

$$f(x) + f(0) = \phi(x)$$

$$\Rightarrow \quad f(x) + 0 = \phi(x) \qquad \because f(0) = 0$$

Hence $f(x) + f(y) = f\left(\dfrac{x+y}{1-xy}\right)$

Example. 18. The roots of the equation in λ

$$(\lambda - x)^3 + (\lambda - y)^3 + (\lambda - z)^3 = 0$$

are u, v, w Prove that $\dfrac{\partial(u,v,w)}{\partial(x,y,z)} = -2\dfrac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$

Solution. Here u, v, w are the roots of the equation $\lambda^3 - (x+y+z)\lambda^2 + (x^2+y^2+z^2)\lambda - \dfrac{1}{3}(x^3+y^3+z^3) = 0$

Let $x + y + z = \xi,\ x^2 + y^2 + z^2 = \eta,\ \dfrac{1}{3}(x^3 + y^3 + z^3) = \zeta$ (1)

and then $u + v + w = \xi,\ vw + wu + uv = \eta,\ uvw = \zeta$ (2)

Then from (1)

$$\frac{\partial(\zeta,\eta,z)}{\partial(x,y,\zeta)} = \begin{vmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \\ x^2 & y^2 & z^2 \end{vmatrix} = 2\,(y-z)\,(z-x)\,(x-y) \tag{3}$$

Again, from (3), we have

$$\frac{\partial(\xi,\eta,\zeta)}{\partial(v,u,w)} = \begin{vmatrix} 1 & 1 & 1 \\ v+w & w+u & u+v \\ vw & wu & uv \end{vmatrix} = -(v-w)\,(w-h)\,(u-v) \tag{4}$$

Then from (3) and (4)

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \frac{\partial(u,v,w)}{\partial(\xi,\eta,\zeta)} \cdot \frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,\zeta)} = -2\frac{(y-z)(z-x)(x-y)}{(u-w)(w-u)(u-v)}$$

Example. 19. If α, β, γ are the roots of the equation in k :

$$\frac{x}{a+k} + \frac{y}{b+k} + \frac{z}{c+k} = 1,$$

then

$$\frac{\partial(x,y,z)}{\partial(\alpha,\beta,\gamma)} = -\frac{(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)}{(a-b)(b-c)(c-a)}$$

Solution. The equation in k is

$$k^3 + k^2(a+b+c-x-y-z) + k[ab + bc + ca - x(b+c) - y(c+a) - z(a+b)] + abc - bcx - cay - abz = 0 \tag{*}$$

Now α, β, γ are the roots of this equation. Therefore

$$\alpha + \beta + \gamma = -(a + b + c) + x + y + z$$

$$\alpha\beta + \beta\gamma + \gamma\alpha = ab + bc + ca - x(b+c) - y(c+a) - z(a+b)$$

and

$$\alpha\beta\gamma = -abc + bcx + cay + abz.$$

Then, we have

$$1 = \frac{\partial x}{\partial \alpha} + \frac{\partial y}{\partial \alpha} + \frac{\partial z}{\partial \alpha}$$

$$1 = \frac{\partial x}{\partial \beta} + \frac{\partial y}{\partial \beta} + \frac{\partial z}{\partial \beta}$$

$$1 = \frac{\partial x}{\partial \gamma} + \frac{\partial y}{\partial \gamma} + \frac{\partial z}{\partial \gamma}$$

$$\beta + \gamma = -(b+c)\frac{\partial x}{\partial \alpha} - (c+a)\frac{\partial y}{\partial \alpha} - (a+b)\frac{\partial z}{\partial \alpha}$$

$$\gamma+\alpha = -(b+c)\ \frac{\partial x}{\partial \beta}-(c+a)\frac{\partial y}{\partial \beta}-(a+b)\frac{\partial z}{\partial \beta}$$

$$\alpha+\beta = -(b+c)\ \frac{\partial x}{\partial \gamma}-(c+a)\frac{\partial y}{\partial \gamma}-(a+b)\frac{\partial z}{\partial \gamma}$$

$$\beta\gamma = bc\ \frac{\partial x}{\partial \alpha}+ca\frac{\partial y}{\partial \alpha}+ab\frac{\partial z}{\partial \alpha}$$

$$\gamma\alpha = bc\ \frac{\partial x}{\partial \beta}+ca\frac{\partial y}{\partial \beta}+ab\frac{\partial z}{\partial \beta}$$

$$\alpha\beta = bc\ \frac{\partial x}{\partial \gamma}+ca\frac{\partial y}{\partial \gamma}+ab\frac{\partial z}{\partial \gamma}$$

Now,
$$\begin{vmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial y}{\partial \alpha} & \frac{\partial z}{\partial \alpha} \\ \frac{\partial x}{\partial \beta} & \frac{\partial y}{\partial \beta} & \frac{\partial z}{\partial \beta} \\ \frac{\partial x}{\partial \gamma} & \frac{\partial y}{\partial \gamma} & \frac{\partial z}{\partial \gamma} \end{vmatrix} \begin{vmatrix} 1 & 1 & 1 \\ -(b+c) & -(c+a) & -(a+b) \\ bc & ca & ab \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & 1 \\ \beta+\gamma & \gamma+\alpha & \alpha+\beta \\ \beta\gamma & \gamma\alpha & \alpha\beta \end{vmatrix}$$

Hence

$$\frac{\partial(x,y,z)}{\partial(\alpha,\beta,\gamma)}\ (b-c)\ c-a)\ (a-b) = -(\alpha-\beta)\ (\beta-\gamma)\ (\gamma-\alpha)$$

$$\Rightarrow \qquad \frac{\partial(x,y,z)}{\partial(\alpha,\beta,\gamma)} = -\frac{(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)}{(b-c)(c-a)(a-b)}$$

Second Method. After the step (*) let $a+b+c-(x+y+z) = \xi$, $ab + bc + ca - x(b+a) - y(c+a) - z(a+b) = \eta$

$$abc - bcx - cay - abz = \zeta \qquad (1)$$

$$\alpha + \beta + \gamma = -\xi,\ \alpha\beta + \beta\gamma + \gamma\alpha = \eta,\ \alpha\beta\gamma = -\zeta \qquad (2)$$

then
$$\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)} = \begin{vmatrix} -1 & -1 & -1 \\ -(b+c) & -(c+a) & -(a+b) \\ -bc & -ca & -ab \end{vmatrix} \text{ and } \frac{\partial(\xi,\eta,\zeta)}{\partial(\alpha,\beta,\gamma)} = \begin{vmatrix} -1 & -1 & -1 \\ \beta+\gamma & \gamma+\alpha & \alpha+\beta \\ -\beta\gamma & -\gamma\alpha & -\alpha\beta \end{vmatrix}$$

$$= (a-b)\ (b-c)\ (c-a) = -\ (\alpha-\beta)\ (\beta-\gamma)\ (\gamma-\alpha)$$

Therefore
$$\frac{\partial(x,y,z)}{\partial(\alpha,\beta,\zeta)} = \frac{\partial(x,y,z)}{\partial(\xi,\eta,\zeta)}\cdot\frac{\partial(\xi,\eta,z\zeta}{\partial(\alpha,\beta,\gamma)} = -\frac{(\alpha-\beta)(\beta-\alpha)(\gamma-\alpha)}{(a-b)(b-c)(c-a)}$$

Example. 20. Prove that the three functions U, V, W are connected by an identical functional relation if

$$U = x + u - z,\ V = x-y+z,\ W = x^2 + y^2 + z^2 - 2yz$$

and find the functional relation.

Solution. Here

$$\frac{\partial(U,V,W)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z} \\ \frac{\partial W}{\partial x} & \frac{\partial W}{\partial y} & \frac{\partial W}{\partial z} \end{vmatrix} = \begin{vmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2x & 2(y-z) & 2(z-y) \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 2x & 2(y-z) & 0 \end{vmatrix} = 0$$

Hence there exists some functional relation between U, V and W.

Moreover,

$$U + V = 2x$$

$$U - V = 2(y-z)$$

and $(U+V)^2 + (U-V)^2 = 4(x^2 + y^2 + z^2 - 2yz)$

$$= 4W$$

which is the required functional relation.

Example. 21. Let V be a function of the two variables, x and y. Transform the expression

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}$$

by the formulae of plane polar transformation.

$$x = r\cos\theta,\ y = r\sin\theta.$$

Solution. We are given a function V which is function of x and y and therefore it is a function of r and θ. From $x = r\cos\theta$ & $y = r\sin\theta$, we have

$$r = \sqrt{x^2 + y^2}\ ,\ \ \theta = \tan^{-1} y/x$$

Now

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial r}.\frac{\partial r}{\partial x} + \frac{\partial V}{\partial u}.\frac{\partial \theta}{\partial x}$$

$$= \cos\theta\frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta} \qquad \left(\because \quad \frac{\partial r}{\partial x} = \cos\theta, \frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r}\right)$$

and

$$\frac{\partial V}{\partial y} = \frac{\partial V}{\partial r}.\frac{\partial r}{\partial y} + \frac{\partial V}{\partial \theta}.\frac{\partial \theta}{\partial y}$$

$$= \sin\theta\frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta} \qquad \left(\because \quad \frac{\partial r}{\partial y} = \sin\theta, \frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r}\right)$$

Therefore

$$\frac{\partial}{\partial x} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)$$

$$\frac{\partial}{\partial y} = \left(\sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)$$

Hence
$$\frac{\partial^2 V}{\partial x^2} = \left(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)$$

$$= \cos\theta \frac{\partial}{\partial r}\left(\cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)$$

$$= \cos\theta\left(\cos\theta \frac{\partial^2 V}{\partial r^2} + \frac{\sin\theta}{r^2}\frac{\partial V}{\partial \theta} - \frac{\sin\theta}{r}\frac{\partial^2 V}{\partial x \partial \theta}\right)$$

$$-\frac{\sin\theta}{r}\left(\cos\theta \frac{\partial^2 V}{\partial x \partial r} - \sin\theta \frac{\partial V}{\partial r} - \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta} - \frac{\sin\theta}{r}\frac{\partial^2 V}{\partial \theta^2}\right)$$

$$= \cos^2\theta \frac{\partial^2 V}{\partial r^2} - 2\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 V}{\partial r \partial \theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 V}{\partial \theta^2}$$

$$+ \frac{\sin^2\theta}{r}\frac{\partial V}{\partial r} + \frac{2\sin\theta\cos\theta}{r^2}\frac{\partial V}{\partial \theta} \tag{1}$$

and
$$\frac{\partial^2 V}{\partial y^2} = \left(\sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\sin\theta \frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}\right)$$

$$= \sin\theta \frac{\partial}{\partial r}\left(\sin\theta \frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}\right) + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}\right)$$

$$= \sin\theta\left(\sin\theta \frac{\partial^2 V}{\partial r^2} - \frac{\cos\theta}{r^2}\frac{\partial V}{\partial \theta} + \frac{\cos\theta}{r}\frac{\partial^2 V}{\partial r \partial \theta}\right)$$

$$+ \frac{\cos\theta}{r}\left(\sin\theta \frac{\partial^2 V}{\partial \theta \partial r} + \cos\theta \frac{\partial V}{\partial \theta} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta} + \frac{\cos\theta}{r}\frac{\partial^2 V}{\partial \theta^2}\right)$$

$$= \sin^2\theta \frac{\partial^2 V}{\partial r^2} + \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 V}{\partial r \partial \theta} - \frac{\cos\theta\sin\theta}{r^2}\frac{\partial V}{\partial \theta}$$

$$+ \frac{\cos\theta\sin\theta}{r}\frac{\partial^2 V}{\partial \theta \partial r} + \frac{\cos^2\theta}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{\cos^2\theta}{r}\frac{\partial V}{\partial r}$$

$$-\frac{\sin\theta\cos\theta}{r^2}\frac{\partial V}{\partial \theta} \tag{2}$$

Adding (1) and (2) we obtain

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{\partial^2 V}{\partial r^2} + \frac{1}{r}.\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2}$$

which is the required result.

Example. 22. Transform the expression

$$\left(x\frac{\partial Z}{\partial x}+y\frac{\partial Z}{\partial y}\right)^2 + (a^2-x^2-y^2)\left\{\left(\frac{\partial Z}{\partial x}\right)^2+\left(\frac{\partial Z}{\partial y}\right)^2\right\}$$

by the substitution $x = r\cos\theta$, $y = r\sin\theta$

Solution. If V is a function of x, y, then

$$\frac{\partial V}{\partial r}=\frac{\partial V}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial V}{\partial y}.\frac{\partial y}{\partial r}=\frac{x}{r}\frac{\partial V}{\partial x}+\frac{y}{r}\frac{\partial V}{\partial y}$$

$$\Rightarrow \qquad r\frac{\partial V}{\partial r}=x\frac{\partial V}{\partial x}+y\frac{\partial V}{\partial y}=\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial x}\right)v$$

$$\Rightarrow \qquad r\frac{\partial}{\partial r}=x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}$$

Similarly $$\frac{\partial}{\partial\theta}=x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}$$

Now $$\frac{\partial Z}{\partial x}=\frac{\partial Z}{\partial r}.\frac{\partial r}{\partial x}+\frac{\partial z}{\partial\theta}.\frac{\partial\theta}{\partial x}=\cos\theta\frac{\partial z}{\partial r}-\frac{\sin\theta}{r}\frac{\partial z}{\partial\theta} \qquad (1)$$

$$\frac{\partial Z}{\partial y}=\sin\theta\frac{\partial Z}{\partial r}+\frac{\cos\theta}{r}.\frac{\partial Z}{\partial\theta} \qquad (2)$$

Therefore $$\left(\frac{\partial Z}{\partial x}\right)^2+\left(\frac{\partial Z}{\partial y}\right)^2=\left(\frac{\partial Z}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial Z}{\partial\theta}\right)^2$$

and the given expression is equal to

$$\left(r\frac{\partial Z}{\partial r}\right)^2+(a^2-r^2)\left[\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial Z}{\partial\theta}\right)^2\right]$$

$$= a^2\left(\frac{\partial Z}{\partial r}\right)^2+\left(\frac{a^2}{r^2}-1\right)\left(\frac{\partial Z}{\partial\theta}\right)^2.$$

Example. 23. If $x = r\cos\theta$, $y = r\sin\theta$, prove that

$$(x^2-y^2)\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}\right)+4xy\frac{\partial^2 u}{\partial x\partial y}=r^2\frac{\partial^2 u}{\partial r^2}-r\frac{\partial u}{\partial r}-\frac{\partial^2 u}{\partial\theta^2}$$

where u is any twice differentiable function of x and y.

Solution. We have

$$\frac{\partial u}{\partial r}.\frac{\partial x}{\partial x}+\frac{\partial u}{\partial y}.\frac{\partial y}{\partial r}$$

$$= \cos\theta \frac{\partial u}{\partial x} + \sin\theta \frac{\partial u}{\partial y} = \frac{x}{r}\frac{\partial u}{\partial x} + \frac{y}{r}\frac{\partial u}{\partial y}$$

$\Rightarrow$ $$r\frac{\partial u}{\partial r} = x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} \qquad (1)$$

Therefore $$r\frac{\partial}{x}\left(i\frac{\partial u}{\partial r}\right) = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right)$$

$$= x\frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right) + y\frac{\partial}{\partial y}\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right)$$

$$= x^2\frac{\partial^2 u}{\partial x^2} + xy\frac{\partial^2 u}{\partial x\partial y} + xy\frac{\partial^2 u}{\partial y\partial x} + y^2\frac{\partial^2 u}{\partial y^2} + x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}$$

Therefore

$\therefore$ $$r^2\frac{\partial^2 r}{\partial r^2} + r\frac{\partial u}{\partial r} = x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} + x\frac{\partial y}{\partial x} + y\frac{\partial u}{\partial y} \qquad (2)$$

$$r^2\frac{\partial^2 u}{\partial r^2} = x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial xvy} + y^2\frac{\partial^2 u}{\partial y^2} \qquad \text{using (1)}$$

Again, $$\frac{\partial u}{\partial \theta} = \frac{\partial u}{\partial x}.\frac{\partial x}{\partial \theta} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial \theta}$$

$$= x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}$$

Therefore $$\frac{\partial^2 u}{\partial \theta^2} = \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\left(x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}\right)$$

$$= x\frac{\partial}{\partial y}\left(x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}\right) - y\frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}\right)$$

$$= x^2\frac{\partial^2 u}{\partial y^2} - 2xy\frac{\partial^2 u}{\partial u\partial x} + y^2\frac{\partial^2 u}{\partial x^2} - x\frac{\partial u}{\partial x} - y\frac{\partial u}{\partial y} \qquad (3)$$

From (1), (2) and (3) we get the required result.

Example 24. If $x = r\cos\theta$, $y = r\sin\theta$, show that

$$\frac{\partial^2 \theta}{\partial x\partial y} = -r^{-2}\cos 2\theta$$

Solution. We have

$$x = r\cos\theta,\ y = r\sin\theta$$

Then $$dx = dr\cos\theta - r\sin\theta\, d\theta$$

$$dy = dr\sin\theta + r\cos\theta\, d\theta$$

$$\therefore \qquad r d\theta = \cos\theta\, dy - \sin\theta\, dx$$

$$dr = \cos\theta\, du + \sin\theta\, dy$$

$$\therefore \qquad dr d\theta + r\, d^2\theta = -\sin\theta\, d\theta\, dy - \cos\theta\, d\theta\, dx$$

$$= -(\sin\theta\, dy + \cos\theta\, dx)\, d\theta$$

$$= -(\sin\theta\, dy + \cos\theta\, dx)\, \frac{1}{r}(\cos\theta\, dy - \sin\theta\, dx)$$

$$\Rightarrow \qquad r\, d^2\theta = -\frac{1}{r}(\sin\theta\, dy + \cos\theta\, dx)(\cos\theta\, dy - \sin\theta\, dx) - \frac{1}{r}(\cos\theta\, dx + \sin\theta\, dy)$$

$$(\cos\theta\, dy - \sin\theta\, dx)$$

$$\Rightarrow \qquad d^2\theta = -\frac{1}{r^2}[(\sin\theta dy + \cos\theta\, dx\, (\cos\theta\, dy - \sin\theta\, dx) + (\cos dx + \sin\theta\, dy)$$

$$(\cos\theta\, dy - \sin\theta\, dx)]$$

$$= -\frac{1}{r^2}[\sin\theta\cos\theta\, dy^2 - \sin^2\theta\, dx\, dy + \cos^2\theta\, dx\, dy - \sin\theta\cos\theta\, dx^2$$

$$+ \cos^2\theta\, dx\, dy - \sin\theta\cos dx + \sin\theta\cos\theta\, dy^2 - \sin^2\theta\, dxdy]$$

$$= -\frac{1}{r^2}[2\cos^2\theta\, dx\, dy - 2\sin^2\theta\, dx\, dy + 2\sin\theta\cos\theta\, dy^2 - 2\sin\theta\cos\theta\, dx]$$

$$\Rightarrow \qquad \frac{\partial^2\theta}{\partial x^2}dx^2 + 2\frac{\partial^2\theta}{\partial x\partial y}dxdy + \frac{\partial^2\theta}{\partial y^2}dy^2$$

$$= -\frac{1}{r^2}[2\cos^2\theta\, dx\, dy - 2\sin^2\theta\, dx\, dy + 2\sin\theta\cos\theta\, dy^2 - 2\sin\theta\cos\theta\, dx]$$

Now
$$\frac{\partial^2\theta}{\partial x^2} = \frac{2\cos\theta\sin\theta}{r^2}, \quad \frac{\partial^2\theta}{\partial y^2} = \frac{-2\sin\theta\cos\theta}{r^2}$$

Therefore

$$\frac{\partial^2\theta}{\partial x\partial y} = -\frac{1}{2r^2}[2\cos^2\theta - 2\sin^2\theta]$$

$$= -\frac{1}{r^2}\cos 2\theta = -\frac{\cos 2\theta}{r^2}.$$

3

PART A : THE RIEMANN - STIELTJES INTEGRAL

3.1 We have been dealing with Riemann integrals in our undergraduate level studies in mathematics. The aim of this chapter is to consider a more general concept than that of Riemann. This concept is known as Riemann – Stieltjes integral which involve two function f and α. In what follows, we shall consider only real – valued functions.

3.2 Definitions and Notations

Definition. Let [a, b] be a given interval. By a partition (or subdivision) P of [a, b] , we mean a finite set of points

$$P = \{x_0, x_1, \ldots\ldots, x_n\}$$

such that

$$a = x_0 \le x_1 \le x_2 \le \ldots\ldots\; x_{n-1} \le x_n = b.$$

Definition. A partition P* of [a, b] is said to be finer than P (or a refinement of P) if $P^* \supseteq P$, that is, if every point of P is a point of P*.

Definition. The P_1 and P_2 be two partitions of an interval [a, b]. Then a partition P* is called their common refinement if $P^* = P_1 \cup P_2$.

Definition. The length of the largest subinterval of a partition $P = \{x_0, x_1, \ldots, x_n\}$ of [a, b] is called the **Norm** (or **Mess**) of P. We denote norm of P by | P | . Thus

$$|P| = \max \Delta x_i = \max\, [x_i - x_{i-1} : i = 1,2,\ldots,n]$$

We notice that if $P^* \supseteq P$, then $|P^*| \le |P|$. Thus refinement of a partition decreases its norm.

Let f be a bounded real function defined on [a, b]. Corresponding to each partition P of [a,b], we put

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

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

Let α be a monotonically increasing function on [a, b]. Then α is bounded on [a, b] since $\alpha(a)$ and $\alpha(b)$ are finite.

Corresponding to each partition P of [a, b], we put

$$\Delta \alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$

The monotonicity of α implies that $\Delta \alpha_i \ge 0$.

For any real valued bounded function f on [a, b], we take

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

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

Where m_i and M_i are bounds of f defined above. The sums L(P, f, α) and U(P, f, α) are respectively called Lower Stieltjes sum and Upper Stieltjes sum corresponding to the partition P. We further define

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

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

where lub and glb are taken over all possible partitions P of [a, b]. Then $\underline{\int_a^b} f\, d\alpha$ and $\overline{\int_a^b} f\, d\alpha$ are respectively called Lower integral and Upper integrals of f with respect to α.

If the lower and upper integrals are equal, then their common value, denoted by $\int_a^b f\, d\alpha$, is called the **Riemann – Stieltjes integral** of f with respect to α, over [a, b] and in that case we say that f is integrable with respect to α, in the Riemann sense and we write $f \in \Re(\alpha)$.

The functions f and α are known as the **integrand** and the **integrator** respectively.

In the special case, when $\alpha(x) = x$, the Riemann - Stieltjes integral reduces to Riemann – integral. In such a case we write L(P, f), U(P, f), $\underline{\int_a^b} f$, $\overline{\int_a^b} f$ and $f \in \Re$ respectively in place of L(P, f, α), U(P, f, α), f dα, $\underline{\int_a^b}$ $\overline{\int_a^b} f\, d\alpha$ and $f \in \Re(\alpha)$.

Clearly, the numerical value of $\int f\, d\alpha$ depends only on f, α, a and b and does not depend on the symbol x. In fact x is a “dummy variable” and may be replaced by any other convenient symbol.

3.3. In this section, we shall study characterization of upper and lower Stieltjes sums, and upper and lower Stieltjes integrals.

The next theorem shows that for increasing function α, refinement of the partition increases the lower sums and decreases the upper sums.

Theorem 1. If P* is a refinement of a partition P of [a, b], then

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

Proof. Suppose first that P* contains exactly one more point than the partition P of [a, b]. Let this point be x* and let this point lie in the subinterval $[x_{i-1}, x_i]$. Let

$$W_1 = \text{glb } f(x) \qquad (x_{i-1} \le x \le x^*)$$
$$W_2 = \text{glb } f(x) \qquad (x^* \le x \le x_i)$$

Then $w_1 \ge m_i$ and $w_2 \ge m_i$ where

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

Hence

$$\begin{aligned} L(P^*, f, \alpha) - L(P,f,\alpha) &= w_1[\alpha(x^*) - \alpha(x_{i-1})] - w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_{i-1})] \\ &= (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \\ &\ge 0 \end{aligned}$$

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

If P* contains k points more than P, we repeat the above reasoning k times.

The proof for $U(P^*, f, \alpha) \le U(P, f, \alpha)$ is analogous.

Theorem 2. If α is monotonically increasing on [a, b], them for any two partitions $P_1$and P_2, we have

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

Proof. Let P be the common refinement of P_1 and P_2, that is, $P = P_1 \cup P_2$. Then we have, using Theorem1,

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

Remark. It also follows from this theorem that

$m[\alpha(b)-\alpha(a)] \le L(P_1, f, \alpha) \le U(P_2, f, \alpha) \le M[\alpha(b)-\alpha(a)]$,

where m and M are as usual inf and sup of f on [a, b].

Theorem 3. If α is increasing on [a, b], then

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

Proof. Let P* be the common refinement of two partitions P_1 and P_2. Then , by Theorem 1,

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

Hence

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

We keep P_2 fixed, and take lub over all P_1. We obtain

$$\underline{\int} f\,d\alpha \le U(P_2, f, \alpha)$$

Taking glb over all P_2, we get

$$\underline{\int} f\,d\alpha \le \overline{\int} f\,d\alpha.$$

Example. Let $\alpha(x) = x$ and define f on [0, 1] by

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

Then for every partition P of [0, 1], we have

$$m_i = 0,\ M_i = 1,$$

because every subinterval $[x_{i-1}, x_i]$ contain both rational and irrational number. Therefore

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

$$= 0$$

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

$$= \sum_{i=1}^{n} (x_i - x_{i-1}) = x_n - x_0 = 1 - 0$$

$$= 1$$

Hence, in this case

$$\underline{\int} f\,d\alpha \le \overline{\int} f\,d\alpha.$$

Theorem 4. Let $\alpha \nearrow$ on [a, b]. Then $f \in \Re(\alpha)$ if and only if for every $\in > 0$ there exists a partition P such that

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

Proof. Suppose first that for every P we have

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

This gives us

$$\left[U(P, f, \alpha) - \overline{\int_a^b} f\,d\alpha\right] + \left[\overline{\int_a^b} f\,d\alpha - \underline{\int_a^b} f\,d\alpha\right] + \left[\underline{\int_a^b} f\,d\alpha - L(P, f, \alpha)\right] < \in.$$

Since, each one of the three numbers

$U(P, f, \alpha) - \overline{\int} f\,d\alpha$, $\overline{\int} f\,d\alpha - \underline{\int} f\,d\alpha$, $\underline{\int} f\,d\alpha - L(P, f, \alpha)$ is non-negative, we have

$$0 \leq \overline{\int_a} f\, d\alpha - \underline{\int} f\, d\alpha < \in.$$

Since $\in$ is arbitrary positive number, we note that the non-negative number $\overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha$ is less than every positive number and hence

$$\overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha = O$$

which yields

$$\underline{\int} f\, d\alpha = \overline{\int} f\, d\alpha$$

and so $f \in \Re(\alpha)$.

Conversely, suppose that $f \in \Re(\alpha)$ and that $\in > 0$ be given. Then

$$\overline{\int} f\, d\alpha = \underline{\int} f\, d\alpha = \int f\, d\alpha$$

and there exist partitions $P_1$and P_2 such that

(3.3.1) $\qquad U(P_2, f, \alpha) < \overline{\int} f\, d\alpha + \frac{\in}{2} = \int f\, d\alpha + \frac{\in}{2}$

(3.3.2) $\qquad L(P_1, f, \alpha) > \underline{\int} f\, d\alpha - \frac{\in}{2} = \int f\, d\alpha - \frac{\in}{2}$

Let P the common refinement of P_1 and P_2.Then

$$U(P, f, \alpha) \leq U(P_2, f, \alpha)$$

and

$$L(P, f, \alpha) \leq L(P, f, \alpha)$$

Thus the relation (3.3.1) and (3.3.2) reduce to

(3.3.3) $\qquad U(P, f, \alpha) < \int f\, d\alpha + \frac{\in}{2}$

(3.3.4) $\qquad L(P, f, \alpha) > \int f\, d\alpha - \frac{\in}{2}$

Combining (3.3.3) and (3.3.4), we obtain

$$\int f\, d\alpha - \frac{\in}{2} < L(P, f, \alpha) < U(P, f, \alpha) < \int f\, d\alpha + \frac{\in}{2}$$

which yields

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

This completes the proof of the theorem.

3.4. In this section, we shall discuss integerability of continuos and monotonic functions alongwith properties of Riemann-Stieltjes integrals.

Theorem 5. If f is continuous on [a, b], then (α).

(i) $\quad f \in \Re(\alpha)$

(ii) $\quad$ to every $\in > 0$ there corresponds a $\delta > 0$ such that

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f\, d\alpha \right| < \in$$

for every partition P of [a, b] with $| P | < \delta$ and for all $t_i \in [x_{i-1}, x_i]$.

Proof. (i) Let $\in > 0$ and select $\eta > 0$ such that

(3.4.1) $[\alpha(b) - \alpha(a)]\, \eta > \in$

which is possible by monotonicity of α on [a, b]. Also f is continuous on compact set [a, b]. Hence f is uniformly continuous on [a, b]. Therefore there exists a $\delta > 0$ such that

(3.4.2) $| f(x) - f(t) | < \eta$ whenever $| x - t | < \delta$ for all $x \in [a, b]$, $t \in [a, b]$.

Choose a partition P with $| P | < \delta$. Then (3.4.2) implies

$$M_i - m_i \le \eta \qquad (i = 1, 2, \ldots, n)$$

Hence

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} M_i \Delta\alpha_i - \sum_{i=1}^{n} m_i \Delta\alpha_i \\ &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \le \eta \sum_{i=1}^{n} \Delta\alpha_i \\ &= \eta \sum_{i=1}^{n} [\alpha_i(x_i) - \alpha(x_{i-1})] \\ &= \eta[\alpha(b) - \alpha(a)] \\ &< \eta . \frac{\in}{\eta} = \in, \end{aligned}$$

which is necessary and sufficient condition for $f \in \Re(\alpha)$.

(ii) We have

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

and

$$L(P, f, \alpha) \le \int_a^b f d\alpha \le U(P, f, \alpha)$$

Since $f \in \Re(\alpha)$, for each $\in > 0$ there exists $\delta > 0$ such that for all partition P with $| P | < \delta$, we have

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

Thus

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f\, d\alpha \right| < U(P, f, \alpha) - L(P, f, \alpha)$$

$$< \in$$

Thus for continuous functions f, $\lim_{|P| \to 0} \sum_{i=1}^{n} f(t_i)\Delta\alpha_i$ exits and is equal to $\int_a^b f\, d\alpha$.

Theorem 6. If f is monotonic on [a, b] and if α is both monotonic and continuous on [a, b], then $f \in \Re(\alpha)$.

Proof. Let $\in$ be a given positive number. For any positive integer n, choose a partition P of [a, b] such that

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n} \quad (i = 1, 2, \ldots, n).$$

This is possible since α is continuous and monotonic on [a, b] and so assumes every value between its bounds $\alpha(a)$ and $\alpha(b)$. If is sufficient to prove the result for monotonically increasing function f, the proof for monotonically decreasing function being analogous. The bounds of f in $[x_{i-1}, x_i]$ are then

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

Hence

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

$$= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} (M_i - m_i)$$

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

$$= \frac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)]$$

$< \in$ for large n.

Hence $f \in \Re(\alpha)$.

Example. Let f be a function defined by

$f(x^*) = 1$ and $f(x) = 0$ for $x \neq x^*$, $a \leq x^* \leq b$.

Suppose α is increasing on [a, b] and is continuous at x^*. Then $f \in \Re(\alpha)$ over [a, b] and $\int_a^b f\, d\alpha$ $= 0$

Solution. Let $P = \{x_0, x_1, \ldots x_n\}$ be a partition of [a, b] and let $x^* \in \Delta x_i$. Since α is continuous at x^*, to each $\in > 0$ there exists $\delta > 0$ such that

$$|\alpha(x) - \alpha(x^*)| < \frac{\in}{2} \text{ whenever } |x - x^*| < \delta$$

Again since α is an increasing function,

$$\alpha(x) - \alpha(x^*) < \frac{\in}{2} \text{ for } 0 < x - x^* < \delta$$

and

$$\alpha(x^*) - \alpha(x) < \frac{\in}{2} \text{ for } 0 < x^* - x < \delta$$

Then for a partition P of [a, b],

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$$
$$= \alpha(x) - \alpha(x^*) + \alpha(x^*) - \alpha(x_{i-1})$$
$$< \frac{\in}{2} + \frac{\in}{2} = \in.$$

Therefore $\sum_{i=1}^{n} f(t_1)\Delta\alpha_i = \begin{cases} 0 \text{ if } t_i \neq x^* \\ \Delta\alpha_i, t_i = x^*. \end{cases}$

that is,

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - 0 \right| < \in$$

Hence

$$\lim_{|P|\to 0} \sum_{i=1}^{n} f(t_i)\Delta\alpha_i = \int_a^b f\, d\alpha = 0.$$

and so $f \in \Re(\alpha)$ and $\int_a^b f\, d\alpha = 0$.

Theorem 7. Let $f_1 \varepsilon \Re(\alpha)$ and $f_2 \varepsilon \Re(\alpha)$ on [a,b], then $(f_1 + f_2) \varepsilon \Re(\alpha)$ and

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

Proof. Let $P = [a = x_0, x_1, \ldots\ldots x_n = b]$ be any partition of [a, b]. Suppose further that M_i', m_i', M_i'', m_i'' and M_i, m_i are the bounds of f_1, f_2 and $f_1 + f_2$ respectively in the subinterval $[x_{i-1}, x_i]$. If $\alpha_1, \alpha_2 \in [x_{i-1}, x_i]$, then

$$\begin{aligned}
&[f_1(\alpha_2) + f_2(\alpha_2)] - [f_1(\alpha_1) + f_2(\alpha_1)] \\
&\quad \le | f_1(\alpha_2) - f_1(\alpha_1) | + | f_2(\alpha_2) - f_2(\alpha_1)] \\
&\quad \le (M_i' - m_i') + (M_i'' - m_i'')
\end{aligned}$$

Therefore, since this hold for all $\alpha_1, \alpha_2 \in [x_{i-1}, x_i]$, we have

(3.4.3) $$M_i - m_i \le (M_i' - m_i') + (M_i'' - m_i'')$$

Since $f_1, f_2 \in \Re(\alpha)$, there exits a partition P_1 and P_2 of [a, b] such that

(3.4.4) $$\begin{cases} U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha) < \dfrac{\in}{2} \\ U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha) < \dfrac{\in}{2} \end{cases}$$

These inequalities hold if P_1 and P_2 are replaced by their common refinements P.
Thus using (3.4.3), we have for $f = f_1 + f_2$,

$$\begin{aligned}
U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\
&\le \sum_{i=1}^{n} (M_i' - m_i')\Delta\alpha_i + \sum_{i=1}^{n} (M_i'' - m_i'')\Delta\alpha_i \\
&< \frac{\in}{2} + \frac{\in}{2} \quad \text{(using 3.4.4)} \\
&= \in.
\end{aligned}$$

Hence $f = f_1 + f_2 \in \Re(\alpha)$.

Further, we note that

$M_i' - m_i'' \le m_i \le M_i \le M_i' + M_i''$

Multiplying by $\Delta\alpha_I$ and adding for I = 1, 2, …..,n, we get

(3.4.5) $$\begin{aligned} L(P, f_1, \alpha) + L(P, f_2, \alpha) \le L(P, f, \alpha) &\le U(P, f, \alpha) \\ &\le U(P, f_1, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha) \end{aligned}$$

Also

(3.4.6) $$U(P, f_1, \alpha) < \int_a^b f_1\, d\alpha + \frac{\in}{2}$$

(3.4.7) $$U(P, f_2, \alpha) < \int_a^b f_2\, d\alpha + \frac{\in}{2}$$

Combining (3.4.5), (3.4.6) and (3.4.7), we have

$$\begin{aligned}
\int_a^b f\, d\alpha &\le U(P, f, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha) \\
&< \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha + \frac{\in}{2} + \frac{\in}{2}
\end{aligned}$$

Since $\in$ is arbitrary positive number, we have

(3.4.8) $$\int_a^b f\, d\alpha \le \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$$

Preceding with $(-f_1)$, $(-f_2)$ in place of f_1 and f_2 respectively, we have

$$\int_a^b (-f)\, d\alpha \le \int_a^b (-f_1)\, d\alpha + \int_a^b (-f_2)\, d\alpha$$

or

(3.4.9) $$\int_a^b f\, d\alpha \ge \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$$

Now (3.4.8) and (3.4.9) yield

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

Theorem 8. If $f \in \Re(\alpha)$ and $f \in \Re(\beta)$ then $f \in \Re(\alpha + \beta)$ and

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

Proof. Since $f \in \Re(\alpha)$ and $f \in \Re(\beta)$, there exists partition P_1 and P_2 such that

$$U(P_1, f, \alpha) - L(P_1, f, \alpha) < \frac{\in}{2}$$

$$U(P_2, f, \beta) - L(P_2, f, \beta) < \frac{\in}{2}$$

These inequalities hold if P_1 and P_2 are replaced by their common refinement P.
Also

$$\Delta(\alpha_i + \beta_i) = [\alpha(x_i) - \alpha(x_{i-1})] + [\beta(x_i) - \beta(x_{i-1})]$$

Hence, if M_i and m_i are bounds of f in (x_{i-1}, x_i),

$$U(P, f, (\alpha + \beta)) - L(P, f, (\alpha + \beta)) = \sum_{i=1}^{n} (M_i - m_i)\Delta(\alpha_i + \beta_i)$$

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

$$< \frac{\in}{2} + \frac{\in}{2} = \in.$$

Hence $f \in \Re(\alpha + \beta)$.
Further

$$U(P, f, \alpha) < \int_a^b f\, d\alpha + \frac{\in}{2}$$

$$U(P, f, \alpha) < \int_a^b f\, d\beta + \frac{\in}{2}$$

and

$$U(P, f, \alpha+\beta) = \sum M_i\, \Delta\alpha_i + \sum M_i\, \Delta\beta_i$$

Also, then

$$\int_a^b f\, d(\alpha + \beta) \le U(P, f, \alpha + \beta) = U(P, f, \alpha) + U(P, f, \beta)$$

$$< \int_a^b f\, d\alpha + \frac{\in}{2} + \int_a^b f\, d\beta + \frac{\in}{2}$$

$$= \int_a^b f\, d\alpha + \int_a^b f\, d\beta + \in$$

Since $\in$ is arbitrary positive number, therefore

$$\int_a^b f\, d(\alpha + \beta) \le \int_a^b f\, d\alpha + \int_a^b f\, d\beta.$$

Replacing f by –f, this inequality is reversed and hence

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

Theorem 9. If $f \in \Re(\alpha)$ on [a, b], then $f \in \Re(\alpha)$ on [a, c] and $f \in \Re(\alpha)$ on [c, b] where c is a point of [a, b] and

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

Proof. Since $f \in \Re(\alpha)$, there exits a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \in, \in > 0.$$

Let P* be a refinement of P such that P* = P U{c}. Then

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

which yields

(3.4.10) $$U(P^*, f, \alpha) - L(P^*, f, \alpha) \le U(P, f, \alpha) - L(P, f, \alpha) < \in$$

Let P_1 and P_2 denote the sets of points of P* between [a, c], [c, b] respectively. Then P_1 and P_2 are partitions of [a, c] and [c, b] and $P^* = P_1 \cup P_2$. Also

(3.4.11) $$U(P^*, f, \alpha) = U(P_1, f, \alpha) + U(P_2, f, \alpha)$$

and

(3.4.12) $$L(P^*, f, \alpha) = L(P_1, f, \alpha) + L(P_2, f, \alpha)$$

Then (3.4.10), (3.4.11), and (3.4.11) imply that

$$U(P^*, f, \alpha) - L(P^*, f, \alpha) = [U(P_1, f, \alpha) - L(P_1, f, \alpha)] + [U(P_2, f, \alpha) - L(P_2, f, \alpha)] < \in$$

Since each of $U(P_1, f, \alpha) - L(P_1, f, \alpha)$ and $U(P_2, f, \alpha) - L(P_2, f, \alpha)$ is non – negative, it follows that

$$U(P_1, f, \alpha) - L(P_1, f, \alpha) < \in$$

and

$$U(P_2, f, \alpha) - L(P_2, f, \alpha) < \in$$

Hence f is integrable on [a, c] and [c, b].

Taking inf for all partitions, the relation (3.4.11) yields

(3.4.13) $$\overline{\int} f\, d\alpha \ge \overline{\int_a^c} f\, d\alpha + \overline{\int_c^b} f\, d\alpha$$

But since f in integrable on [a, c] and [c, b], we have

(3.4.14) $$\int_a^b f(x)\, d\alpha \ge \int_a^c f\, d\alpha + \int_c^b f\, d\alpha$$

The relation (3.4.12) similarly yields

(3.4.15) $\int_a^b f\,d\alpha \le \int_a^c f\,d\alpha + \int_c^b f\,d\alpha$

Hence (3.4.14) and (3.4.15) imply that

$$\int_a^b f\,d\alpha = \int_a^c f\,d\alpha + \int_c^b f\,d\alpha$$

Theorem 10. If $f \in \Re(\alpha)$, then

(i) $cf \in \Re(\alpha)$ and $\int_a^b (cf)\,d\alpha = c\int_a^b f\,d\alpha$, for every constant c,

(ii) If in addition $|f(x)| \le K$ on [a, b], then

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

Proof. (i) Let $f \in \Re(\alpha)$ and let $g = cf$. Then

$$U(P, g, \alpha) = \sum_{i=1}^{n} M_i'\Delta\alpha_i = \sum_{i=1}^{n} cM_i\Delta\alpha_i$$

$$= c\sum_{i=1}^{n} M_i\Delta\alpha_i$$

$$= c\,U(P, f, \alpha)$$

Similarly

$$L(P, g, \alpha) = c\,L(P, f, \alpha)$$

Since $f \in \Re(\alpha)$, $\exists$ a partition P such that for every $\in > 0$,

$$U(P, f, \alpha) - L(P, f, \alpha) < \frac{\in}{c}$$

Hence

$$U(P, g, \alpha) - L(P, g, \alpha) = c\,[\,U(P, f, \alpha) - L(P, f, \alpha)]$$

$$< c\,\frac{\in}{c} = \in.$$

Hence $g = c\,f \in \Re(\alpha)$.

Further, since $U(P, f, \alpha) < \int_a^b f\,d\alpha + \frac{\in}{2}$,

$$\int_a^b g\,d\alpha \le U(P, g, \alpha) = c\,U(P, f, \alpha)$$

$$< c\left(\int_a^b f\,d\alpha + \frac{\in}{2}\right)$$

Since $\in$ is arbitrary

$$\int_a^b g\,d\alpha \le c\int_a^b f\,d\alpha$$

Replacing f by –f, we get

$$\int_a^b g\,d\alpha \ge c\int_a^b f\,d\alpha$$

Hence $\int_a^b (cf)\, d\alpha = c \int_a^b f\, d\alpha$

(ii) If M and m are bounds of $f \in \Re(\alpha)$ on [a, b], then it follows that

(3.4.16) $m[\alpha(b) - \alpha(a)] \le \int_a^b f\, d\alpha \le M[\alpha(b) - \alpha(a)]$ for $b \ge a$.

In fact, if a = b, then (3.4.16) is trivial. If b > a, then for any partition P, we have

$$m[\alpha(b) - \alpha(a)] \le \sum_{i=1}^{n} m_i \Delta\alpha_i = L(P, f, \alpha)$$
$$\le \int_a^b f\, d\alpha$$
$$\le U(P, f, \alpha) = \sum M_i \Delta\alpha_i$$
$$\le M\,(b - a)$$

which yields

(3.4.17) $\quad m\,[\alpha(b) - \alpha(a)] \le \int_a^b f\, d\alpha \le M\,(b - a)$

Since $|f(x)| \le k$ for all $x \in (a, b)$, we have

$$-k \le f(x) \le k$$

so if m and M are the bounds of f in (a, b),

$$-k \le m \le f(x) \le M \le k \text{ for all } x \in (a, b).$$

If $b \ge a$, then $\alpha(b) - \alpha(a) \ge 0$ and we have by (3.4.17)

$$-k[\alpha(b) - \alpha(a)] \le m[\alpha(b) - \alpha(a)] \le \int_a^b f\, d\alpha$$
$$\le M[\alpha(b) - \alpha(a)] \le k[\alpha(b) - \alpha(a)]$$

Hence

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

Theorem 11. Suppose $f \in \Re(\alpha)$ on [a, b], $m \le f \le M$, φ is continuous on [m, M] and $h(x) = \phi[f(x)]$ on [a, b]. Then $h \in \Re(\alpha)$ on [a, b].

Proof. Let $\in > 0$. Since φ is continuous on closed and bounded interval [m, M], it is uniformly continuous on [m, M]. Therefore there exists $\delta > 0$ such that $\delta < \in$ and

$$|\phi(s) - \phi(t)| < \in \text{ if } |s - t| \le \delta,\ s, t \in [m, M].$$

Since $f \in \Re(\alpha)$, there is a partition $P = \{x_0, x_1, \ldots\ldots, x_n\}$ of [a, b] such that

(3.4.18) $U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$.

Let M_i, m_i and M^*_i, m_i^* be the lub, g. l. b of f(x) and $\phi(x)$ respectively in $[x_{i-1}, x_i]$. Divide the number 1,2,…..,n into two classes :

$$i \in A \text{ if } M_i - m_i < \delta$$

and

$$i \in B \text{ if } M_i - m_i \ge \delta.$$

For $i \in A$, our choice of δ implies that $Mi^* - m_i^* \le \in$. Also, for $i \in B$, $M_i^* - m_i^* \le 2k$ where $k = \text{lub} |\varphi(t)|$, $t \in [m, M]$. Hence, using (3.4.18), we have

(3.5.19) $\quad \delta \sum_{i\in B} \Delta\alpha_i \le \sum_{i\in B} (M_i - m_i)\Delta\alpha_i < \delta^2$

so that $\sum_{i\in B}\Delta\alpha_i < \delta$. Then we have

$$U(P, h, \alpha) - L(P, h, \alpha) = \sum_{i\in A}(M_i * - m_i *)\Delta\alpha_i + \sum_{i\in B}(M_i * - m_i *)\Delta\alpha_i$$
$$\leq \in[\alpha(b) - \alpha(a)] + 2\,k\delta$$
$$< [\alpha(b) - \alpha(a)] + 2k]$$

Since $\in$ was arbitrary,

$$U(P, h, \alpha) - L(P, h, \alpha) < \in^*, \in^* > 0.$$

Hence $h \in f(\alpha)$.

Theorem 12. If $f \in \Re(\alpha)$ and $g \in \Re(\alpha)$ on [a, b], then $f\,g \in \Re$, $|\,f\,| \in \Re(\alpha)$ and

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

Proof. Let ϕ be defined by $\phi(t) = t^2$ on (a,b]. Then $h(x) = \phi[f(x)] = f^2 \in \Re(\alpha)$ by Theorem 11. Also

$$fg = \frac{1}{4}\,[(f + g)^2 - (f - g)^2].$$

Since $f, g \in \Re(\alpha)$, $f + g \in \Re(\alpha)$, $f - g \in \Re(\alpha)$. Then, $(f + g)^2$ and $(f - g)^2 \in \Re(\alpha)$ and so their difference multiplied by $\frac{1}{4}$ also belong to $\Re(\alpha)$ proving that $fg \in \Re$.

If we take $\phi(f) = |\,t\,|$, again Theorem 11 implies that $|\,f\,| \in \Re(\alpha)$. We choose $c = \pm 1$ so that

$$c\int f\,d\alpha \geq 0$$

Then

$$\left|\int f\,d\alpha\right| = c\int f\,d\alpha = \int c\,f\,d\alpha \leq \int |\,f\,|\,d\alpha$$

because $cf \leq |\,f\,|$.

3.5. Riemann-Stieltjes integral as limit of sums. In this section, we shall show that Riemann-Stieltjes integral $\int f\,d\alpha$ can be considered as the limit of a sequence of sums in which M_i, m_i involved in the definition of $\int f\,d\alpha$ are replaced by values of f.

Definition. Let $P = \{a = x_0, x_1, \ldots\ldots, x_n = b\}$ be a partition of [a, b] and let points $t_1, t_2, \ldots.., t_n$ be such that $t_I \in [x_{i-1=}, x_i]$. Then the sum

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

is called a **Riemann-Stieltjes sum of f with respect to α.**

Definition. We say that

$$\lim_{|P|\to 0} S(P, d, \alpha) = A$$

If for every $\in > 0$, there exists a $\delta > 0$ such that $|\,P\,| < \delta$ implies

$$|\,S(P, f, \alpha) - A\,| < \in.$$

Theorem 13. If $\lim_{|P|\to 0} S(P, f, \alpha)$ exists, then $f \,\varepsilon\, \Re\,(\alpha)$ and

$$\lim_{|P|\to 0} S(P, f, \alpha) = \int_a^b f\,d\alpha.$$

Proof. Suppose $\lim_{|P|\to 0} S(P, f, \alpha)$ exists and is equal to A. Then given $\in > 0$ there exists a $\delta>0$ such that $|\,P| < \delta$ implies

$$|\,S(P, f, \alpha) - A\,| < \frac{\in}{2}$$

or

(3.5.1) $A - \frac{\in}{2} < S(P, f, \alpha) < A + \frac{\in}{2}$

If we choose partition P satisfying $|P| < \delta$ and if we allow the points t_i to range over $[x_{i-1}, x_i]$, taking lub and glb of the numbers $S(P, f, \alpha)$ obtained in this way, the relation (3.5.1) gives

$$A - \frac{\in}{2} \leq L(P, f, \alpha) \leq U((P, f, \alpha) \leq (U, f, \alpha) \leq A + \frac{\in}{2}$$

and so

$$U((P, f, \alpha) - L(P, f, \alpha) < \frac{\in}{2} + \frac{\in}{2} = \in$$

Hence $f \in \Re(\alpha)$. Further

$$A - \frac{\in}{2} \leq L(P, f, \alpha) \leq \int f d\alpha \leq U(P, f, \alpha) \leq A + \frac{\in}{2}$$

which yields

$$A - \frac{\in}{2} \leq \int f\, d\alpha \leq A + \frac{\in}{2}$$

or

$$\int f\, d\alpha = A = \lim_{|P| \to 0} S(P, f, \alpha).$$

Theorem 14. If

(i) f is continuous, then

$$\lim_{|P| \to 0} S(P, f, \alpha) = \int_a^b f\, d\alpha$$

(ii) $f \in \Re(\alpha)$ and α is continuous on [a, b], then

$$\lim_{|P| \to 0} S(P, f, \alpha) = \int_a^b f\, d\alpha$$

Proof. Part (i) is already proved in Theorem 5(ii) of this chapter.

(ii) Let $f \in \Re(\alpha)$, α be continuous and $\in > 0$. Then there exists a partition P* such that

(3.5.2) $U(P^*, f, \alpha) < \int f\, d\alpha + \frac{\in}{4}$

Now, α being uniformly continuous, there exists $\delta_1 > 0$ such that for any partition P of [a,b] with $|P| < \delta_1$, we have

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) < \frac{\in}{4Mn} \text{ for all i}$$

where n is the number of intervals into which [a, b] is divided by P*. C onsider the sum U(P, f, α). Those intervals of P which contain a point of P* in their interior contribute no more than

(3.5.3) $(n - 1) \max \Delta\alpha_i . M < \frac{(n-1) \in M}{4Mn} < \frac{\in}{4}$.

Then (3.5.2) and (3.5.3) yield

(3.5.4) $U(P, f, \alpha) < \int f\, d\alpha + \frac{\in}{2}$

for all P with $| P | < \delta_1$.
Similarly, we can show that there exists a $\delta_2 > 0$ such that

(3.5.5) $L(P, f, \alpha) > \int f\, d\alpha - \frac{\in}{2}$

for all P with $| P | < \delta_2$.
Taking $\delta = \min(\delta_1, \delta_2)$, it follows that (3.5.3) and (3.5.4) hold for every P such that $| P | < \delta$.
Since

$$L(P, f, \alpha) \le S(P, f, \alpha) < U(P, f, \alpha)$$

(3.5.4) and (3.5.5) yield

$$S(P, f, \alpha) < \int f\, d\alpha + \frac{\in}{2}$$

and

$$S(P, f, \alpha) < \int f\, d\alpha - \frac{\in}{2}$$

Hence

$$| S(P, f, \alpha) - \int f\, d\alpha | < \frac{\in}{2}$$

for all P such that $| P | < \delta$ and so

$$\lim_{|P|\to 0} S(P, f, \alpha) = \int f\, d\alpha$$

This completes the proof of the theorem.

The **Abel's Transformation (Partial Summation Formula)** for sequences reads as follows:
Let $<a_n>$ and $<b_n>$ be sequences and let

$$A_n = a_0 + a_1 + \ldots\ldots + a_n \qquad (A_{-1} = 0),$$

then

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

3.6. Integration and Differentiation. In this section, we show that integration and differentiation are inverse operations.

Definition. If $f \in \Re$ on [a, b], than the function F defined by

$$F(t) = \int_a^t f(x)\, dx,\ t \in [a, b]$$

is called the "**Integral Function**" of the function f.

Theorem 15. If $f \in \Re$ on [a, b], then the integral function F of f is continuous on [a, b].

Proof. We have

$$F(t) = \int_a^t f(x)\, dx$$

Since $f \in \Re$, it is bounded and therefore there exists a number M such that for all x in [a, b], $|f(x)| \le M$.
Let $\in$ be any positive number and c any point of [a, b]. Then

$$F(c) = \int_a^c f(x)\,dx,\ F(c+h) \int_a^{c+h} f(x)\,dx.$$

Therefore

$$|F(c+h) - F(c)| = |\int_a^{c+h} f(x)\,dx - \int_a^c f(x)\,dx$$

$$= |\int_a^{c+h} f(x)\,dx|$$

$$\leq M|h|$$

$$< \in \text{ if } |h| < \frac{\in}{M}$$

Thus $|(c+h) - c| < \delta = \dfrac{\in}{M}$ implies $|F(c+h) - F(c) < \in$. Hence F is continuous at any point $C \in [a, b]$ and is so continuous in the interval [a, b].

Theorem 16. If f is continuous on [a, b], then the integral function F is differentiable and $F'(x_0) = f(x_0)$, $x \in [a, b]$.

Proof. Let f be continuous at x_0 in [a, b]. Then there exists $\delta > 0$ for every $\in > 0$ such that

(3.6.1) $\quad |f(t) - f(x_0)| < \in$

whenever $|t - x_0| < \delta$. Let $x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta$ and $a \leq s < t \leq b$, then

$$\frac{F(t)-F(s)}{t-s} - f(x_0) = |\frac{1}{t-s}\int_s^t f(x)\,dx - f(x_0)|$$

$$= |\frac{1}{t-s}\int_s^t f(x)\,dx - \frac{1}{t-s}\int_s^t f(x_0)\,dx|$$

$$= \frac{1}{t-s}|\int_s^t [f(x) - f(x_0)]dx| \leq \frac{1}{t-s}|\int_s^t f(x) - f(x_0)|dx < \in,$$

(using (3.6.1)).

Hence $F'(x_0) = f(x_0)$. This completes the proof of the theorem

Definition. A derivable function F such that F′ is equal to a given function f in [a, b] is called **Primitive of** f.

Thus the above theorem asserts that "Every continuous function f possesses a Primitive, viz the integral function $\int_a^t f(x)\,dx$"

Furthermore, the continuity of a function is not necessary for the existence of primitive. In other words, the function possessing primitive are not necessary continuous. For example, consider the function f on [0, 1] defined by

$$f(x) = \begin{cases} 2x\sin\frac{1}{x} - \cos\frac{1}{x}, x \neq 0 \\ 0,\ x = 0 \end{cases}$$

It has primitive

$$F(x) = \begin{cases} x^2 \sin\frac{1}{x}, x \neq 0 \\ 0, x = 0 \end{cases}$$

Clearly $F'(x) = f(x)$ but f(x) is not continuous at x = 0, i.e., f is not continuous in [0,1].

Theorem 17.(Fundamental Theorem of the Integral Calculus). If $f \in \mathfrak{R}$ on [a, b] and if there is a differential function F on [a, b] such that $F' = f$, then

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

Proof. Let P be a partition of [a, b] and choose t_i (I = 1, 2, ……,n) such that $x_{i-1} \le t_i \le x_i$. Then, by Lagrange's Mean Value Theorem, we have

$$F(x_i) - F(x_{i-1}) = (x_i - x_{i-1})\,F'(t_i) = (x_i - x_{i-1})\,f(t_i) \qquad \text{(since } F' = f\text{)}.$$

Further

$$\begin{aligned} F(b) - F(a) &= \sum_{i=1}^{n} [F(x_i) - F(x_{i-1})] \\ &= \sum_{i=1}^{n} f(t_i)\,(x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} f(t_i)\,\Delta x_i \end{aligned}$$

and the last sum tends to $\int_a^b f(x)\,dx$ as $|P| \to 0$, by Theorem 13 taking $\alpha(x) = x$. Hence

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

This completes the proof of the Theorem.

The next theorem tells us that the symbol $d\alpha(x)$ can be replaced by $\alpha'(x)\,dx$ in the Riemann – Stieltjes integral $\int_a^b f(x)\,d\alpha(x)$. This is the situation in which Riemann – Stieltjes integral reduces to Riemann integral.

Theorem 18. If $f \in \mathfrak{R}$ and $\alpha' \in \mathfrak{R}$ on [a, b], then $f \in \mathfrak{R}(\alpha)$ and

$$\int_a^b f\,d\alpha = \int_a^b f(x)\,\alpha'(x)\,dx.$$

Proof. Since $f \in \mathfrak{R}$, $\alpha' \in \mathfrak{R}$, it follows that their product $f\alpha' \in \mathfrak{R}$. Let $\in > 0$ be given. Choose M such that $|f| \le M$. Since $f\alpha' \in \mathfrak{R}$ and $\alpha' \in \mathfrak{R}$, using Theorem 14(ii) for integrator as x, we have

(3.6.2) $$\left| \sum f(t_i)\,\alpha'(t_i)\,\Delta x_i - \int f\alpha' \right| < \in$$

if $|P| < \delta_1$ and $x_{i-1} \le t_i \le x_i$ and

(3.6.3) $$\left| \sum \alpha'(t_i)\,\Delta x_i - \int \alpha' \right| < \in$$

if $|P| < \delta_2$ and $x_{i-1} \le t_i \le x_i$. Letting t_i vary in (3.6.3), we have

(3.6.4) $$\left| \sum \alpha'(s_i)\,|\,\Delta x_i - \int \alpha' \right| < \in$$

if $|P| < \delta_2$ and $x_{i-1} \le s_i \le x_i$. From (3.6.3) and (3.6.4) it follows that

$$\begin{aligned} &\left| \sum \alpha'(t_i)\,|\,\Delta x_i - \int \alpha' + \int \alpha' - \sum \alpha'(s_i)\,\Delta x_i \right| \\ &\le \left| \sum \alpha'(t_i)\,\Delta x_i - \int \alpha' \right| + \left| \sum \alpha'(S_i)\,\Delta x_i - \int \alpha' \right| \\ &< \in + \in = 2\in \end{aligned}$$

or

(3.6.5) $\sum |\alpha'(t_i) - \alpha'(s_i)| \; \Delta x_I < 2\in$

if $| P | < \delta_2$ and $x_{i-1} \le t_i \le x_i$, $x_{i-1} \le s_i \le x_i$.

Now choose a partition P so that $| P | < \delta = \min(\delta_1, \delta_2)$ and choose $t_i \in [x_{i-1}, x_i]$. By Mean Value Theorem,

$$\Delta \alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(S_i)\,(x_i - x_{i-1})$$
$$= \alpha'(S_i)\; \Delta x_i$$

Then, we have

(3.6.6) $\sum f(t_i)\; \Delta \alpha_i = \sum f(t_i)\, \alpha'(t_i)\; \Delta x_i + \sum f(t_i)[\, \alpha'(S_i) - \alpha'(t_i)]\; \Delta x_i.$

Thus, by (3.6.2) and (3.6.5), it follows that

$$|\sum f(t_i)\; \Delta \alpha_i - \int f\,\alpha' \,| = |\sum f(t_i)\,\alpha'(t_i)\; \Delta x_i - \int f\,\alpha'$$
$$+ \sum f(t_i)[\,\alpha'(S_i) - \alpha'(t_i)]\; \Delta x_i\,|$$
$$< \in + 2 \in M = \in(1 + 2M)$$

Hence

$$\lim_{|p|\to 0} \sum f(t_i)\; \Delta x_i = \int_a^b f(x)\,\alpha'(x)\,dx$$

or

$$\int_a^b f\,d\alpha = \int f(x)\,\alpha'(x)\,dx$$

Example. Evaluate (i) $\int_0^2 x^2\,dx^2$, (ii) $\int_0^2 [x]\,dx^2$

Solution. We know that

$$\int_a^b f\,d\alpha = \int_a^b f\,(x)\,\alpha'(x)\,dx$$

Therefore

$$\int_0^2 x^2\,dx^2 = \int_0^2 x^2(2x)\,dx^2 = \int_0^2 2x^3\,dx$$
$$= 2\left|\,\frac{x^4}{4}\right|_0^2 = 8 \text{ Ans.}$$

and

$$\int_0^2 [x]\,dx^2 = \int_0^2 [x]\,2x\,dx$$
$$= \int_0^1 [x]\,2x\,dx + \int_1^2 [x]\,2x\,dx$$
$$= 0 + \int_1^2 2x\,dx = 0 + 2\left|\,\frac{x^2}{2}\right|_1^2$$
$$= 0 + 3 = 3 \text{ Ans.}$$

We now establish a connection between the integrand and the integrator in a Riemann – Stieltjes integral. We shall show that existence of $\int f\,d\alpha$ implies the existence of $\int \alpha\,df$.

We recall that Abel's transformation (Partial Summation Formula) for sequences reads as follows:

"Let $<a_n>$ and $<b_n>$ be two sequences and let $A_n = a_0 + a_1 + \ldots + a_n$ $(A_{-1} = 0)$. Then

(3.6.7) $$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p.”$$

Theorem (Integration by parts). If $f \in \Re(\alpha)$ on [a, b], then $\alpha \in \Re(f)$ on [a, b] and

$$\int f(x)\, d\alpha(x) = f(b)\, \alpha(b) - f(a)\, \alpha(a) - \int \alpha(x)\, df(x)$$

(Due to analogy with (3.6.7), the above expression is also known as **Partial Integration Formula**).

Proof. Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be a partition of [a, b]. Choose $t_1, t_2, \ldots, t_n$ such that $x_{i-1} \le t_i \le x_i$ and take $t_0 = a$, $t_{n+1} = b$. Suppose Q is the partition $\{t_0, t_1, \ldots, t_{n+1}\}$ of [a, b]. By partial summation, we have

$$S(P, f, \alpha) = \sum_{i=1}^{n} f(t_i)[\alpha(x_i) - \alpha(x_{i-1})] = f(b)\, \alpha(b) - f(a)\, \alpha(a) - \sum_{i=1}^{n+1} \alpha(x_{i-1})[f(t_i) - f(t_{i-1})]$$

$$= f(b)\, \alpha(b) - f(a)\, \alpha(a) - S(Q, \alpha, f)$$

since $t_{i-1} \le x_{i-1} \le t_i$. If $|P| \to 0$, $|Q| \to 0$, then

$S(P, f, \alpha) \to \int f\, d\alpha$ and $S(Q, \alpha, f) \to \int \alpha\, df$.

Hence

$$\int f\, d\alpha = f(b)\, \alpha(b) - f(a)\, \alpha(a) - \int \alpha\, df$$

3.7. Mean Value Theorems For Riemann – Stieltjes Integrals. In this , section, we establish Mean Value Theorems which are used to get estimate value of an integral rather than its exact value.

Theorem 19 (First Mean Value Theorem for Riemann – Stieltjes Integral). If f is continuous and real valued and α is monotonically increasing on [a, b], then there exists a point x in [a, b] such that

$$\int f\, d\alpha = f(x)\, [\alpha(b) - \alpha(a)]$$

Proof. If $\alpha(a) = \alpha(b)$, the theorem holds trivially, both sides being 0 in that case (α become constant and so $d\alpha = 0$). Hence we assume that $\alpha(a) < \alpha(b)$. Let

$$M = \text{lub } f(x),\ m = \text{glb } f(x).\ a \le x \le b$$

Then

$$m \le f(x) \le M$$

or

$$m[\alpha(b) - \alpha(a)] \le \int f\, d\alpha \le M[\alpha(b) - \alpha(a)]$$

Hence there exists some c satisfying $m \le c \le M$ such that

$$\int_a^b f\, d\alpha = c[\alpha(b) - \alpha(a)]$$

Since f is continuous, there is a point $x \in [a, b]$ such that $f(x) = c$ and so we have

$$\int_a^b f(x)\, d\alpha(x) = f(x)[\alpha(b) - \alpha(a)]$$

This completes the proof of the theorem.

Theorem 20 (Second Mean – value Theorem for Riemann – Stieltjes Integral). Let f be monotonic and α real and continuous. Then there is a point $x \in [a, b]$ such that

$$\int_a^b f\, d\alpha = f(a)[\alpha(x) - \alpha(a)] + f(b)[\alpha(b) - \alpha(x)]$$

Proof. By Partial Integration Formula, we have

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

The use of First Mean –Value Theorem of Riemann – Stieltjes integral yields that there is x in [a, b] such that

$$\int_a^b \alpha df = \alpha(x)[f(b) - f(a)]$$

Hence, for some $x \in [a, b]$, we have.

$$\int_a^b f\, d\alpha = f(b)\alpha(b) - f(a)\alpha(a) - \alpha(x)[f(b) - f(a)]$$

$$= f(a)[\alpha(x) - \alpha(a)] + f(b)[\alpha(b) - \alpha(x)]$$

which proves the theorem.

3.8. We discuss now change of variable. In this direction we prove the following result.

Theorem 21. Let f and ϕ be continuous on [a, b]. If φ is strictly increasing on $[\alpha, \beta]$, where $a = \phi(\alpha)$, $b = \phi(b)$, then

$$\int_a^b f(x)\, dx = \int_\alpha^\beta f(\phi(y))\, d\phi(y)$$

(this corresponds to change of variable in $\int_a^b f(x)\, dx$ by taking $x = \varphi(y)$].

Proof. Since ϕ is strictly monotonically increasing, it is invertible and so

$\alpha = \phi^{-1}(a)$, $\beta = \phi^{-1}(b)$.

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be any partition of [a, b] and $Q = \{\alpha = y_0, y_1, \ldots, y_n = \beta\}$ be the corresponding partition of $[\alpha, \beta]$, where $y_i = \phi^{-1}(x_i)$. Then

$$\begin{aligned}\Delta x_i &= x_i - x_{i-1}\\ &= \phi(y_i) - \phi(y_{i-1})\\ &= \Delta\phi_i\end{aligned}$$

Let for any $c_i \in \Delta x_i$, $d_i \in \Delta y_i$, where $c_i = \varphi(d_i)$. Putting $g(y) = f[\varphi(y)]$, we have

(3.8.1) $$S(P, f) = \sum_{i=1}^{n} f(c_i)\, \Delta x_i$$

$$= \sum_i f(\phi(d_i))\Delta\phi_i$$

$$= \sum_i g(d_i)\, \Delta\varphi_i$$

$$= S(Q, g, \varphi)$$

Continuity of f implies that $S(P, f) \to \int_a^b f(x)\, dx$ as $|P| \to 0$ and continuity of g implies that

$$S(Q, g, \varphi) \to \int_a^b g(y)\, d\varphi \text{ as } |Q| \to 0.$$

Since uniform continuity of φ on [a, b] implies that $|Q| \to 0$ as $|P| \to 0$. Hence letting $|P| \to 0$ in (3.8.1), we have

$$\int_a^b f(x)\, dx = \int_\alpha^\beta g(y)\, d\phi = \int_\alpha^\beta f[(\phi y)]\, d\phi(y)$$

This completes the proof of the theorem.

3.9. Integration of Vector – Valued Functions. Let $f_1, f_2, \ldots, f_k$ be real valued functions defined on [a, b] and let **f** = $(f_1, f_2, \ldots, f_k)$ be the corresponding mapping of [a, b] into $\mathbf{R}^k$.
Let α be a monotonically increasing function on [a, b]. If $f_i \in \Re(\alpha)$ for i = 1, 2, ….., k, we say that $\mathbf{f} \in \Re(\alpha)$ and then the integral of **f** is defined as

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

Thus $\int_a^b \mathbf{f}\, d\alpha$ is the point in $\mathbf{R}^k$ whose ith coordinate is $\int f_i\, d\alpha$.

It can be shown that if $\mathbf{f} \in \Re(\alpha)$, $\mathbf{g} \in \Re(\alpha)$,
then

(i) $$\int_a^b (\mathbf{f} + \mathbf{g})\, d\alpha = \int_a^b \mathbf{f}\, d\alpha + \int_a^b \mathbf{g}\, d\alpha$$

(ii) $$\int_a^b \mathbf{f}\, d\alpha = \int_a^c \mathbf{f}\, d\alpha + \int_a^b \mathbf{f}\, d\alpha, \quad a < c < b.$$

(iii) if $\mathbf{f} \in \Re(\alpha_1)$, $\mathbf{f} \in \Re(\alpha_2)$, then $\mathbf{f} \in \Re(\alpha_1 + \alpha_2)$

and

$$\int_a^b \mathbf{f}\, d(\alpha_1 + \alpha_2) = \int_a^b \mathbf{f}\, d\alpha_1 + \int_a^b \mathbf{f}\, d\alpha_2$$

To prove these results, we have to apply earlier results to each coordinate of **f**. Also, fundamental theorem of integral calculus holds for vector valued function **f** . We have

Theorem 22. If **f** and **F** map[a, b] into $\Re^k$, if $\mathbf{f} \in \Re(\alpha)$ if $\mathbf{F}' = \mathbf{f}$, then

$$\int_a^b \mathbf{f}(t)\, dt = \mathbf{F}(b) - \mathbf{F}(a)$$

Theorem 23. If **f** maps [a, b] into $\mathbf{R}^k$ and if $f \in \mathbf{R}(\alpha)$ for some monotonically increasing function α on [a, b], then $|\mathbf{f}| \in \mathbf{R}(\alpha)$ and

$$\left| \int_a^b \mathbf{f}\, d\alpha \right| \le \int_a^b |\mathbf{f}|\, d\alpha.$$

Proof. Let

$$\mathbf{f} = (f_1, \ldots, f_k).$$

Then

$$|\mathbf{f}| = (f_1^2 + \ldots + f_h^{\,h})^{1/2}$$

Since each $f_i \in \mathbf{R}(\alpha)$, the function $f_i^2 \in \mathbf{R}(\alpha)$ and so their sum $f_1^2 + \ldots + f_k^2 \in \mathbf{R}^{(\alpha)}$. Since x^2 is a continuous function of x, the square root function is continuous on [0, M] for every real M. Therefore $|\mathbf{f}| \in \mathbf{R}(\alpha)$.

Now, let $\mathbf{y} = (y_1, y_2, \ldots, y_k)$, where $y_i = \int f_i\, d\alpha$, then

$$\mathbf{y} = \int \mathbf{f}\, d\alpha$$

and

$$|\mathbf{y}|^2 = \sum_i y_i^2 = \sum y_i \int f_i\, d\alpha$$

$$= \int \left(\sum y_i f_i\right) d\alpha$$

But, by Schwarz inequality

$$\sum y_i f_i(t) \le |\mathbf{y}|\,|\mathbf{f}(t)| \quad (a \le t \le b)$$

Then

(3.9.1) $$|\mathbf{y}|^2 \le |\mathbf{y}| \int |f|\,d\alpha$$

If $\mathbf{y} = \mathbf{0}$, then the result follows. If $\mathbf{y} \ne \mathbf{0}$, then divide (3.9.1) by $|\mathbf{y}|$ and get

$$|\mathbf{y}| \le \int |\mathbf{f}|\,d\alpha$$

or $$\int_a^b |\mathbf{f}|\,d\alpha \le \int |\mathbf{f}|\,d\alpha.$$

3.10. Rectifiable Curves. The aim of this section is to consider application of results studied in this chapter to geometry.

Definition. A continuous mapping γ of an interval [a, b] into $\mathbf{R}^k$ is called a curve in $\mathbf{R}^k$.
If $\gamma : [a, b] \to \mathbf{R}^k$ is continuous and one – to – one, then it is called an arc.
If for a curve $r : [a, b] \to \mathbf{R}^k$,

$$r(a) = r(b)$$

but

$$r(t_1) \ne r(t_2)$$

for every other pair of distinct points t_1, t_2 in [a, b], then the curve γ is called a simple closed curve.

Definition. Let $\mathbf{f} : [a, b] \to \mathbf{R}^k$ be a map. If $P = \{x_0, x_1, \ldots, x_n\}$ is a partition of [a, b], then

$$V(\mathbf{f}, a, b) = \text{lub} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

where the lub is taken over all possible partitions of [a, b], is called total variation of **f** on [a,b].
The function **f** is said to be of bounded variation on [a, b] if $V(\mathbf{f}, a, b) < +\infty$.
Definition. A curve $\gamma : [a, b] \to \mathbf{R}^k$ is called rectifiable if γ is of bounded variation. The length of a rectifiable curve γ is defined as total variation of γ, i.e, $V(\gamma, a, b)$. Thus length of rectifiable curve $\gamma = \text{lub} \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|$ for the partition $(a = x_0 < x_1 < \ldots < x_n = b)$.

The ith term $|\gamma(x_i) - \gamma(x_{i-1})|$ in this sum is the distance in $\mathbf{R}^k$ between the points $r(x_{i-1})$ and $r(x_i)$. Further $\sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|$ is the length of a polygon whose vertices are at the points $\gamma(x_0), \gamma(x_1), \ldots, \gamma(x_n)$. As the norm of our partition tends to zero those polygons approach the range of γmore and more closely.

Theorem 24. Let γ be a curve in $\mathbf{R}^k$. If γ' is continuous on [a, b], then γ is rectifiable and has length

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

Proof. It is sufficient to show that $\int |\gamma'| = V(\gamma, a, b)$. So, let $\{x_0, \ldots, x_n\}$be a partition of [a, b]. Using Fundamental Theorem of Calculus for vector valued function, we have

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

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

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

Thus

$$\text{(3.10.1)} \qquad V(\gamma, a, b) \leq \int |\gamma'|.$$

To prove the reverse inequality, let $\in$ be a positive number. Since γ' is uniformly continuous on [a, b], there exists $\delta > 0$ such that

$$|\gamma'(s) - \gamma'(t)| < \in, \text{ if } |s - t| < \delta.$$

If mesh (norm) of the partition P is less then δ and $x_{i-1} \leq t \leq x_i$, then we have

$$|\gamma'(t)| \leq |\gamma'(x_i)| + \in,$$

so that

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)|\, dt - \in \Delta x_i \leq \gamma'(x_i) | \Delta x_I$$

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

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

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

Adding these inequalities for i = 1, 2, …., n, we get

$$\int_a^b |\gamma'(t)|\, dt \leq \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})| + 2 \in (b - a)$$

$$= V(\gamma, a, b) + 2\in (b - a)$$

Since $\in$ is arbitrary, it follows that

$$\text{(3.10.2)} \quad \int_a^b |\gamma'(t)|\, dt \leq V(\gamma, a, b)$$

Combining (3.10.1) and (3.10.2), we have

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

Hence the length of r is $\int_a^b |\gamma'(t)|\, dt$.

PART B : THEORY OF MEASURE AND INTEGRATION

3.11. In this section we shall define Lebesgue Measure, which is a generalization of the idea of length.

Definition. **The length l(I) of an interval I with end points a and b is defined as the difference of the end points. In symbols, we write**

$$l\,(I) = b - a.$$

Definition. A function whose domain of definition is a class of sets is called a Set Function.
For example, length is a set function. The domain being the collection of all intervals.

Definition. An extended real – valued set function μ defined on a class **E** of sets is called Additive if $A \in E$, $B \in$ **E**, $A \cup B \in$ **E** and $A \cap B = \phi$, imply

$$\mu\,(A \cup B) = \mu\,(A) + \mu\,(B)$$

Definition. An extended real valued set function μ defined on a class **E** of sets is called **finitely additive** if for every finite disjoint class $\{A_1, A_2, \ldots, A_n\}$ of sets in E, whose union is also in **E**, we have

$$\mu\,\left(\bigcup_{i=1}^{n} Ai\right) = \sum_{i=1}^{n} \mu(Ai)$$

Definition. An extended real – valued set function μ defined on a class **E** of sets is called countably additive it for every disjoint sequence $\{A_n\}$ of sets in **E** whose union is also in **E**, we have

$$\mu\,\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

Definition. Length of an open set is defined to be the sum of lengths of the open intervals of which it is composed of. Thus, if $\in$ is an open set, then

$$l(G) = \sum_n l(I_n),$$

where

$$G = \bigcup_n I_n,\ I_{n1} \cap I_{n2} = \varphi \text{ if } n_1 \neq n_2.$$

Definition. The Lebesgue Outer Measure or simply the outer measure m* of a set A is defined as

$$m^*(A) = \inf_{A \subseteq \cup I_n} \sum l(I_n),$$

where the infimum is taken over all finite or countable collections of intervals $\{I_n\}$ such that $A \subseteq I_n$.
Since the lengths are positive numbers, it follows from the definition of m* that $m^*(A) \geq 0$ and that $m^* \phi = 0$.
Further, if A is a singleton, then $m^* A = 0$ and also if $A \subseteq B$, then $M^* A \leq M^* B$.

Theorem 25. Outer measure is translation invariant.

Proof. Let A be a set. We shall show that $m^*(A) = m^*(A+x)$,
where $A + x = \{y + x : y \in A\}$.
Let $\{I_n\}$ be collection of intervals $\{I_n\}$ such that $A \subseteq \cup I_n$. Then, by the definition of outer measure, for $\in > 0$, we have

(3.11.1) $$m^*(A) \geq \sum l(I_n) - \in.$$

But, $$A + x \subseteq \cup (I_n + x),$$
so by (3.11.1)

$$m^*(A + x) \leq \sum l(I_n + x) = \sum l(I_n) \leq m^*A + \in$$

Since $\in$ is arbitrary positive number, we have

(3.11.2) $$m^*(A + x) \leq m^*(A)$$

On the other hand

$$A = A + x - x$$

and so

(3.11.3) $m^*(A) \leq m^*(A + x)$

Combining (3.11.2) and (3.11.3), the required result follows.

Theorem 26. **The outer measure of an interval is its length.**

Proof. First assume that I is a bounded closed interval [a, b]. Since for every +ve real number $\in$ the open interval $(a - \in, b + \in)$ covers I, it follows that

$$m^* I \leq l(a - \in, b + \in)$$
$$= b - a + 2\in$$

Since this is true for every $\in < 0$, we must have

$$m^* I \leq b - a = l(I)$$

For this special case I = [a, b], it remains to prove that $m^* I \geq b - a$. Let $\{I_n\}$ be countable collection of open intervals covering I. Then it is sufficient to establish

$$\sum_n l(I_n) \geq b - a$$

Since [a, b] is compact, by Heine Borel Theorem, we can select a finite number of open intervals from $\{I_n\}$such that their union contain I = [a, b]. Let these finite intervals be $J_1, J_2, \ldots, J_p$.Then since $\bigcup_{i=1}^{p} J_i \supset [a, b]$ it is sufficient to prove that

$$\sum_{i=1}^{p} l(J_i) \geq b - a.$$

Since $a \in I$, there exists an open interval (a_1, b_1) from the above mentioned finite number of intervals such that $a_1 < a < b_1$. If $b_1 \leq b$, then $b_1 \in I$. Since b_1 is not covered by the open interval (a_1, b_1), there is an open interval (a_2, b_2) in the finite collection $J_1, \ldots, J_p$ with $a_2 < b_1 < b_2$. Continuing in this fashion we obtain a sequence

$$(a_1, b_1), (a_2, b_2), \ldots (a_n, b_n)$$

in the collection $J_1, J_2, \ldots, J_p$ satisfying

$a_i < b_{i-1} < b_i$

for every i = 2, ..,n.

Since the collection is finite, out process must terminate with an (a_n, b_n) satisfying $b \in (a_n, b_n)$

Then we have

$$\sum_n l(I_n) \geq \sum_{i=1}^{n} l(a_i, b_i)$$
$$= (b_n - a_n) + (b_{n-1} - a_{n-1}) + \ldots.+(b_2 - a_2) + (b_1 - a_1)$$
$$= b_n - (a_n - b_{n-1}) \ldots\ldots(a_2 - b_1) - a_1$$

Since each expression in the bracket is –ve, it follows that

$$\sum_n l(I_n) > b_n - a_1 > b - a.$$

Hence the theorem is proved in this case.

Next, let I be any bounded interval with end points a and b. For every positive real number $\in$, we have

$$[a + \in, b - \in] \subset I \subset [a, b]$$

Therefore

$$b - a - 2\in \leq m^* [a + \in, b - \in]$$
$$\leq m^* I \leq m^* [a, b]$$
$$= b - a.$$

Since this holds for every $\in > 0$, we must have

$$m^* I = b - a = l(I)$$

Finally, let I be unbounded. Then, for every real number r, I contains a bounded interval H of length

$$l\,(H) \geq r.$$

Therefore by the above result

$$m^* I \geq m^* H = l(H) \geq r.$$

Since this holds for every $r \in R$, we must have

$$m^* I = \infty = l(I)$$

This completes the proof of the theorem.

Theorem27. Let $\{A_n\}$ be a countable collection of sets of real numbers. Then

$$m^* (\cup A_n) \leq \sum m^* A_n.$$

Proof. If one of the sets A_n has infinite outer measure, the inequality holds trivially. So suppose $m^* A_n$ is finite. Then, given $\in > 0$, there exists a countable collection $\{ I_{n,\,i}\}$ of open intervals such that $A_n \subset \bigcup_i I_{n,\,i}$ and

$$\sum_i l(I_{n,i}) < m^* A_n + \frac{\in}{2^n},$$

by the definition of $m^* A_n$.

Now the collection $[I_{n,\,i}]_{n,i} = \bigcup_n [I_{n,\,i}]_i$ is countable, being the union of a countable number of countable collections, and covers $\cup A_n$. Thus

$$m^* (\cup A_n) \leq \sum_{n,i} l\,(I_{n,\,i})$$
$$= \sum_n \sum_i l\,(I_{n,\,i})$$
$$< \sum_n \left(m^* A_n + \frac{\in}{2^n} \right)$$
$$= \sum_n m^* A_n + \sum_n \frac{\in}{2^n}$$

$$= \sum_n m^* A_n + \in \sum_n \frac{1}{2^n}$$
$$= \sum_n m^* A_n + \in$$

Since $\in$ is an arbitrary positive number, it follows that

$$m^* (\cup A_n) \leq \sum m^* A_n$$

Cor 1. If A is countable, $m^* A = 0$.

Proof. We know that a countable set is the union of a countable family of singleton. Therefore $A = \cup [x_n]$, which yields

$$m^* A = m^* [\cup(x_n)] \leq \sum m^* [x_n] \text{(by the above theorem)}$$

But as already pointed out outer measure of a singleton is zero. Therefore it follows that

$$m^* A \leq 0$$

Since outer measure is always a non – negative real number, $m^* A = 0$.

Cor 2. Every interval is not countable.

Proof. We know that outer measure of an interval I is equal to its length. Therefore it follows from Cor. 1 that every interval is not countable.

Cor 3. If $m^* A = 0$, then $m^* (A \cup B) = m^* B$.

Proof. **Using the above proposition**

$$m^*(A \cup B) \leq m^* A + m^* B$$
$$= 0 + m^* B \qquad \text{(i)}$$

Also $B \subset A \cup B$

Therefore $\quad m^* B \leq m^* (A \cup B) \qquad$ (ii)

From (i) and (ii) it follows that

$m^* B = m^* (A \cup B)$

Note:- Because of the property $m^* (\cup A_n) \leq \sum m^* A_n$, the function m^* is said to be **countably subadditive**. It would be much better if m^* were also countably additive, that is, if

$$m^* (\cup A_n) = \sum m^* A_n.$$

for every countable collection$[A_n]$ of disjoint sets of real numbers. If we insist on countable additivity, we have to restrict the domain of the function m^* to some subset ***m*** of the set 2^R of all subsets of **R**. The members of ***m*** are called the measurable subsets of **R**. That is, to do so we suitably reduce the family of sets on which m^* is defined. This is done by using the following definition due to Carathedory.

Definition. A set E of real numbers is said to be m^* measurable, if for every set $A \in \mathbf{R}$, we have

$$m^* A = m^* (A \cap E) + m^* (A \cap E^c)$$

Since

$$A = (A \cap E) \cup (A \cap E^c),$$

It follows from the definition that

$$m^* A = m^* [(A \cap E) \cup (A \cap E^c) \leq m^* (A \cap E) + m^* (A \cap E^c)$$

Hence, the above definition reduces to:

A set $E \in \mathbf{R}$ is measurable if and only if for every set $A \in \mathbf{R}$, we have

$m^* A \geq m^* (A \cap E) + m^* (A \cap E^c)$.

For example ϕ is measurable.

Theorem 28. If $m^* E = 0$, then E is measurable.

Proof. Let A be any set. Then $A \cap E \subset E$ and so

$$m^* (A \cap E) \leq m^* E = 0 \qquad \text{(i)}$$

Also $A \supset A \cap E^c$, and so

$$m^* A \geq m^* (A \cap E^c) = m^* (A \cap E^c) + m^* (A \cap E)$$

as $\quad m^* (A \cap E) = 0$ by (i)

Hence E is measurable.

Theorem29. If a set E is measurable, then so is its complement E^c.

Proof. The definition is symmetrical with respect to E^c , and so if E is measurable, its complement E^c is also measurable.

Theorem30. Union of two measurable sets is measurable.

Proof. Let E_1 and E_2 be two measurable sets and let A be any set. Since E_2 is measurable, we have

$$m^* (A \cap E_1^c) = m^* (A \cap E_1^c \cap E_2) + m^* (A \cap E_1^c \cap E_2^c) \qquad \text{(i)}$$

and since

$$A \cap (E_1 \cup E_2) = (A \cap E) \cup [A \cap E_2 \cap E_1^c] \qquad \text{(ii)}$$

Therefore by (ii) we have

$$m^* [A \cap (E_1 \cup E_2)] \leq m^* (A \cap E_1) + m^* [A \cap E_2 \cap E_1^c] \qquad \text{(iii)}$$

Thus

$$\begin{aligned} & m^* [A \cap (E_1 \cup E_2)] + m^* (A \cap E_1^c \cup E_2^c) \\ & \quad \leq m^* (A \cap E_1) + m^* (A \cap E_2 \cup E_1^c) + m^* (A \cap E_1^c \cap E_2^c) \\ & \quad = m^* (A \cap E_1) + m^* (A \cap E_1^c) \quad \text{(by (i))} \\ & \quad \leq m^* A \qquad \text{(since } E_1 \text{ is measurable)} \end{aligned}$$

i.e. $m^* (A \cap (E_1 \cup E_2)) + m^* (A \cap (E_1 \cup E_2)^c) \leq m^* A$

Hence $E_1 \cup E_2$ is measurable.

Cor. If E_1 and E_2 are measurable, then $E_1 \cap E_2$ is also measurable.

In fact we note that E_1, E_2 are measurable $\Rightarrow E_1^c, E_2^c$ are measurable $\Rightarrow E_1^c \cup E_2^c$ is measurable $\Rightarrow (E_1^c \cup E_2^c)^c = E_1 \cap E_2$ is measurable.

Similarly, it can be shown that if E_1 and E_2 are measurable, then $E_1^c \cap E_2^c$ is also measurable.

Definition. Algebra or Boolean Algebra: - A collection **A** of subsets of a set X is called an algebra of sets or a Boolean Algebra if

(i) $A, B \in \mathbf{A} \Rightarrow A \cup B \in \mathbf{A}$

(ii) $A \in \mathbf{A} \Rightarrow A^c \in \mathbf{A}$

(iii) For any two members A and B of **A**, the intersection $A \cap B$ is in **A**.

Because of De Morgan's formulae (i) and (ii) are equivalent to (ii) and (iii).

It follows from the above definition that the collection **M** of all measurable sets is an algebra. The proof is an immediate consequence of Theorems 29 and 30.

Definition. By a **Boolean σ - algebra** or simply a **σ - algebra** or **Borel field** of a collection of sets, we mean a Boolean Algebra **A** of the collection of the sets such that union of any countable collection of members of this collection is a member of **A**.

From De Morgan's formula an algebra of sets is a σ - algebra or Borel field if and only if the intersection of any countable collection of members of **A** is a member of **A**.

Lemma 1. Let A be any set, and $E_1, E_2, \ldots, E_n$ a finite sequence of disjoint measurable sets. Then

$$m^* \left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

Proof. We shall prove this lemma by induction on n. The lemma is trivial for n = 1. Let n > 1 and suppose that the lemma holds for n – 1 measurable sets E_i.

Since E_n is measurable, we have

$$m^* (X) = m^* (X \cap E_n) + m^* (X \cap E_n^c)$$

for every set $X \in \mathbf{R}$. In particular we may take

$$X = A \cap \left[\bigcup_{i=1}^{n} E_i\right].$$

Since $E_1, E_2, \ldots, E_n$ are disjoint, we have

$$X \cap E_n = A \cap [\bigcup_{i=1}^{n} E_i] \cap E_n = A \cap E_n$$

and

$$X \cap E_n^c = A \cap [\bigcup_{i=1}^{n} E_i] \cap E_n^c = A \cap [\bigcup_{i=1}^{n-1} E_i]$$

Hence we obtain

$$m^* X = m^*(A \cap E_n) + m^*(A \cap [\bigcup_{i=1}^{n-1} E_i]) \qquad \text{(i)}$$

But since the lemma holds for n – 1 we have

$$m^*(A \cap [\bigcup_{i=1}^{n-1} E_i]) = \sum_{i=1}^{n-1} m^*(A \cap E_i)$$

Therefore (i) reduces to

$$m^* X = m^*(A \cap E_n) + \sum_{i=1}^{n-1} m^*(A \cap E_i)$$

$$= \sum_{i=1}^{n} m^*(A \cap E_i).$$

Hence the lemma.

Lemma 2. Let **A** be an algebra of subsets and $\{E_i \mid i \in N\}$ a sequence of sets in **A**. Then there exists a sequence $[D_i \mid i \in N]$ of disjoint members of **A** such that

$$D_i \subset E_i \qquad (i \in N)$$

$$\bigcup_{i \in N} D_i = \bigcup_{i \in N} E_i$$

Proof. For every $i \in N$, let

$$\begin{aligned} D_n &= E_n \setminus (E_1 \cup E_2 \cup \ldots\ldots \cup E_{n-1}) \\ &= E_n \cap (E_1 \cup E_2 \cup \ldots\ldots \cup E_{n-1})^c \\ &= E_n \cap E_1^c \cap E_2^c \cap \ldots\ldots \cap E_{n-1}^c \qquad \text{(i)} \end{aligned}$$

Since the complements and intersections of sets in **A** are in **A**, we have each $D_n \in$ **A**. By construction, we obviously have

$$D_i \subset E_i \quad (i \in N) \qquad \text{(ii)}$$

Let D_n and D_m be two such sets, and suppose $m < n$. Then $D_m \subset E_m$, and so

$$\begin{aligned} D_m \cap D_n &\subset E_m \cap D_n \\ &= E_m \cap E_n \cap E_1^c \cap \ldots.. E_m^c \cap \ldots. \cap E_{n-1}^c \qquad \text{(using (i))} \\ &= (E_m \cap E_m^c) \cap \ldots. \\ &= \phi \cap \ldots\ldots\ldots \\ &= \phi \qquad \text{(iii)} \end{aligned}$$

The relation (i) implies

$$\bigcup_{i \in N} D_i \subset \bigcup_{i \in N} E_i$$

It remains to prove that

$$\bigcup_{i \in N} D_i \supset \bigcup_{i \in N} E_i.$$

For this purpose let x be any member of $\bigcup_{i \in N} E_i$. Let n denotes the least natural number satisfying $x \in E_n$. Then we have

$$x \in E_n \setminus (E_1 \cup \dots \cup E_{n-1}) = D_n \subset \bigcup_{i \in N} D_n.$$

This completes the proof.

Theorem 31. The collection **M** of measurable sets is a σ - algebra.

Proof. We have proved already that **M** is an algebra of sets and so we have only to prove that **M** is closed with respect to countable union. By the lemma proved above each set E of such countable union must be the union of a sequence $\{D_n\}$ of pairwise disjoint measurable sets. Let A be any set, and let

$E_n = \bigcup_{i \in 1}^{n} D_i \subset E$. Then E_n is measurable and $E_n^c \supset E^c$. Hence

$$m^* A = m^*(A \cap E_n) + m^*(A \cap E_n^c) \geq m^*(A \cap E_n) + m^*(A \cap E_n^c).$$

But by lemma 1,

$$m^*(A \cap E_n) = m^*[A \cap (\bigcup_{i \in 1}^{n} D_i)] = \sum_{i=1}^{n} m^*(A \cap D_i)$$

Therefore

$$m^* A \geq \sum_{i=1}^{n} m^*(A \cap D_i) + m^*(A \cap E^c)$$

Since the left hand side of the inequality is independent of n, we have

$$m^* A \geq \sum_{i=1}^{\infty} m^*(A \cap D_i) + m^*(A \cap E^c)$$

$$\geq m^*(\bigcup_{i \in 1}^{\infty} [A \cap D_i]) + m^*(A \cap E^c) \text{ (by countably subadditivity of } m^*)$$

$$= m^*(A \cap \bigcup_{i \in 1}^{\infty} D_i) + m^*(A \cap E^c)$$

$$= m^*(A \cap E) + m^*(A \cap E_n^c).$$

which implies that E is measurable. Hence the theorem.

Lemma 3. The interval (a, ∞) is measurable

Proof. Let A be any set and

$$A_1 = A \cap (a, \infty)$$
$$A_2 = A \cap (a, \infty)^c = A \cap (-\infty, a].$$

Then we must show that

$$m^* A_1 + m^* A_2 \leq m^*. A.$$

If $m^* A = \infty$, then there is nothing to prove. If $m^* A$ **M** $< \infty$, then given $\in > 0$ there is a countable collection $\{I_n\}$ of open intervals which cover A and for which

$$\sum l(I_n) \leq m^* A + \in \qquad \text{(i)}$$

Let $I_n' = I_n \cap (a, \infty)$ and $I_n'' = I_n \cap (-\infty, a)$. Then I_n' and I_n'' are intervals (or empty) and

$$l(I_n) = l(I_n') + l(I_n'') = m^*(I_n') + m^*(I_n'') \qquad \text{(ii)}$$

Since $A_1 \subset U I_n'$, we have

$$m^* A_1 \leq m^*(U I_n') \leq \sum m^* I_n', \qquad \text{(iii)}$$

and since, $A_2 \subset U I_n'$, we have

$$m^* A_2 \leq m^*(U I_n'') \sum m^* I_n'', \qquad \text{(iv)}$$

Adding (iii) and (iv) we have

$$m^* A_1 + m^* A_2 \le \sum m^* I_n' + \le \sum m^* I_n''$$

$$= \sum (m^* I_n' + \le m^* I_n'')$$

$$= \sum l(I_n) \qquad [\text{by (ii)}]$$

$$\le m^* A + \in \qquad [\text{by (i)}].$$

But $\in$ was arbitrary positive number and so we must have

$$m^* A_1 + m^* A_2 \le m^* A.$$

Definition. The collection **ß** of Borel sets is the smallest σ - algebra which contains all of the open sets.

Theorem 32. Every Borel set is measurable. In particular each open set and each closed set is measurable.

Proof. We have already proved that (a, ∞) is measurable. So we have

$$(a, \infty)^c = (-\infty, a] \text{ measurable.}$$

Since $(-\infty, b) = \bigcup_{n=1}^{\infty} (-\infty, b - \frac{1}{n}]$ and we know that countable union of measurable sets is measurable, therefore $(-\infty, b)$ is also measurable. Hence each open interval,

$(a, b) = (-\infty, b) \cap (a, \infty)$ is measurable, being the intersection of two measurable sets.

But each open set is the union of countable number of open intervals and so must be measurable (The measurability of closed set follows because complement of each measurable set is measurable).

Let **M** denote the collection of measurable sets and C the collection of open sets. Then $C \subset$ M. Hence **ß** is also a subset of **M** since it is the smallest σ - algebra containing C. So each element of **ß** is measurable. Hence each Borel set is measurable.

Definition. If E is a measurable set, then the outer measure of E is called the Lebesgue Measure of E ad is denoted by mE.

Thus, m is the set function obtained by restricting the set function m* to the family **M** of measurable sets. Two important properties of Lebesgue measure are summarized by the following theorem.

Theorem 33. Let $\{E_n\}$ be a sequence of measurable sets. Then

$$m(\cup E_i) \le \sum m E_i$$

If the sets E_n are pairwise disjoint, then

$$m(\cup E_i) = \sum m E_i .$$

Proof. The inequality is simply a restatement of the subadditivity of m*.

If $\{E_i\}$is a finite sequence of disjoint measurable sets. So we apply lemma 1 replacing A by **R**. That is , we have

$$m^*(\mathbf{R} \cap [\bigcup_{i=1}^{n} E_i]) = \sum_{i=1}^{n} m^*(R \cap E_i)$$

$$\Rightarrow m^*(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} m^* E_i$$

and so m is finitely additive.

Let $\{E_i\}$ be an infinite sequence of pairwise disjoint sequence of measurable sets. Then

$$\bigcup_{i=1}^{\infty} E_i \supset \bigcup_{i=1}^{n} E_i.$$

And so

$$m(\bigcup_{i=1}^{\infty} E_i) \geq m(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} m\, E_i$$

Since the left hand side of this inequality is independent of n, we have

$$m(\bigcup_{i=1}^{\infty} E_i) \geq \sum_{i=1}^{\infty} m\, E_i$$

The reverse inequality follows from countable subadditivity and we have

$$m(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m\, E_i.$$

Hence the theorem is proved.

Theorem 34. Let $\{E_n\}$ be an infinite sequence of measurable sets such that $E_{n+1} \subset E_n$ for each n. Let $mE_1 < \infty$. Then

$$m(\bigcap_{i=1}^{\infty} E_n) = \lim_{n \to \infty} mE_n.$$

Proof. Let $E = \bigcap_{i=1}^{\infty} E_i$ and let $F_i = E_i - E_{i-1}$. Then since $\{E_n\}$ is a decreasing sequence. We have

$$\cap F_i = \phi.$$

Also we know that if A and B are measurable sets then their difference $A - B = A \cap B^c$ is also measurable. Therefore each F_i is measurable. Thus $\{F_i\}$ is a sequence of measurable pairwise disjoint sets.
Now

$$\bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} (E_i - E_{i+1})$$

$$= \bigcup_{i=1}^{\infty} (E_i \cap E_{i+1}{}^c)$$

$$= E_1 \cap (\cup E_i{}^c)$$

$$= E_1 \cap (\bigcap_{i=1}^{\infty} E_i)^c$$

$$= E_1 \cap E^c$$

$$= E_1 - E$$

Hence

$$m(\bigcup_{i=1}^{\infty} F_i) = m(E_1 - E)$$

$$\Rightarrow \sum_{i=1}^{\infty} m\, F_i = m(E_1 - E)$$

$$\Rightarrow \sum_{i=1}^{\infty} m(E_i - E_{i+1}) = m(E_1 - E) \qquad \text{(i)}$$

Since $E_1 = (E_1 - E) \cup E$, therefore

$$mE_1 = m(E_1 - E) + m(E)$$

$$\Rightarrow m E_1 - m E = m(E_1 - E), \text{ (since } m E \leq m E_1 < \infty) \qquad \text{(ii)}$$

Again

$$E_i = (E_i - E_{i+1}) \cup E_{i+1}$$

$$\Rightarrow m E_i = m(E_i - E_{i+1}) + m E_{i+1}$$

$$\Rightarrow m E_i - m E_{i+1} = m(E_i - E_{i+1}) \text{ (since } E_{i+1} \subset E_i) \qquad \text{(iii)}$$

Therefore (i) reduces to

$$m E_1 - m E = \sum_{i=1}^{\infty} (m E_i - m E_{i+1}) \text{ (using (ii) and (iii))}$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} (m E_i - m E_{i+1})$$

$$= \lim_{n\to\infty} [m E_1 - m E_2 + mE_2 - m E_3 \ldots\ldots - m E_{n+1}]$$

$$= \lim_{n\to\infty} [m E_1 - m E_{n+1}]$$

$$= m E_1 - \lim_{n\to\infty} m E_{n+1}$$

$$\Rightarrow m E = \lim_{n\to\infty} m E_n$$

$$\Rightarrow m\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} m E_n$$

Theorem 35. Let $\{E_n\}$ be an increasing sequence of measurable sets, that is, a sequence with $E_n \subset E_{n+1}$ for each n. Let mE_1 be finite, then

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} m E_n.$$

Proof. The sets $E_1, E_2 - E_1, E_3 - E_2, \ldots, E_n - E_{n-1}$, are measurable and are pairwise disjoint. Hence

$$E_1 \cup (E_2 - E_1) \cup \ldots \cup (E_n - E_{n-1}) \cup \ldots$$

is measurable and

$$m[E_1 \cup (E_2 - E_1) \cup \ldots \cup (E_n - E_{n-1}) \cup \ldots]$$

$$= m E_1 + \sum_{i=2}^{n} m(E_i - E_{i-1}) = m E_1 + \lim_{n\to\infty} \sum_{i=2}^{n} m(E_i - E_{i-1})$$

But $E_1 \cup (E_2 - E_1) \cup \ldots \cup (E_n - E_{n-1}) \cup \ldots\ldots$ is precisely $\bigcup_{i=1}^{\infty} E_n$

Moreover,

$$\sum_{i=2}^{n} m(E_i - E_{i-1}) = \sum_{i=2}^{n} (m E_i - m E_{i-1})$$

$$= (m E_2 - m E_1) + (m E_3 - m E_2) + \ldots\ldots + (m E_n - m E_{n-1})$$

$$= m E_n - m E_1$$

Thus we have

$$m\left[\bigcup_{i=1}^{\infty} E_i\right] = m E_1 + \lim_{n\to\infty} [m E_n - m E_1]$$

$$= \lim_{n\to\infty} m E_n.$$

Definition. The symmetric difference of the sets A and B is the union of the sets A – B and B – A. It is denoted by $A \Delta B$

Theorem 36. If m $(E_1 \Delta E_2) = 0$ and E_1 is measurable, then E_2 is measurable. Moreover $m E_2 = m E_1$.

Proof. We have

$$E_2 = [E_1 \cup (E_2 - E_1)] - (E_1 - E_2) \qquad (i)$$

By hypothesis, both $E_2 - E_1$ and $E_1 - E_2$ are measurable and have measure zero. Since E_1 and $E_2 - E_1$ are disjoint, $E_1 \cup (E_2 - E_1)$ is measurable and $m[E_1 \cup (E_2 - E_1)] = m E_1 + 0 = m E_1$. But, since

$$E_1 - E_2 \subset [E_1 \cup (E_2 - E_1)],$$

it follows from (i) that E_2 is measurable and

$$\begin{aligned} m E_2 &= m[E_1 \cup (E_2 - E_1)] - m(E_1 - E_2) \\ &= m E_1 - 0 = m E_1. \end{aligned}$$

This completes the proof.

Definition. Let x and y be real numbers in [0, 1]. Then **sum modulo 1 of x and y**, denoted by $x \overset{o}{+} y$, is defined by

$$x \overset{o}{+} y = \begin{cases} x + y \ if \ x + y < 1 \\ x + y - 1 \ if \ x + y \geq 1. \end{cases}$$

It can be seen that $\overset{o}{+}$ is a commutative and associative operation which takes pair of numbers in[0, 1) into numbers in [0, 1).

If we assign to each $x \in [0, 1)$ the angle $2\pi x$ then addition modulo 1 corresponds to the addition of angles.

If E is a subset of [0, 1), we define the **translation module 1** of E to be the set

$$E \overset{o}{+} y = [z \mid z = x \overset{o}{+} y \text{ for some } x \in E].$$

If we consider addition modulo 1 as addition of angles, translation module 1 by y corresponds to rotation through an angle of 2π y.

We shall now show that Lebesgue measure is invariant under translation modulo 1.

Lemma. Let $E \subset [0, 1)$ be a measurable set. Then for each $y \in [0, 1)$ the set $E \overset{o}{+} y$ is measurable and $m(E \overset{o}{+} y) = m E$.

Proof. Let $E_1 = E \cap [0, 1 - y)$ and $E_2 = E \cap [1 - y, 1)$. Then E_1 and E_2 are disjoint measurable sets whose union is E, and so

$$m E = m E_1 + m E_2.$$

We observe that

$$\begin{aligned} E_1 \overset{o}{+} y &= \{x \overset{o}{+} y : x \in E_1] \\ &= \begin{cases} x + y \ if \ x + y < 1 \\ x + y - 1 \ if \ x + y \geq 1. \end{cases} \qquad x \in E_1 \end{aligned}$$

But for $x \in E_1$, we have $x + y < 1$ and so

$$E_1 \overset{o}{+} y = \{x + y, x \in E_1\} = E_1 + y.$$

and hence $E_1 \overset{o}{+} y$ is measurable. Thus

$$m(E_1 \overset{o}{+} y) = m(E_1 + y) = m(E_1),$$

since m is translation invariant. Also $E_2 \overset{o}{+} y = E_2 + (y - 1)$ and so $E_2 \overset{o}{+} y$ is measurable and $m(E_2 \overset{o}{+} y) = m E_2$. But

$$E \overset{o}{+} y = (E_1 \overset{o}{+} y) \cup (E_2 \overset{o}{+} y)$$

and the sets $(E_1 \overset{o}{+} y)$ and $(E_2 \overset{o}{+} y)$ are disjoint measurable sets. Hence $E \overset{o}{+} y$ is measurable and

$$m(E \overset{o}{+} y) = m[(E_1 \overset{o}{+} y) \cup (E_2 \overset{o}{+} y)$$

$= m(E_1 \overset{o}{+} y) + m(E_2 \overset{o}{+} y)$
$= m(E_1) + m(E_2)$
$= m(E).$

This completes the proof of the lemma.

Construction of a non – measurable set. If x – y is a rational number, we say that x and y are equivalent and write x ~ y. It is clear that x ~ x; x ~ y $\Rightarrow$ y ~ x and x ~ y, y ~ z $\Rightarrow$ x ~ z. Thus '~' is an equivalence relation and hence partitions [0, 1) into equivalence classes, that is, classes such that any two elements of one class differ by a rational number, while any two elements of different classes differ by an irrational number. By the axiom of choice (Let *C* be any collection of non – empty sets. Then there is a function F defined on *C* which assign to each set A $\in$ *C* on element F(A) in A.) there is a set P which contains exactly one element from each equivalence class. Let $\langle r_i \rangle_{i=0}^{\infty}$ be an enumeration of the rational numbers in [0, 1) with $r_0 = 0$ and define

$P_i = P \overset{o}{+} r_i$. (translator modulo 1 of P)

Then $P_0 = P$. Let $x \in P_i \cap P_j$. Then

$$x = p_i + r_i = p_j + r_j.$$

with p_i and p_j belonging to P. But $p_i - p_j = r_j - r_i$ is a rational number, whence $p_i \sim p_j$. Since P has only one element from each equivalence class, we must have i = j. This implies that if $i \neq j$, $P_i \cap P_j = \phi$, that is , that $\langle P_i \rangle$ is a pair wise disjoint sequence of sets. On the other hand, each real number x in [0, 1) is in some equivalence class and so is equivalent to an element in P. But if x differs from an element in P by the rational number r_i, them $x \in P_i$. Thus $\cup P_i$ = [0, 1). Since each P_i is a translation modulo 1 of P, each P_i will be measurable if P is and will have the same measure. But if this were the case,

$$m[0, 1) = \sum_{i=1}^{\infty} m P_i = \sum_{i=1}^{\infty} m P,$$

and the right hand side is either zero or infinite, depending on whether m P is zero or positive. But this is impossible since m[o, 1) = 1, and consequently P cannot be measurable.

Definition. An outer measure is said to be **regular** if for any set A contained in whole space X we can find a measurable set B such that

$$B \supset A \text{ and } m^*A = m^* B = m B.$$

Theorem37. Let m* be a regular outer measure such that $m^* X < \infty$. Then the necessary and sufficient condition for a set E to be measurable is that

$$m^* X = m^* E + m^* E^c .$$

Proof. The condition is necessary :- Since E is measurable, for any set A we have

$$m^* A = m^*(A \cap E) + m^*(A \cap E^c)$$

Replacing A by X we have

$$m^* X = m^*(X \cap E) + m^*(X \cap E^c)$$
$$= m^* E + m^* E^c.$$

The condition is sufficient :- Let A be any set. Since m* is regular we can find a measurable set $B \supset A$ such that

$$m^* A = m^* B = m B.$$

Now B being measurable, we have

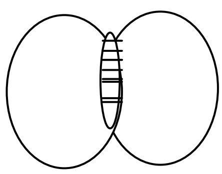
$$m^* E = m^*(E \cap B) + m^*(E \cap B^c) \qquad \text{(i)}$$
$$m^* E^c = m^*(E^c \cap B) + m^*(E^c \cap B^c) \qquad \text{(ii)}$$

Adding (i) & (ii) we have

$$m^* E + m^* E^c = m^*(E \cap B) + m^*(E \cap B^c) + m^*(E^c \cap B) + m^*(E^c \cap B^c)$$
$$\Rightarrow m^* X = m^*(E \cap B) + m^*(E \cap B^c) + m^*(E^c \cap B) + m^*(E^c \cap B^c)$$

Now consider

E B



$$m^*(E \cap B) + m^*(E^c \cap B).$$

Since

$$B = (E \cap B) \cup (E^c \cap B),$$

Therefore

$$m^* B \leq m^*(E \cap B) + m^*(E^c \cap B)$$

Hence EAB

$$m^* X \geq m^*(E \cap B^c) + m^* B + m^*(E^c \cap B^c)$$

Again

$$m^* B^c \leq m^*(E \cap B^c) + m^*(E^c \cap B^c) \qquad \text{(iii)}$$

Hence

$$m^* X \geq m^* B + m^* B^c$$
$$= m B + m B^c = m^* X$$

Thus

$$m^*(E \cap B) + m^*(E \cap B^c) + m^*(E^c \cap B) + m^*(E^c \cap B^c)$$
$$= m B + m B^c$$

That is,

$$m^*(E \cap B) + m^*(E \cap B^c) + m^*(E^c \cap B) + m^*(E^c \cap B^c) - m B^c$$
$$= m B$$

Using (iii), the last expression reduces to

$$m^*(E \cap B) + m^*(E^c \cap B) \leq m B.$$

INNER MEASURE

Definition. Let F be a closed set. Then inner measure of a set E, denoted by $m_* E$, is defined by

$$m_* E = \sup_{F \subset E} \{| F |\}$$

where | . | denotes the length.

Definition. A subset E of (a, b) is said to measurable if

$$m^* E = m_* E.$$

The relation between inner and outer measure is $m_* E = (b - a) - m^* E^c$.

Let $[E_n]$ be a sequence of sets. Then

Definition. The set of those elements which belong to E_n for infinitely many values of n is called the lim sup of the sequence of sets $\{E_n\}$. We denote it by $\overline{\lim}\ E_n$.

Lemma. $\overline{\lim}\ E_n = \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n.$

Proof. Let $x \in \overline{\lim}\ E_n$. Then

$$x \in \bigcup_{n \geq m} E_n \qquad \text{for all } m$$

$$\Rightarrow x \in \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n \Rightarrow \overline{\lim}\ E_n \subset \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n$$

conversely, if $x \in \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n$, then

$$x \in \bigcup_{n \geq m} E_n \qquad \text{for all } m$$

$\Rightarrow x \in E_n$ for some $n \geq m$

$\Rightarrow x \in E_n$ for infinite values of n.

$\Rightarrow x \in \overline{\lim}\ E_n$

$$\Rightarrow \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n \subset \overline{\lim}\ E_n$$

Hence $\overline{\lim}\, E_n = \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} E_n$.

Definition. The set of those elements which belong to all E_n except a finite number of them is called lim in f of the sequence of sets $\{E_n\}$. We denote it by $\underline{\lim}\, E_n$.

It may be proved that $\underline{\lim}\, E_n = \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} E_n$. For, let

$$x \in \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} E_n \Rightarrow x \in \bigcap_{m=1}^{\infty} E_n \text{ for some m}$$

$$\Rightarrow x \in \underline{\lim}\, E_n.$$

Similarly, we can show that if $x \in \lim E_n$ then $x \in \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} E_n$.

Theorem 38. If $\{E_n\}$ is a sequence of measurable sets, then $\overline{\lim}\, E_n$ and $\underline{\lim}\, E_n$ are also measurable.

Proof. We have shown above that

$$\overline{\lim}\, E_n = \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} E_n.$$

Since $\{E_n\}$ is a sequence of measurable sets, the right hand side is measurable and so $\overline{\lim}\, E_n$ is measurable.
Similarly, since

$$\underline{\lim}\, E_n = \bigcup_{m=1}^{\infty} \bigcap_{n \geq m} E_n$$

it is obvious that right hand side and so $\underline{\lim}\, E_n$ is measurable.

Example. The cantor set is uncountable with outer measure zero.

Solution. We already know that cantor ser is uncountable. Let C_n denote the union of the closed intervals left at the nth stage of the construction. We note that C_n consists of 2^n closed intervals, each of length 3^{-n}Therefore

$$m^* C_n \leq 2^n . 3^{-n} \qquad (\because m^* C_n = m^*(\cup F_n) = \sum m^* F_n)$$

But any point of the cantor set C must be in one of the intervals comprising the union C_n, for each $n \in N$, and as such $C \subset C_n$ for all $n \in N$. Hence

$$m^* C \leq m^* C_n \leq \left(\frac{2}{3}\right)^n$$

This being true for each $n \in N$, letting $n \to \infty$ gives $m^* C = 0$.

Example. If $E_1$and E_2 are any measurable sets, show that

$$M(E_1 \cup E_2) + m(E_1 \cap E_2) = m(E_1) + m(E_2).$$

Proof. Let A be any set. Since E_1 is measurable,

$$m^* A = m^*(A \cap E_1) + m^*(A \cap E_1^c).$$

We set $A = E_1 \cup E_2$ and we have

$$m^*(E_1 \cup E_2) = m^*[(E_1 \cup E_2) \cap E_1] + m^*[(E_1 \cup E_2) \cap E_1^c]$$

Adding $m (E_1 \cap E_2)$ to both sides we have

$$m(E_1 \cup E_2) + m(E_1 \cap E_2) = m E_1 + m[(E_1 \cup E_2) \cap E_1^c] + m(E_1 \cap E_2) \quad (1)$$

But

$$E_2 = [(E_1 \cup E_2) \cap E_1^c] \cup (E_1 \cap E_2).$$

Therefore

$$m\{[(E_1 \cup E_2) \cap E_1^c] \cup (E_1 \cap E_2)\} = m E_2$$

Hence (1) reduces to

$$M(E_1 \cup E_2) + m(E_1 \cap E_2) = mE_1 + m\,E_2.$$

Theorem39. Let E be any set. Then given $\in > 0$, there is an open set $O \supset E$ such that

$$m^*\,O < m^*\,E + \in.$$

Proof. There exists a countable collection $[I_n]$ of open intervals such that $E \subset \bigcup_n I_n$ and

$$\sum_{n=1}^{\infty} l(I_n) < m^*\,E + \in.$$

Put $\quad O = \bigcup_{n=1}^{\infty} I_n.$

Then O is an open set and

$$\begin{aligned} m^*\,O &= m^*\Big(\bigcup_{n=1}^{\infty} I_n\Big) \\ &\le \sum_{n=1}^{\infty} m^*\,I_n \\ &= \sum_{n=1}^{\infty} l(I_n) < m^*\,E + \in. \end{aligned}$$

Theorem 40. Let E be a measurable set. Given $\in > 0$, there is an open set $O \supset E$ such that $m^*(O\backslash E) < \in$.

Proof. Suppose first that $m\,E < \infty$. Then by the above theorem there is an open set $O \supset E$ such that

$$m^*\,O < m^*\,E + \in$$

Since the sets O and E are measurable, we have

$$m^*(O\backslash E) = m^*\,O - m^*\,E < \in.$$

Consider now the case when $m\,E = \infty$. Write the set **R** of real numbers as a union of disjoint finite intervals; that is, $\mathbf{R} = \bigcup_{n=1}^{\infty} I_n$. Then, if $E_n = E \cap I_n$, $m(E_n) < \infty$. We can, thus, find open sets $O_n \supset E_n$ such that

$$m^*(O_n - E_n) < \frac{\in}{2^n}.$$

Define $O = \bigcup_{n=1}^{\infty} O_n$. Clearly O is an open set such that $O \supset E$ and satisfies

$$O - E = \bigcup_{n=1}^{\infty} O_n - \bigcup_{n=1}^{\infty} E_n \subset \bigcup_{n=1}^{\infty} (O_n - E_n)$$

Hence

$$m^*(O - E) \le \sum_{n=1}^{\infty} m^*(O_n\backslash E_n) < \in.$$

4

MEASURABLE FUNCTIONS AND LEBESGUE INTEGRAL

PART A : MEASURABLE FUNCTIONS

4.1. Definition. Let E be a measurable set and f a function defined on E. Then f is said to be measurable (Lebesgue function) if for any real α any **one** of the following four conditions is satisfied.

(a) $\{x \mid f(x) > \alpha\}$ is measurable
(b) $\{x \mid f(x) \geq \alpha\}$ is measurable
(c) $\{x \mid f(x) < \alpha\}$ is measurable
(d) $\{x \mid f(x) \leq \alpha\}$ is measurable.

We show first that these four conditions are equivalent. First of all we show that (a) and (b) are equivalent. Since

$$\{x \mid f(x) > \alpha\} = \{x \mid f(x) \leq \alpha\}^c$$

and also we know that complement of a measurable set is measurable, therefore (a) $\Rightarrow$ (d) and conversely.

Similarly since (b) and (c) are complement of each other, (c) is measurable if (b) is measurable and conversely.

Therefore, it is sufficient to prove that (a) $\Rightarrow$ (b) and conversely.

Firstly we show that (b) $\Rightarrow$ (a) .
The set $\{x \mid f(x) \geq \alpha\}$ is given to be measurable.
Now

$$\{x \mid f(x) > \alpha\} = \bigcup_{n=1}^{\infty} \{x \mid f(x) \geq \alpha + \frac{1}{n}\}$$

But by (b), $\{x \mid f(x) \geq \alpha + \frac{1}{n}\}$ is measurable. Also we know that countable union of measurable sets is measurable. Hence $\{x \mid f(x) > \alpha\}$ is measurable which implies that (b) $\Rightarrow$ (a). Conversely, let (a) holds. We have

$$\{x \mid f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \{x \mid f(x) > \alpha - \frac{1}{n}\}$$

The set $\{x \mid f(x) > \alpha - \frac{1}{n}\}$ is measurable by (a). Moreover, intersection of measurable sets is also measurable. Hence $\{x \mid f(x) \geq \alpha\}$ is also measurable. Thus (a) $\Rightarrow$ (b).
Hence the four conditions are equivalent.

Lemma. If α is an extended real number then these four conditions imply that $\{x \mid f(x) = \alpha\}$ is also measurable.

Proof. Let α be a real number, then

$$\{x \mid f(x) = \alpha\} = \{x \mid f(x) \geq \alpha\} \cap \{x \mid f(x) \leq \alpha\}.$$

Since $\{x \mid f(x) \geq \alpha\}$ and $\{x \mid f(x) \leq \alpha\}$ are measurable by conditions (b) and (d), the set $\{x \mid f(x) = \alpha\}$ is measurable being the intersection of measurable sets.

Suppose $\alpha = +\infty$. Then

$$\{x \mid f(x) = \infty\} = \bigcap_{n=1}^{\infty} \{x \mid f(x) \geq n\}$$

which is measurable by the condition (b) and the fact that intersection of measurable sets is measurable.

Similarity when $\alpha = -\infty$, then

$$\{x \mid f(x) = -\infty\} = \bigcap_{n=1}^{\infty}\{x \mid f(x) \le -n\}$$

which is again measurable by condition (d).

Hence the result follows.

Second definition of Measurable functions

We see that

$$\{x \mid f(x) > \alpha\}$$

is inverse image of $(\alpha, \infty]$. Similarly the sets

$[x \mid f(x) \ge \alpha\}$, $\{x \mid f(x) < \alpha\}$, $\{x \mid f(x)\ \varepsilon\ \alpha\}$ are inverse images of $[\alpha, \infty]$, $[-\infty, \alpha)$ and $[-\infty, \alpha]$ respectively.

Hence we can also define a measurable function as follows.

A function f defined on a measurable set E is said to be measurable if for any real α any one of the four conditions is satisfied :

(a) The inverse image $f^{-1}(\alpha, \infty]$ of the half-open interval $(\alpha, \infty]$ is measurable.

(b) For every real α, the inverse image $f^{-1}[\alpha, \infty]$ of the closed interval $[\alpha, \infty]$ is measurable.

(c) The inverse image $f^{-1}[-\infty, \alpha)$ of the half open interval $[-\infty, \alpha)$ is measurable.

(d) The inverse image $f^{-1}[-\infty, \alpha]$ of the closed interval $[-\infty, \alpha]$ is measurable.

Remark 1. It is immediate that a necessary and sufficient condition for measurability is that $\{x \mid a \le f(x) \le b\}$ should be measurable for all a, b [including the case $a = -\infty$, $b = +\infty$], for any set of this form can be written as the intersection of two sets

$$\{x \mid f(x) \ge a\} \cap \{x \mid f(x) \le b\},$$

if f is measurable, each of these is measurable and so is $\{x \mid a \le f(x) \le b\}$. Conversely any set of the form occurring in the definition can easily be expressed in terms of the sets of the form $\{x \mid a \le f(x) \le b\}$.

Remark 2. Since (α, ∞) is an open set, we may define a measurable function as "A function f defined on a measurable set E is said to be measurable if for every open set G in the real number system, $f^{-1}(G)$ is a measurable set.

Definition. Characteristic function of a set E is defined by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x\ \varepsilon\ E \\ 0 & \text{if } x \notin E \end{cases}$$

This is also known as **indicator** function.

Example of a Measurable function

Let E be a set of rationals in [0, 1]. Then the characteristic function $\chi_E(x)$ is measurable.

Proof. For the set of rationals in the given interval, we have

$$\chi_E(x) = \begin{cases} 1 & \text{if } x\ \varepsilon\ E \\ 0 & \text{if } x \notin E \end{cases}$$

It is sufficient to prove that the set

$$\{x \mid \chi_E(x) > \alpha\}$$

is measurable for any real α.

Let us suppose first that $\alpha \ge 1$. Then

$$\{x \mid \chi_E(x) > \alpha\} = \{x \mid \chi_E(x) > 1\}$$

Hence the set $\{x \mid \chi_E(x) > \alpha\}$ is empty in this very case. But outer measure of any empty set is zero. Hence for $\alpha \ge 1$, the set $\{x \mid \chi_E(x) > \alpha\}$ and so $\chi_E(x)$ is measurable.

Further let $0 \le \alpha < 1$. Then

$$\{x \mid \chi_E(x) > \alpha\} = E$$

But E is countable and therefore measurable. Hence $\chi_E(x)$ is measurable.

Lastly, let $\alpha < 0$. Then

$$\{x \mid \chi_E(x) > \alpha\} = [0, 1]$$

and therefore measurable. Hence the result.

Example 2. A continuous is measurable.

Proof. If the function f is continuous, then $f^{-1}(\alpha, \infty)$ is also open. But every open set is measurable. Hence every continuous function is measurable.

We may also argue as follows:

If f is continuous then

$$\{x \mid f(x) < \alpha, x \varepsilon (a, b)\}$$

is closed and hence

$$\{x \mid f(x) \geq \alpha\} = \{x \mid f(x) < \alpha\}^c$$

is open and so measurable.

* All the ordinary functions of analysis may be obtained by limiting process from continuous function and so are measurable.

Example 3. A constant function with a measurable domain is measurable.

Solution. Let E be a measurable set and let $f : E \to R^*$ be a constant function definition by $f(x) = K$(constant). Then for any real α, we have

$$\{x : f(x) > \alpha\} = \begin{cases} E & \text{if } \alpha < k \\ \phi & \text{if } \alpha \geq k \end{cases}$$

Since both E and ϕ are measurable, it follows that the set $\{x : f(x) > \alpha\}$ and hence f is measurable.

Theorem 1. For any real c and two measurable real-valued functions f, g the four functions f + c, cf, f+g, fg are measurable.

Proof. We are given that f is a measurable function and c is any real number. Then for any real number α

$$\{x \mid f(x) + c > \alpha\} = \{x \mid f(x) > \alpha - c\}$$

But $\{x \mid f(x) > \alpha - c\}$ is measurable by the condition (a) of the definition. Hence $\{x \mid f(x) + c > \alpha\}$ and so $f(x) + c$ is measurable.

We next consider the function cf. In case c = 0, cf is the constant function 0 and hence is measurable since every constant function is continuous and so measurable. In case c > 0 we have

$$\{x \mid cf(x) > \alpha\} = \{x \mid f(x) > \frac{\alpha}{c}\} = f^{-1}\left(\frac{\alpha}{c}, \infty\right],$$

and so measurable.

In case c < 0, we have

$$\{x \mid cf(x) > r\} = \{x \mid f(x) < \frac{r}{c}\}$$

and so measurable.

Now if f and g are two measurable real valued functions defined on the same domain, we shall show that f+g is measurable. To show iit, it is sufficient to show that the set $\{x \mid f(x) + g(x) > \alpha\}$ is measurable.

If $f(x) + g(x) > \alpha$, then $f(x) > \alpha - g(x)$ and by the Cor. of the axiom of Archimedes there is a rational number r such that

$$\alpha - g(x) < r < f(x)$$

Since the functions f and g are measurable, the sets

$$\{x \mid f(x) > r\} \text{ and } \{x \mid g(x) > \alpha - r\}$$

are measurable. Therefore, there intersection

$$S_r = \{x \mid f(x) > r\} \cap \{x \mid g(x) > \alpha - r\}$$

is also measurable.

It can be shown that

$$\{x \mid f(x) + g(x) > \alpha\} = U\{S_r \mid r \text{ is a rational}\}$$

Since the set of rational is countable and countable union of measurable sets is measurable, the set $U\{S_r \mid r$ is a rational$\}$ and hence $\{x \mid f(x) + g(x) > \alpha\}$ is measurable which proves that $f(x) + g(x)$ is measurable.

From this part it follows that $f-g = f+(-g)$ is also measurable, since when g is measurable $(-g)$ is also measurable.

Next we consider fg.

The measurability of fg follows from the identity

$$fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2] ,$$

if we prove that f^2 is measurable when f is measurable. For this it is sufficient to prove that

$$\{ x \in E \mid f^2(x) > \alpha \} , \quad \alpha \text{ is a real number,}$$

is measurable.

Let α be a negative real number. Then it is clear that the set $\{x \mid f^2(x) > \alpha \}$ = E(domain of the measurable function f). But E is measurable by the definition of f . Hence $\{x \mid f^2(x) > \alpha \}$ is measurable when $\alpha < 0$.

Now let $\alpha \geq 0$, then

$$\{x \mid f^2(x) > \alpha \} = \{x \mid f(x) > \sqrt{\alpha}\} \cup \{x \mid f(x) < -\sqrt{\alpha}\}$$

Since f is measurable, it follows from this equality that

$$\{x \mid f^2(x) > \alpha \}$$

is measurable for $\alpha \geq 0$.

Hence f^2 is also measurable when f is measurable.

Therefore, the theorem follows from the above identity, since measurability of f and g imply the measurability of f + g . From this we may also conclude that f/g ($g \neq 0$) is also measurable.

Theorem 2. If f is measurable, then |f| is also measurable.

Proof. It suffices to prove the measurability of the set

$$\{x \mid |f(x)| > \alpha \}, \quad \text{where } \alpha \text{ is any real number.}$$

If $\alpha < 0$, then

$$\{x \mid |f(x)| > \alpha\} = E \text{ (domain of f)}$$

But E is assumed to be measurable. Hence $\{x \mid |f(x)| > \alpha \}$ is measurable for $\alpha < 0$.
If $\alpha \geq 0$ then

$$\{x \mid |f(x)| > \alpha \} = \{x \mid f(x) > \alpha \} \cup \{x \mid f(x) < -\alpha \}$$

The right hand side of the equality is measurable since f is measurable. Hence $\{x \mid f(x)| > \alpha\}$ is also measurable.

Hence the theorem is proved.

Theorem 3. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions. Then
$\sup\{f_1, f_2,.., f_n\}$, $\inf. \{f_1, f_2,\ldots, f_n \}$,
$\sup\limits_n f_n$, $\inf\limits_n f_n$, $\lim f_n$ and $\lim f_n$ are measurable.

Proof. Define a function

$$M(x) = \sup\{f_1, f_2,\ldots, f_n\}$$

We shall show that $\{x \mid M(x) > \alpha\}$ is measurable.

In fact

$$\{x \mid M(x) > \alpha \} = \bigcup_{i=1}^{n}\{ x|f_i(x) > \alpha \}$$

Since each f_i is measurable, each of the set $\{x \mid f_i(x) > \alpha \}$ is measurable and therefore their union is also measurable. Hence $\{x \mid M(x) > \alpha \}$ and so M(x) is measurable.

Similarly we define the function

$$m(x) = \inf\{f_1, f_2,\ldots, f_n\}$$

Since $m(x) < \alpha$ iff $f_i(x) < \alpha$ for some i we have

$$\{x \mid m(x) < \alpha \} = \bigcup_{i=1}^{n}\{ x|f_i(x) < \alpha \}$$

and since $\{x \mid f_i(x) < \alpha\}$ is measurable on account of the measurability of f_i, it follows that $\{x \mid m(x) < \alpha\}$ and so m(x) is measurable.

Define a function

$M'(x) = \sup_n f_n(x) = \sup \{f_1, f_2, \ldots, f_n, \ldots)$

We shall show that the set

$$\{x \mid M'(x) > \alpha\}$$

is measurable for any real α.

Now

$$\{x \mid M'(x) > \alpha\} = \bigcup_{i=1}^{\infty} \{x \mid f_n(x) > \alpha\}$$

is measurable, since each f_n is measurable.

Similarly if we define

$$m'(x) = \inf_n f_n(x),$$

then

$$\{x \mid m'(x) < \alpha\} = \bigcup_{i=1}^{\infty} \{x \mid f_n(x) < \alpha\}$$

and therefore measurability of f_n implies that of $m'(x)$. Now since

$$\overline{\lim}\, f_n = \lim \sup_n f_n = \inf_k \{\sup_{n \geq k} f_n\}$$

and $$\underline{\lim}\, f_n = \sup_k \{\inf_{n \geq k} f_n\},$$

the upper and lower limits are measurable.

* Finally if the sequence is convergent, its limit is the common value of lim f_n and lim f_n and hence is measurable.

Definition. Let f and g be measurable functions. Then we define

$$f^+ = \text{Max}(f, 0)$$

$$f^- = \text{Max}(-f, 0)$$

$$f \vee g = \frac{f + g + |f - g|}{2} \quad \text{i.e. Max}(f, g)$$

and

$$f \wedge g = \frac{f + g - |f - g|}{2} \quad \text{i.e. min}(f, g)$$

Theorem 4. Let f be a measurable function. Then $\overset{+}{f}$ and $\overset{-}{f}$ are both measurable.

Proof. Let us suppose that $f > 0$. Then we have

$$\overset{+}{f} = f \text{ and } \overset{-}{f} = 0 \qquad \text{(i)}$$

So in this case we have

$$f = \overset{+}{f} - \overset{-}{f}$$

Now let us take f to be negative. Then

$$\left.\begin{aligned} \overset{+}{f} &= \text{Max}(f, 0) = 0 \\ \overset{-}{f} &= \text{Max}(-f, 0) = -f \end{aligned}\right\} \qquad \text{(ii)}$$

Therefore on subtraction

$$f = \overset{+}{f} - \overset{-}{f}$$

In case $f = 0$, then

$$\overset{+}{f} = 0, \quad \overset{-}{f} = 0 \qquad \text{(iii)}$$

Therefore $f = \overset{+}{f} - \overset{-}{f}$

Thus for all f we have

$$f = \overset{+}{f} - \overset{-}{f} \qquad \text{(iv)}$$

Also adding the components of (i) we have

$$f = |f| = \overset{+}{f} + \overset{-}{f} \qquad \text{(v)}$$

since f is positive.
And from (ii) when f is negative we have

$$\overset{+}{f} + \overset{-}{f} = 0 - f = -f = |f| \qquad \text{(vi)}$$

In case f is zero, then

$$\overset{+}{f} + \overset{-}{f} = 0 + 0 = 0 = |f| \qquad \text{(vii)}$$

That is for all f, we have

$$|f| = \overset{+}{f} + \overset{-}{f} \qquad \text{(viii)}$$

Adding (iv) and (viii) we have

$$f + |f| = 2\overset{+}{f}$$

$$\Rightarrow \overset{+}{f} = \frac{1}{2}(f + |f|) \qquad \text{(ix)}$$

Similarly on subtracting we obtain

$$\overset{-}{f} = \frac{1}{2}(|f| - f) \qquad \text{(x)}$$

Since measurability of f implies the measurability of | f | it is obvious from (ix) and (x) that $\overset{+}{f}$ and f^- are measurable.

Theorem 5. If f and g are two measurable functions, then $f \vee g$ and $f \wedge g$ are measurable.

Proof. We know that

$$f \vee g = \frac{f + g\,|f - g|}{2}$$

$$f \wedge g = \frac{f + g - |f - g|}{2}$$

Now measurability of f $\Rightarrow$ measurability of |f|. Also if f and g are measurable, then f+g, f−g are measurable. Hence $f \vee g$ and $f \wedge g$ are measurable.

We now introduce the terminology "almost everywhere" which will be frequently used in the sequel.

Definition. A statement is said to hold almost everywhere in E if and only if it holds everywhere in E except possibly at a subset D of measure zero.

Examples

(a) Two functions f and g defined on E are said to be equal almost everywhere in E iff f(x) = g(x) everywhere except a subset D of E of measure zero.

(b) A function defined on E is said to be continuous almost everywhere in E if and only if there exists a subset D of E of measure zero such that f is continuous at every point of E−D.

Theorem 6. (a) If f is a measurable function on the set E and $E_1 \subset E$ is measured set, then f is a measurable function on E_1.
(b) If f is a measurable function on each of the sets in a countable collection $\{E_i\}$ of disjoint measurable sets, then f is measurable.

Proof. (a) For any real α, we have $\{x \in E_1, f(x) > \alpha\} = \{x \in E; f(x) > \alpha\} \cap E_1$. The result follows as the set on the right-hand side is measurable.

(b) Write $E = \bigcup_{i=1}^{\infty} E_i$. Clearly, E, being the union of measurable set is measurable. The result now follows, since for each real α, we have $E = \{ x \,\varepsilon\, E, f(x) > \alpha\} = \bigcup_{i=1}^{\infty} E_i \ f(x) > \alpha \}$

Theorem 7. Let f and g be any two functions which are equal almost everywhere in E. If f is measurable so is g.

Proof. Since f is measurable, for any real α the set $\{x \mid f(x) > \alpha \}$ is measurable. We shall show that the set $\{x \mid g(x) > \alpha \}$ is measurable. To do so we put

$$E_1 = \{x \mid f(x) > \alpha \}$$

and

$$E_2 = \{x \mid g(x) > \alpha \}$$

Consider the sets

$$E_1 - E_2 \text{ and } E_2 - E_1$$

Since f = g almost everywhere, measures of these sets are zero. That is, both of these sets are measurable. Now

$$E_2 = [E_1 \cup (E_2-E_1)] - (E_1-E_2)$$
$$= [E_1 \cup (E_2-E_1) \cap (E_1-E_2)^c$$

Since E_1, E_2-E_1 and $(E_1-E_2)^c$ are measurable therefore it follows that E_2 is measurable. Hence the theorem is proved.

Cor. Let $\{f_n\}$ be a sequence of measurable functions such that $\lim_{n\to\infty} f_n = f$ almost everywhere. Then f is a measurable function.

Proof. We have already proved that if $\{f_n\}$ is a sequence of measurable functions then $\lim_{n\to\infty} f_n$ is measurable. Also it is given that $\lim_{n\to\infty} f_n = f$ a.e. Therefore using the above theorem it follows that f is measurable.

Theorem 8. Characteristic function χ_A is measurable if and only if A is measurable.

Proof. Let A be measurable. Then

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \,\varepsilon\, A \\ 0 & \text{if } x \not\varepsilon\, A \text{ i.e. } x \,\varepsilon\, A^c \end{cases}$$

Hence it is clear from the definition that domain of χ_A is $A \cup A^c$ which is measurable due to the measurability of A. Therefore, if we prove that the set $\{x \mid \chi_A(x) > \alpha \}$ is measurable for any real α, we are through.

Let $\alpha \geq 0$. Then

$$\{x \mid \chi_A(x) > \alpha \} = \{x \mid \chi_A(x) = 1\}$$
$$= A \text{(by the definition of Ch. fn.)}$$

But A is given to be measurable. Hence for $\alpha \geq 0$. The set $\{x \mid \chi_A(x) > \alpha \}$ is measurable.

Now let us take $\alpha < 0$. Then

$$\{x \mid \chi_A(x) > \alpha \} = A \cup A^c$$

Hence $\{x \mid \chi_A(x) > \alpha\}$ is measurable for $\alpha < 0$ also, since $A \cup A^c$ has been proved to be measurable. Hence if A is measurable, then χ_A is also measurable.

Conversely, let us suppose that $\chi_A(x)$ is measurable. That is, the set $\{x \mid \chi_A(x) > \alpha\}$ is measurable for any real α.

Let $\alpha \geq 0$. Then

$$\{x \mid \chi_A(x) > \alpha\} = \{x \mid \chi_A(x) = 1\} = A$$

Therefore, measurability of $\{x \mid \chi_A(x) > \alpha\}$ implies that of the set A for $\alpha \geq 0$.

Now consider $\alpha < 0$. Then

$$\{x \mid \chi_A(x) > \alpha\} = A \cup A^c$$

Thus measurability of $\chi_A(x)$ implies measurability of the set $A \cup A^c$ which imply A is measurable.

Remark. With the help of above result, the existence of non-measurable function can be demonstrated. In fact, if A is non-measurable set then χ_A cannot be measurable.

Theorem 9. If a function f is continuous almost everywhere in E, then f is measurable.

Proof. Since f is continuous almost everywhere in E, there exists a subset D of E with m*D = 0 such that f is continuous at every point of the set C = E−D. To prove that f is measurable, let α denote any given real number. It suffices to prove that the inverse image

$$B = f^{-1}(\alpha, \infty) = \{x \in E \mid f(x) > \alpha\}$$

of the interval (α, ∞) is measurable.

For this purpose, let x denote an arbitrary point in $B \cap C$. Then $f(x) > \alpha$ and f is continuous at x. Hence there exists an open interval U_x containing x such that $f(y) > \alpha$ hold for every point y of $E \cap U_x$. Let

$$U = \bigcup_{x \in B \cap C} U_x$$

Since $x \in E \cap U_x \subset B$ holds for every $x \in B \cap C$, we have

$$B \cap C \subset E \cap \cup \subset B$$

This implies

$$B = (E \cap U) \cup (B \cap D)$$

As an open subset of R, U is measurable. Hence $E \cup U$ is measurable. On the other hand, since

$$m^*(B \cap D) \le m^*D = 0,$$

$B \cap D$ is also measurable. This implies that B is measurable. This completes the proof of the theorem.

Definition. A function ϕ, defined on a measurable set E, is called **simple** if there is a finite disjoint class $\{E_1, E_2, \dots, E_n\}$ of measurable sets and a finite set $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ of real numbers such that

$$f(x) = \begin{cases} \alpha_i & \text{if } x \in E_i, i = 1,2,\dots,n \\ 0 & \text{if } x \notin E_1 \cup E_2 \cup \dots \cup E_n \end{cases}$$

Thus, a function is simple if it is measurable and takes only a finite number of different values.
The simplest example of a simple function is the characteristic function χ_E of a measurable set E.

Definition. A function f is said to be a step function if

$$f(x) = C_i, \quad \xi_{i-1} < x < \xi_i$$

for some subdivision of [a, b] and some constants C_i.

Clearly, a step function is a simple function.

Theorem 10. Every simple function ϕ on E is a linear combination of characteristic functions of measurable subsets of E.

Proof. Let ϕ be a simple function and $c_1, c_2, \dots, c_n$ denote the non-zero real numbers in its image $\phi(E)$. For each i = 1,2,…, n, let

$$A_i = \{x \in E : \phi(x) = C_i\}$$

Then we have

$$\phi = \sum_{i=1}^{n} C_i \chi_{A_i}$$

On the other hand, if $\phi(E)$ contains no non-zero real number, then $\phi = 0$ and is the characteristic function χ_ϕ of the empty subset of E.

It follows from Theorem 10 that simple functions, being the sum of measurable functions, is measurable.

Also, by the definition, if f and g are simple functions and c is a constant, then f +c, cf, f+g and fg are simple.

Theorem 10 (Approximation Theorem). For every non-negative measurable function f, there exists a non-negative non-decreasing sequence $\{f_n\}$ of simple functions such that

$$\lim_{n \to \infty} f_n(x) = f(x), \quad x \in E$$

In the general case if we do not assume non-negativeness of f , then we say

For every measurable function f, there exists a sequence $\{f_n\}$, $n \in N$ of simple function which converges (pointwise) to f .
i.e. "Every measurable function can be approximated by a sequence of simple functions."

Proof. Let us assume that $f(x) \ge 0$ and $x \in E$. Construct a sequence

$$f_n(x) = \begin{cases} \dfrac{i-1}{2^n} & \text{for} \dfrac{i-1}{2^n} \le f(x) < \dfrac{i}{2^n}, i = 1,2,..., n\,2^n \\ n & \text{for } f(x) \ge n \end{cases}$$

for every n ε N.
If we take n = 1, then

$$f_1(x) = \begin{cases} \dfrac{i-1}{2} & \text{for} \dfrac{i-1}{2} \le f(x) < \dfrac{i}{2}, i = 1,2, \\ 1 & \text{for } f(x) \ge 1 \end{cases}$$

That is, $$f_1(x) = \begin{cases} 0 & \text{for } 0 \le f(x) < \dfrac{1}{2} \\ \dfrac{1}{2} & \text{for} \dfrac{1}{2} \le f(x) < 1 \\ 1 & \text{for } f(x) \ge 1 \end{cases}$$

Similarly taking n = 2, we obtain

$$f_2(x) = \begin{cases} \dfrac{i-1}{4} \text{for} \dfrac{i-1}{4} \le f(x) < \dfrac{i}{4}, i = 1,2,...,8 \\ 2 \qquad \text{for } f(x) \ge 2 \end{cases}$$

That is,

$$f_2(x) = \begin{cases} 0 & \text{for } 0 \le f(x) < \dfrac{1}{4} \\ \dfrac{1}{4} & \text{for } \dfrac{1}{4} \le f(x) < \dfrac{1}{2} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{7}{4} & \text{for} \dfrac{7}{4} \le f(x) < 2 \\ 2 & \text{for } f(x) \ge 2 \end{cases}$$

Similarly we can write $f_3(x)$ and so on. Clearly all f_n are positive whenever f is positive and also it is clear that $f_n \le f_{n+1}$. Moreover f_n takes only a finite number of values. Therefore $\{f_n\}$ is a sequence of non-negative, nondecreasing functions which assume only a finite number of values.

Let us denote

$$E_{ni} = f^{-1}\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right] = \left\{x \varepsilon E \mid \frac{i-1}{2^n} \le f(x) < \frac{i}{2^n}\right\}$$

and

$$E_n = f^{-1}[n, \infty] = \{x \varepsilon E \mid f(x) \ge n\}$$

Both of them are measurable. Let

$$f_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n_i}} + n\chi_{E_n}$$

for every n ε N .

Now $\sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n_i}}$ is measurable, since E_{n_i} has been shown to be measurable and characteristic function of a measurable set is measurable. Similarly χ_{E_n} is also measurable since E_n is measurable. Hence each f_n is measurable. Now we prove the convergence of this sequence.

Let $f(x) < \infty$. That is f is bounded. Then for some n we have

$$\frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n}$$

$$\Rightarrow \quad \frac{i-1}{2^n} - \frac{i-1}{2^n} \leq f(x) - \frac{i-1}{2^n} < \frac{i}{2^n} - \frac{i-1}{2^n},$$

$$\Rightarrow \quad 0 \leq f(x) - \frac{i-1}{2^n} < \frac{1}{2^n}$$

$$\Rightarrow \quad 0 \leq f(x) - f_n(x) < \frac{1}{2^n} \quad \text{(by the def of } f_n(x))$$

$$\Rightarrow \quad f(x) - f_n(x) < \in$$

$$\Rightarrow \quad \lim_{n\to\infty} f_n(x) = f(x)$$

and this convergence is uniform.

Let us suppose now that f is not bounded. Then since

$$f_n(x) = n \text{ for } f(x) \geq n$$

$$\lim_{n\to\infty} f_n(x) = \infty = f(x)$$

When **we do not assume non-negativenss of the function** then since we know that f^+ and f^- are both non-negative, we have by what we have proved above

$$\overset{+}{f} = \lim_{n\to\infty} \phi'_n(x) \ldots \qquad \text{(i)}$$

$$\overline{f} = \lim_{n\to\infty} \phi''_n(x) \ldots \qquad \text{(ii)}$$

where $\phi'_n(x)$ and $\phi''_n(x)$ are simple functions. Also we have proved already that

$$f = \overset{+}{f} - \overline{f}$$

Now from (i) and (ii) we have

$$\overset{+}{f} - \overline{f} = \lim_{n\to\infty} \phi'_n(x) - \lim_{n\to\infty} \phi''_n(x)$$

$$= \lim_{n\to\infty} (\phi'_n(x) - \phi''_n(x))$$

$$= \lim_{n\to\infty} \phi_n(x)$$

(since the difference of two simple functions is again a simple function). Hence the theorem.

Littlewood's three principles of measurability

The following three principles concerning measure are due to Littlewood.

First Principle. Every measurable set is a finite union of intervals.

Second Principle. Every measurable function is almost a continuous function.

Third Principle. If $\{f_n\}$ is a sequence of measurable function defined on a set E of finite measure and if $f_n(x) \to f(x)$ on E, then $f_n(x)$ converges almost uniformly on E.

First of all we consider third principle. We shall prove Egoroff's theorem which is a slight modification of third principle of Littlewood's.

Theorem 11(Egoroff's Theorem). Let $\{f_n\}$ be a sequence of measurable functions defined on a set E of finite measure such that $f_n(x) \to f(x)$ almost everywhere. Then to each $\in > 0$ there corresponds a measurable subset E_0 of E such that m $E_0^c < \in$ and $f_n(x)$ converges to f(x) uniformly on E_0.

Proof. Since $f_n(x) \to f(x)$ almost everywhere and $\{f_n\}$ is a sequence of measurable functions, therefore f(x) is also a measurable function. Let

$$H = \{ x \mid \lim_{n\to\infty} f_n(x) = f(x) \}$$

Clearly measure of E–H is zero.

For each pair (k, n) of positive integers, let us define the set

$$E_{kn} = \bigcap_{m=n}^{\infty}\{x \mid |f_m(x) - f(x)| < \frac{1}{k} \}$$

(Since each f_m–f is a measure function, the sets E_{kn} are measurable).

Then for each k, if we put

$$E' = \bigcup_{n=1}^{\infty} E_{kn}$$

Then it is clear that

$$E' = \bigcup_{n=1}^{\infty} E_{kn} \supset H$$

In fact, if $x \varepsilon H$ then $x \varepsilon E' \Rightarrow H \subset E'$.

We have also

$$E_{k(n+1)} = \bigcap_{m=n+1}^{\infty}\{x \mid |f_m(x) - f(x)| < \frac{1}{k}\}$$

Clearly

$$E_{kn} = E_{k(n+1)} \cap \{x \mid |f_n(x) - f(x)| < \frac{1}{k} \}$$

Hence $E_{k(n+1)}$ cannot be a proper subset of E_{kn}. That is,

$$E_{kn} \subset E_{k(n+1)}$$

Thus for each k the sequence $[E_{kn}]$ is an expanding sequence of measurable sets. Therefore

$$\lim_{n\to\infty} m(E_{kn}) = m(\bigcup_{n=1}^{\infty} E_{kn})$$
$$\geq m(H) = m(E),$$

whence

$$\lim_{n\to\infty} m(E_{kn}^c) = 0 . \qquad \text{(i)}$$

Thus, given $\in > 0$, we have that for each k there is a positive integer n_k such that

$$|m\ E_{kn}^c - 0| < \frac{\in}{2^k}, \ n \geq n_k$$

i.e.
$$|m\ E_{kn}^c| < \frac{\in}{2^k}, \quad n \geq n_k \qquad \text{(ii)}$$

Let

$$E_0 = \bigcap_{k=1}^{\infty} E_k\, n_k ,$$

then E_0 is measurable and

$$m\ E_0^c = m(\bigcap_{k=1}^{\infty} E_{kn_k})^c$$

$$= m\left(\bigcup_{k=1}^{\infty} E_{kn_k}^c\right)$$

$$\leq \sum_{k=1}^{\infty} m\, E_{kn_k}^c$$

$$= \sum_{k=1}^{\infty} \frac{\in}{2^k} \quad \text{(using (ii))}$$

$$= \in \sum_{k=1}^{\infty} \frac{1}{2^k} = \in .$$

It follows from the definition of E_{kn} that for all $m \geq n_k$,

$$| f_m(x) - f(x) | < \frac{1}{k} \qquad \text{(iii)}$$

for every $x \varepsilon E_{kn_k}$. Since $E_0 \subset E_{kn_k}$ for every k, the condition $m \geq n_k$ yields (iii) for every $x \in E_0$. Hence $f_n(x) \to f(x)$ uniformly on E_0. This completes the proof of the theorem.

Now we pass to the **second principle of Littlewood**. This is nothing but approximation of measurable functions by continuous functions. In this connection we shall prove the following theorem known as Lusin Theorem after the name of a Russian Mathematician Lusin, N.N.

Theorem 12 (Lusin's Theorem). Let f be a measurable function defined on [a, b]. Then to each $\in > 0$, there corresponds a measurable subset E_0 of [a, b] such that m $E_0^c < \in$ and f is continuous on E_0.

Proof. Let f be a measurable function defined on [a, b]. We know that every measurable function is the limit of a sequence $\{\phi_n(x)\}$ of simple functions **whose points of discontinuity form a set of measure zero.** Thus we have

$$\lim_{n\to\infty} \phi_n(x) = f, \quad x \varepsilon [a, b]$$

By Egoroff's theorem, to each $\in > 0$ there exists a subset E_0 of [a, b] such that m $E_0^c < \in$ and $\phi_n(x)$ converges to f(x) uniformly on E_0. But we know that if $\{\phi_n(x)\}$ is a sequence of continuous function converging uniformly to a function f(x), then f(x) is continuous. Therefore f(x) is continuous on E_0. This completes the proof of the theorem.

Theorem 13. Let f be a measurable function defined on [a, b] and assume that f takes values $\pm\infty$ on a set of measure zero. Then given $\in > 0$ we can find a continuous function g and a step function h such that

$$|f-g| < \in , \quad (f-h) < \in ,$$

except on a set of measure less than $\in$.

Proof. Let H be a subset of [a, b] where f(x) is not $\pm\infty$. Then by the hypothesis of the theorem mH = m([a, b]). We know that every measurable function can be expressed as a almost everywhere limit of a sequence of step functions which are continuous on a set of measure zero.

That is, we can find a sequence of step functions such that

$$\lim_{n\to\infty} \phi_n(x) = f(x) \text{ a.e. on H.}$$

Let $F \subset H$ such that $\phi_n(x) \to f(x)$ and is continuous everywhere on F.

By Egoroff's theorem for a given $\in > 0$ we can find a subset $F' \subset H$ such that $\phi_n(x) \to f(x)$ uniformly on F′ and

$$M(F - F') < \in$$

But we know that if $\{f_n\}$ is a sequence of continuous function converging uniformly to a function f(x), then f(x) is continuous. Therefore f(x) is continuous on F′.

Define a continuous function g(x) on [a, b] such that

$$g(x) = \begin{cases} 0 \text{ if } x \notin F' \\ f(x) \text{ if } x \in F' \end{cases}$$

Therefore on F′ we have

$$|f-g| < \in$$

We have already shown that

$$m([a, b] - F') < \in .$$

Also we have shown that $\phi_n(x) \to f(x)$ where $\phi_n(x)$ is a sequence of step function, so f(x) is also a step function. Hence the theorem.

In order to prove the **first principle of Littlewood** we prove two theorems on approximations of measurable sets.

Theorem 14. A set E in R is measurable if and only if to each $\in > 0$, there corresponds a pair of sets F, G such that $F \subset E \subset G$, F is closed, G is open and $m(G-F) < \in$.

Proof. Sufficiency :- Taking $\in = \frac{1}{n}$, let the corresponding pair of sets be F_n, G_n with

$$m(G_n - F_n) < \frac{1}{n}$$

Let

$$X = \bigcup_n F_n, \quad Y = \bigcap_n G_n$$

It follows that $Y-X \subset G_n-F_n$ and

$$m(Y-X) \le m(G_n-F_n) < \frac{1}{n}$$

so that

$$m(Y-X) = 0 .$$

Since

$$E-X \subset Y-X ,$$

so

$$m(E-X) = 0 .$$

Therefore, $E - X$ is measurable.

But $E = (E-X) \cup X$. Therefore E is measurable, since X is measurable and $E-X$ is measurable.

Necessity. We now assume that E is measurable. We first prove this part under the assumption that E is bounded. Since E is measurable and bounded, we can choose an open set $G \supset E$ such that

$$m(G) < m(E) + \frac{\in}{2} \qquad \text{(i)}$$

Choose a compact (closed and bounded) set $S \supset E$, and then choose an open set V such that $S-E \subset V$ and

$$m(V) < m(S-E) + \frac{\in}{2} \qquad \text{(ii)}$$

Let $F = S-V$. Then F is closed (since $S-V = S \cap V^c$ which is closed being the intersection of closed sets) and $F \subset E$. We have

$$\begin{aligned} m(F) &= m(S) - m(S \cap V) \\ &\ge m(S) - m(V) \\ &> m(S) - m(S-E) - \frac{\in}{2} \quad \text{(Using (ii))} \\ &= m(E) - \frac{\in}{2} \qquad \text{(iii)} \end{aligned}$$

Then

$$\begin{aligned} m(G-F) &= m(G) - m(F) \\ &= m(G) - m(E) + m(E) - m(F) \\ &< \frac{\in}{2} + \frac{\in}{2} = \in \quad \text{(using (i) and (iii))} \end{aligned}$$

This finishes the proof for the case in which E is bounded.

Now, let E be the measurable but unbounded. Let

$$/ \qquad S_n = \{x \mid |x| \le n\} \quad n \in Z$$

$$E_1 = E \cap S_1$$

$$E_n = E \cap (S_n-S_{n-1}), \quad n \ge 2 .$$

Then

$$E = \bigcup_n E_n ,$$

where each E_n is bounded and measurable.

Using what has already been established, let F_n, G_n be a pair of sets such that $F_n \subset E_n \subset G_n$, F_n is closed, G_n is open, and $m(G_n-F_n) < \frac{\in}{2^n}$. Let $F = \bigcup_n F_n$, $G = \bigcup_n \in_n$. Then $G-F \subset \bigcup_n (G_n-F_n)$ and so

$$m(G-F) \le m\{\bigcup_n (G_n - F_n)\}$$

$$\le \sum_n m(G_n - F_n)$$

$$= \sum_n \frac{\epsilon}{2^n}$$

$$= \epsilon \sum_n \frac{1}{2^n} = \epsilon .$$

We see that G is open and that $F \subset E \subset G$, so all that remains to prove is that F is closed. Suppose $\{x_i\}$ is a convergent sequence (say $x_i \to x$) with $x_i \in F$ for each i. Then $\{x_i\}$ is bounded and so is contained in S_N for certain N . Now $F_n \subset S_n - S_N$ if $n > N$. Therefore, $x_i \in \bigcup_{n=1}^{N} F_n$ for each i. But then the limit x is in $\bigcup_{n=1}^{N} F_n$, for this last set is closed. Therefore F is closed. This finishes the proof.

Definition. If A and B are two sets, then

$$A \triangle B = (A-B) \cup (B-A) .$$

Theorem 15. If E is a measurable set of finite measure in R and if $\epsilon > 0$, there is a set G of the form $G = \bigcup_{n=1}^{N} I_n$ where $I_1, I_2, \ldots, I_N$ are open intervals, such that $m(E \triangle G) < \epsilon$.

Proof. Let us assume at first that E is bounded. Let X be an open interval such that $E \subset X$. There exist Lebesgue covering $\{I_n\}$ and $\{J_n\}$ of E and X–E respectively such that

$$\sum_n | I_n | < m(\epsilon) + \frac{\epsilon}{3} ,$$

$$\sum_n | J_n | < m(X-E) + \frac{\epsilon}{3} ,$$

and such that each I_n and J_n is contained in X. Choose N so that $\sum_{n>N} | I_n | < \frac{\epsilon}{3}$ and define sets G, H, K as follows

$$G = \bigcup_{n=1}^{N} I_n , \quad H = \bigcup_{n>N} I_n , \quad K = G \cap \bigcup_n J_n$$

Observe that $E-G \subset H$ and $G-E \subset K$ so that $E \triangle G \subset H \cup K$ and therefore

$$m(E \triangle G) \le m(H \cup K) \le m(H) + m(K)$$

We know that $m(H) \le \sum_{n>N} m(I_n)$

$$= \sum_{n>N} | I_n |$$

$$< \frac{\epsilon}{3} \qquad \text{(by our choice)}$$

Hence it suffices to prove that $m(K) < \frac{2\epsilon}{3}$. Since

$$K = G \cap \bigcup_n J_n$$

$$= \cup\, G \cap J_n$$

therefore $m(K) = \sum_n m(G \cap J_n)$. So we seek an estimate of $\sum_n m(G \cap J_n)$. Now we can see that

$$X = [\bigcup_n I_n] \cup [\bigcup_n (J_n - G)],$$

whence

$$m(X) = m[\bigcup_n I_n] + m[\bigcup_n (J_n - G)]$$

$$\leq \sum_n |I_n| + \sum_n m(J_n - G)$$

We also have

$$\sum_n |I_n| + \sum_n |J_n| < m(E) + m(X - E) + \frac{2\in}{3}$$

$$= m(X) + \frac{2\in}{3},$$

whence

$$\sum_n |I_n| + \sum_n |J_n| < \sum_n |I_n| + \sum_n m(J_n - G) + \frac{2\in}{3}$$

and therefore, since $J_n = (J_n - G) \cup (J_n \cap G)$,

$$m(K) \leq \sum_n m(G \cap J_n) = \sum_n m(J_n) - \sum_n m(J_n - G)$$

$$< \frac{2\in}{3}$$

Hence when E is bounded

$$m(E\Delta G) < \frac{\in}{3} + \frac{2\in}{3} = \in$$

For the general case, let

$$S_n = \{x \mid |x| \leq n\},$$

$$T_1 = S_1$$

$$T_1 = S_n - S_{n-1}, \ n \geq 2$$

Let $E_n = E \cap S_n$. Then

$$E = \bigcup_{i=1}^{\infty} (E \cap T_i)$$

$$E - E_n = \bigcup_{i=n+1}^{\infty} (E \cap T_i)$$

Because $m(E) < +\infty$, we have

$$m(E - E_n) = \sum_{i=n+1}^{\infty} m(E \cap T_i) \to 0 \text{ as } n \to \infty.$$

But $E\Delta E_n = E - E_n$(since $E\Delta E_n = (E - E_n) \cup (E_n - E)$ and $E_n - E$ is empty) and so $m(E\Delta E_n) \to 0$. Using what has already been proved we can find a sequence G_n which is finite union of open intervals such that $m(E_n \Delta G_n) < \frac{1}{n}$. Now the following inequality is true.

$$m(E\Delta G_n) \leq m(E\Delta E_n) + m(E_n \Delta G_n),$$

since $E\Delta G_n = (E\Delta E_n) \cup (E_n \Delta G_n)$. We see therefore that $m(E\Delta G_n) \to 0$. If $\in > 0$, we shall have $m(E\Delta G_n) < \in$ for a suitable value of n, and then G_n will serve our purpose. This completes the proof of the theorem.

Theorem 16. Let $\in$ be a set with $m^* E < \infty$. Then E is measurable iff given $\in > 0$, there is a finite union B of open intervals such that

$$m^*(E \Delta B) < \in$$

Proof. Suppose E is measurable and let $\in > 0$ be given. The (as already shown) there exists an open set $O \supset E$ such that m^* (O E) $< \frac{\in}{2}$. As m^*E is finite, so is m^*O. Since the open set O can be written as the union of countable (disjoint) open intervals $\{I_i\}$, there exists an $n \in N$ such that

$$\sum_{i=n+1}^{\infty} l(I_i) < \frac{\in}{2} \quad \text{(In fact m* O = } \sum_{i=n+1}^{\infty} l(I_i) < \infty \Rightarrow \sum_{i=n+1}^{\infty} l(I_i) < \frac{\in}{2} \text{ because m* O} < \infty)$$

Set $B = \bigcup_{i=1}^{n} I_i$. Then

$$E \Delta B = (E - B) \cup (B \setminus E) \subset (O \setminus B) \cup (O \setminus E)$$

Hence

$$m^*(E \Delta B) \le m^* \left(\bigcup_{i=n}^{\infty} I_i\right) + m^*(O \setminus E) < \frac{\in}{2} + \frac{\in}{2} = \in .$$

Conversely, assume that for a given $\in > 0$, there exists a finite union $B = \bigcup_{i=n}^{n} I_i$ if open intervals with $m^*(E \Delta B) < \in$. Then using "Let $\in$ be any set. The given $\in > 0$ there exists an open set $O \supset E$ such that $m^* O < m^* E + \in$ there is an open set $O \supset E$ such that

$$m^* O < m^* E + \in \qquad \text{(i)}$$

If we can show that m* (O — E) is arbitrary small, then the result will follow from "Let E be set. Then the following are equivalent (i) E is measurable and (ii) given $\in > 0$ there is an open set $O \supset E$ such that $m^*(O — E) < \in$".

Write

$$S = \bigcup_{i=1}^{n} (I_i \cap O)$$

Then $S \subset B$ and so

$$S \Delta E = (E \setminus S) \cup (S \setminus E) \subset (E — S) \cup (B — E) .$$

However,

$$E \setminus S = (E \cap O^c) \cup (E \cap B^c) = E — B, \text{ because } E \subset O .$$

Therefore

$$S \Delta E \subset (E — B) \cup (B — E) = E \Delta B ,$$

and as such $m^*(S \Delta E) < \in$. However,

$$E \subset S \cup (S \Delta E)$$

and so

$$m^* E < m^* S + m^*(S \Delta E)$$

$$< m^* S + \in \qquad \text{(ii)}$$

Also

$$O — E = (O - S) \cup (S \Delta E)$$

Therefore

$$\begin{aligned} m^*(O \setminus E) &< m^* O - m^* S + \in \\ &< m^* E + \in - m^* S + \in && \text{(using (i))} \\ &< m^* S + \in + \in - m^* S + \in && \text{(using (ii)} \\ &< m^* S + \in + \in - m^* S + \in \\ &= 3 \in . \end{aligned}$$

Hence E is measurable.

"Convergence in Measure"

Definition. A sequence $< f_n >$ of measurable functions is said to converge to f in measure if, given $\in > 0$, there is an N such that for all $n \ge N$ we have

$$m\{x \,|f(x) - f_n(x)| \ge \in \} < \in .$$

F. Riesz Theorem

Theorem 17 (F. Riesz). "Let $< f_n >$ be a sequence of measurable functions which converges in measure to f. Then there is a subsequence $< f_{nk} >$ which converges to **f almost everywhere**."

Proof. Since $< f_n >$ is a sequence of measurable functions which converges in measure to f, for any positive integer k there is an integer n_k such that for $n \ge n_k$ we have

$$m\{x \mid f_n(x) - f(x) \mid \ge \frac{1}{2^k} \} < \frac{1}{2^k}$$

Let

$$E_k = \{x \mid |f_{n_k}(x) - f(x)| \geq \frac{1}{2^k}\}$$

Then if $x \notin \bigcup_{k=i}^{\infty} E_k$, we have

$$|f_{n_k}(x) - f(x)| < \frac{1}{2^k} \quad \text{for } k \geq i$$

and so $f_{n_k}(x) \to f(x)$

Hence $f_{n_k}(x) \to f(x)$ for any $x \notin A = \bigcap_{i=1}^{\infty}\bigcup_{k=i}^{\infty} E_k$

But

$$mA \leq m\left[\bigcup_{k=i}^{\infty} E_k\right]$$

$$= \sum_{k=i}^{\infty} mE_k = \frac{1}{2^{k-1}}$$

Hence measure of A is zero.

Example. An example of a sequence $< f_n >$ which converges to zero in measure on [0, 1] but such that $< f_n(x) >$ does not converge for any x in [0, 1] can be constructed as follows :

Let $n = k + 2^v$, $0 \leq k < 2^v$, and set $f_n(x) = 1$ if $x \in [k2^{-v}, (k+1)2^{-v}]$ and $f_n(x) = 0$ otherwise. Then

$$m\{x \mid |f_n(x)| > \in \} \leq \frac{2}{n} ,$$

and so $f_n \to 0$ in measure, although for any $x \in [0, 1]$, the sequence $< f_n(x) >$ has the value 1 for arbitrarily large values of n and so does not converge.

Definition. A sequence $\{f_n\}$ of a.e. finite valued measurable functions is said to be **fundamental in measure**, if for every $\in > 0$, $m(\{x : |f_n(x) - f_m(x)| \geq \in\}) \to 0$ as n and $m \to \infty$.

Definition. A sequence $\{f_n\}$ of real valued functions is said to be **fundamental a.e.** if there exists a set E_0 of measure zero such that, if $x \notin E_0$ and $\in > 0$, then an integer $n_0 = n_0 = (x, \in)$ can be found with the property that

$$|f_n(x) - f_m(x)| < \in , \text{ whenever } n \geq n_0 \text{ and } m \geq n_0 .$$

Definition. A sequence $\{f_n\}$ of a.e. finite valued measurable functions will be said to converge to the measurable function f **almost uniformly** if, for every $\in > 0$, there exists a measurable set F such that $m(F) < \in$ and such that the sequence $\{f_n\}$ converges to f uniformly on F^c.

In this Language, Egoroff's Theorem asserts that **on a set of finite measure convergence a.e. implies almost uniform convergence.**

The following result goes in the converse direction.

Theorem 18. If $\{f_n\}$ is a sequence of measurable functions which converges to f almost uniformly, then $\{f_n\}$ converges to f a.e.

Proof. Let F_n be a measurable set such that $m(F_n) < \frac{1}{n}$ and such that the sequence $\{f_n\}$ converges to f uniformly on F_n^c, n = 1,2,… If $F = \bigcap_{n=1}^{\infty} F_n$, then

$$m(F) \leq \mu(F_n) < \frac{1}{n} ,$$

so that m(F) = 0 , and it is clear that, for $x \in F^c$, $\{f_n(x)\}$ converges to f(x).

Theorem 19. Almost uniform convergence implies convergence in measure.

Proof. If $\{f_n\}$ converges to f almost uniformly, then for any two positive numbers $\in$ and δ there exists a measurable set F such that $m(F) < \delta$ and such that $|f_n(x) - f(x)| < \in$, whenever x belongs to F^c and n is sufficiently large.

Theorem. If $\{f_n\}$ converges in measure to f, then $\{f_n\}$ is fundamental in measure. If also $\{f_n\}$ converges in measure to g, then f = g .a.e.

Proof. The first assertion of the theorem follows from the relation

$$\{x : |f_n(x) - f_m(x)| \geq \in\} \subset \{x : |f_n(x) - f(x)| \geq \frac{\in}{2}\} \cup \{x : |f_m(x) - f(x)| \geq \frac{\in}{2}\}$$

To prove the second assertion, we have

$$\{x : |f(x) - g(x)| \geq \in\} \subset \{x : f_n(x) - f(x)\,| \geq \frac{\in}{2}\} \cup \{x : |f_n(x) - g(x)| \geq \frac{\in}{2}\}$$

Since by proper choice of n, the measure of both sets on the right can be made arbitrarily small, we have

$$m(\{x : |f(x) - g(x)| \geq \in\}) = 0$$

for every $\in > 0$ which implies that f = g a.e.

Theorem 20. If $\{f_n\}$ is a sequence of measurable functions which is fundamental in measure, then some subsequence $\{f_{n_k}\}$ is almost uniformly fundamental.

Proof. For any positive integer k we may find an integer $\bar{n}(k)$ such that if $n \geq \bar{n}(k)$ and $m \geq \bar{n}(k)$, then

$$m(\{x : |f_n(x) - f_m(x)| \geq \frac{1}{2^k}\}) < \frac{1}{2^k}.$$

We write

$n_1 = \bar{n}(1)$, $n_2 = (n_1+1) \cup \bar{n}(2)$, $n_3 = (n_2+1) \cup \bar{n}(3), \ldots$; then $n_1 < n_2 < n_3 < \ldots$,

So that the sequence $\{f_{n_k}\}$ is indeed on subsequence of $\{k_n\}$. If

$$E_k = \{x : |f_{n_k}(x) - f_{n_k+1}(x)| \geq \frac{1}{2^k}\}$$

and $k \leq i \leq j$, then, for every x which does not belong to $E_k \cup E_{k+1} \cup E_{k+2} \cup \ldots$, we have

$$|f_{n_i}(x) - f_{n_j}(x)| \leq \sum_{m=i}^{\infty} |f_{n_m}(x) - f_{n_{m+1}}(x)| < \sum_{m=i}^{\infty} \frac{1}{2^m} = \frac{1}{2^{i-1}},$$

so that, in other words, the sequence $\{f_{n_i}\}$ is uniformly fundamental on

$E \setminus (E_k \cup E_{k+1} \cup \ldots)$. Since

$$m(E_k \cup E_{k+1} \cup \ldots) \leq \sum_{m=k}^{\infty} m(E_m) < \frac{1}{2^{k-1}}$$

This completes the proof of the theorem.

Theorem 21. If $\{f_n\}$ is a sequence of measurable functions which is fundamental in measure, then there exists a measurable function f such that $\{f_n\}$ converges in measure to f.

Proof. By the above theorem we can find a subsequence $\{f_{n_k}\}$ which is almost uniformly fundamental and therefore fundamental a.e. We write $f(x) = \lim_{k\to\infty} f_{n_k}(x)$ for every x for which the limit exists. We observe that, for every $\in >$ 0 ,

$$\{x : |f_n(x) - f(x)| \geq \in] \subset \{x : |f_n(x) - f_{n_k}(x)| \geq \frac{\in}{2}\} \cup \{x : |f_{n_k}(x) - f(x)| \geq \frac{\in}{2}\}.$$

The measure of the first term on the right is by hypothesis arbitrarily small if n and n_k are sufficiently large, and the measure of the second term also approaches 0 (as $k\to\infty$), since almost uniform convergence implies convergence in measure. Hence the theorem follows.

Remark. Convergence in measure does not necessarily imply convergence pointwise at any point. Let

$$E_{r,k} = [\frac{r-1}{2^k}, \frac{r}{2^k}] \quad (r = 1,2,\ldots, 2^k, \; k = 1,2,\ldots\} ,$$

and arrange these intervals as a single sequence of sets $\{F_n\}$ by taking first those for which $k = 1$, then those with $k = 2$, etc. If m denotes Lebesgue measure on [0,1], and $f_n(x)$ is the indicator function of F_n, then for $0 < \in < 1$,

$$\{x : |f_n(x)| \geq \in\} = F_n$$

so that, for any $\in > 0$, $m \{x : |f_n(x)| \geq \in\} \leq m(F_n) \to 0$. This means that $f_n \to 0$ in measure in [0, 1]. However, at no point $x \in [0, 1]$ does $f_n(x) \to 0$; in fact, since every x is in infinitely many of the sets F_n and infinitely many of the sets $(\Omega - F_n)$ we have

$$\liminf f_n(x) = 0 , \quad \limsup f_n(x) = 1 \quad \text{for all } x \,\varepsilon\, [0, 1].$$

PART B : "THE LEBESGUE INTEGRAL"

4.2. The shortcomings of the Riemann integral suggested the further investigations in the theory of integration. We give a resume of the Riemann Integral first.

Let f be a bounded real-valued function defined on the interval [a,b] and let

$$a = \xi_0 < \xi_1 < \ldots < \xi_n = b$$

be a partition of [a,b]. Then for each partition we define the sums

$$S = \sum_{i=1}^{n} (\xi_i - \xi_{i-1})\, M_i$$

and

$$s = \sum_{i=1}^{n} (\xi_i - \xi_{i-1})\, m_i\ ,$$

where

$$M_i = \sup_{\xi_{i-1} < x \le \xi_i} f(x)\ , \quad m_i = \inf_{\xi_{i-1} < x \le \xi_i} f(x)$$

We then define the upper Riemann integral of f by

$$R\ \overline{\int_a^b} f\ (x)dx = \inf S$$

With the infimum taken over all possible subdivisions of [a,b].

Similarly, we define the lower integral

$$R \underline{\int_a^b} f\ (x)dx = \sup s\ .$$

The upper integral is always at least as large as the lower integral, and if the two are equal we say that f is Riemann integrable and call this common value the Riemann integral of f. We shall denote it by

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

To distinguish it from the Lebesgue integral, which we shall consider later.

By a **step function** we mean a function ψ which has the form

$$\psi(x) = c_i,\ \ \xi_{i-1} < x < \xi_i$$

for some subdivision of [a, b] and some set of constants c_i .

The integral of $\psi(x)$ is defined by

$$\int_a^b \psi(x)dx = \sum_{i=1}^{n} c_i\ (\xi_i - \xi_{i-1})\ .$$

With this in mind we see that

$$R \overline{\int_a^b} f(x)\, dx = \inf \int_a^b \psi(x)dx$$

for all step function $\psi(x) \ge f(x)$.
Similarly,

$$R \underline{\int_a^b} f\ (x)\ dx = \sup \int_a^b \phi(x)dx$$

for all step functions $\phi(x) \le f(x)$.

Example. If

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

then

$$R \overline{\int_a^b} f(x)\,dx = b-a \text{ and } R \underline{\int_a^b} f\,(x)\,dx = 0\,.$$

Thus we see that f(x) is not integrable in the Riemann sense.

4.3. The Lebesgue Integral of a bounded function over a set of finite measure

The example we have cited just now shows some of the shortcomings of the Riemann integral. In particular, we would like a function which is 1 on a measurable set and zero elsewhere to be integrable and have its integral the measure of the set.

The function χ_E defined by

$$\chi_E(x) = \begin{cases} 1 & x \varepsilon E \\ 0 & x \notin E \end{cases}$$

is called the **characteristic function** on E. A linear combination

$$\phi(x) = \sum_{i=1}^{n} a_i \chi_{E_i}(x)$$

is called a **simple** function if the sets E_i are measurable. This representation for ϕ is not unique. However, we note that a function ϕ is simple if and only if it is measurable and assumes only a finite number of values. If ϕ is a simple function and $[a_1,\ldots, a_n]$ the set of non-zero values of ϕ, then

$$\phi = \sum a_i \chi_{A_i},$$

where $A_i = \{x \mid \phi(x) = a_i\}$. **This representation for ϕ is called the canonical representation and it** is characterised by the fact that the A_i are disjoint and the a_i distinct and nonzero.

If ϕ vanishes outside a set of finite measure, we define the integral of ϕ by

$$\int \phi(x)dx = \sum_{i=1}^{n} a_i mA_i$$

when ϕ has the canonical representation $\phi = \sum_{i=1}^{n} a_i \chi_{A_i}$. We sometimes abbreviate the expression for this integral to $\int \phi$. If E is any measurable set, we define

$$\int_E \phi = \int \phi . \chi_E\,.$$

It is often convenient to use representations which are not canonical, and the following lemma is useful.

Lemma. If $E_1, E_2,\ldots, E_n$ are disjoint measurable subset of E then every linear combination

$$\phi = \sum_{i=1}^{n} c_i \chi_{E_i}$$

with real coefficients $c_1, c_2,\ldots,c_n$ is a simple function and

$$\int \phi = \sum_{i=1}^{n} c_i mE_i\,.$$

Proof. It is clear that ϕ is a simple function. Let $a_1, a_2,\ldots, a_n$ denote the non-zero real number in $\phi(E)$. For each j = 1,2,…, n let

$$A_j = \bigcup_{c_i = a_j} E_i$$

Then we have

$$A_j = \phi^{-1}(a_j) = \{x \mid \phi(x) = a_j\}$$

and the canonical representation

$$\phi = \sum_{j=1}^{n} a_j \chi_{A_j}$$

Consequently, we obtain

$$\int\phi = \sum_{j=1}^{n} a_j mA_j$$

$$= \sum_{j=1}^{n} a_j m \quad [\bigcup_{c_i = a_j} E_i\,]$$

$$= \sum_{j=1}^{n} a_j \sum_{c_i = a_j}^{n} mE_i \quad \text{(Since } E_i \text{ are disjoint, additivity of measures applies)}$$

$$= \sum_{j=1}^{n} c_j mE_i$$

This completes the proof of the theorem.

Theorem 22. Let ϕ and ψ be simple functions which vanish outside a set of finite measure. Then

$$\int (a\phi + b\,\psi) = a\int\phi + b\int\psi\ ,$$

and, if $\phi \geq \psi$ a.e, then

$$\int \phi \geq \int \psi\ .$$

Proof. Let $\{A_i\}$ and $\{B_i\}$ be the sets which occur in the canonical representations of ϕ and ψ. Let A_0 and B_0 be the sets where ϕ and ψ are zero. Then the sets E_k obtained by taking all the intersections $A_i \cap B_j$ form a finite disjoint collection of measurable sets, and we may write

$$\phi = \sum_{k=1}^{N} a_k \chi_{E_k}$$

$$\psi = \sum_{k=1}^{N} b_k \chi_{E_k}\ ,$$

and so

$$a\phi + b\psi = a\sum_{k=1}^{N} a_k \chi_{E_k} + b\sum_{k=1}^{N} b_k \chi_{E_k}$$

$$= \sum_{k=1}^{N} aa_k \chi_{E_k} + \sum_{k=1}^{N} bb_k \chi_{E_k}$$

$$= \sum_{k=1}^{N} (aa_k + bb_k)\chi_{E_k}$$

Therefore

$$(a\phi+b\psi) = \sum_{k=1}^{N} (aa_k + bb_k) mE_k$$

$$= \sum_{k=1}^{N} (aa_k) m_{E_k} + \sum_{k=1}^{N} (bb_k) mE_k$$

$$= a\sum_{k=1}^{N} a_k mE_k + b\sum_{k=1}^{N} b_k mE_k$$

$$= a\int \phi + b\int \psi\ .$$

To prove the second statement, we note that

$$\int\phi - \int\psi = \int(\phi-\psi) \geq 0\ ,$$

since the integral of a **simple function which is greater than or equal to zero almost everywhere is non-negative** by the definition of the integral.

Remark. We know that for any simple function ϕ we have

$$\phi = \sum_{k=1}^{N} a_i \chi_{E_i}$$

Suppose that this representation is neither canonical nor the sets E_i's are disjoint. Then using the fact that **characteristic functions are always simple functions** we observe that

$$\begin{aligned}\int\phi &= \int a_1 \chi_{E_1} + \int a_2 \chi_{E_2} + \dots + \int a_n \chi_{E_n} \\ &= a_1 \int \chi_{E_1} + a_2 \int \chi_{E_2} + a_3 \chi_{E_3} + \dots + a_n \int \chi_{E_n} \\ &= a_1 mE_1 + a_2 mE_2 + \dots + \dots + a_n m E_n \\ &= \sum_{k=1}^{N} a_i mE_i\end{aligned}$$

Hence for any representation of ϕ, we have

$$\int\phi = \sum_{k=1}^{N} a_i mE_i$$

Let f be a bounded real-valued function and E a measurable set of finite measure. By analogy with the Riemann integral we consider for simple functions ϕ and ψ the numbers

$$\inf_{\psi \geq f} \int_E \psi$$

and

$$\sup_{\phi \leq f} \int_E \phi,$$

and ask when **these two numbers are equal.** The answer is given by the following proposition :

Theorem 23. Let f be defined and bounded on a measurable set E with mE finite. In order that

$$\inf_{f \leq \psi} \int_E \psi(x)dx = \sup_{f \geq \psi} \int_E \phi(x)dx$$

for all simple functions ϕ and ψ, it is necessary and sufficient that f be measurable.

Proof. Let f be bounded by M and suppose that f is measurable. Then the sets

$$E_k = \left\{ x \mid \frac{KM}{n} \geq f(x) > \frac{(K-1)M}{n} \right\}, \quad -n \leq K \leq n,$$

are measurable, disjoint and have union E. Thus

$$\sum_{k=-n}^{n} mE_k = mE$$

The simple function defined by

$$\psi_n(x) = \frac{M}{n} \sum_{k=-n}^{n} k\chi_{E_k}(x)$$

and

$$\phi_n(x) = \frac{M}{n} \sum_{k=-n}^{n} (k-1)\chi_{E_k}(x)$$

satisfy

$$\phi_n(x) \leq f(x) \leq \psi_n(x)$$

Thus

$$\inf \int_E \psi(x)dx \leq \int_E \psi_n(x)dx = \frac{M}{n} \sum_{k=-n}^{n} km\, E_k$$

and

$$\sup \int_E \phi(x)dx \geq \int_E \phi_n(x)dx = \frac{M}{n}\sum_{k=-n}^{n}(k-1)m\,E_k ,$$

whence

$$0 \leq \inf \int_E \psi(x)dx - \sup_E \int \phi\,(x)dx \leq \frac{M}{n}\sum_{k=-n}^{n} m\,E_k = \frac{M}{n} mE .$$

Since n is arbitrary we have

$$\inf \int_E \psi(x)dx - \sup \int_E \phi(x)dx = 0 ,$$

and the condition is sufficient.

Suppose now that

$$\inf_{\psi \geq f} \int_E \psi\,(x)dx = \sup_{\phi \leq f} \int_E \phi\,(x)dx .$$

Then given n there are simple functions ϕ_n and ψ_n such that

$$\phi_n(x) \leq f(x) \leq \psi_n(x)$$

and

(4.3.1) $$\int \psi_n(x)\,dx - \int \phi_n(x)dx < \frac{1}{n}$$

Then the functions

$$\psi^* = \inf \psi_n$$

and $$\phi^* = \sup \phi_n$$

are measurable and

$$\phi^*(x) \leq f(x) \leq \psi^*(x) .$$

Now the set

$$\Delta = \{x \mid \phi^*(x) < \psi^*(x)\}$$

is the union of the sets

$$\Delta_\nu \{x \mid \phi^*(x) < \psi^*(x) - \frac{1}{\nu}\}.$$

But each Δ_ν is contained in the set $\{x \mid \phi_n(x) < \psi_n(x) - \frac{1}{\nu}\}$, and this latter set by (4.3.1) has measure less than $\frac{\nu}{n}$. Since n is arbitrary, $m\Delta_\nu = 0$ and so $m\Delta = 0$. Thus $\phi^* = \psi^*$ except on a set of measure zero, and $\phi^* = f$ except on a set of measure zero. Thus f is measurable and the condition is also necessary.

Definition. If f is a bounded measurable function defined on a measurable set E with mE finite, we define the Lebesgue integral of f over E by

$$\int_E f(x)dx = \inf \int_E \psi(x)dx$$

for all simple functions $\psi \geq f$.

By the previous theorem, this may also be defined as

$$\int_E f(x)dx = \sup \int_E \phi(x)dx$$

for all simple functions $\phi \leq f$.

We sometimes write the integral as $\int_E f$. If E = [a, b] we write $\int_a^b f$ instead of $\int_{[a,b]} f$

Definition and existence of the Lebesgue integral for bounded functions.

Definition. Let F be a bounded function on E and let E_k be a subset of E. Then we define $M[f, E_k]$ and $m[f, E_k]$ as

$$M[f; E_k] = \underset{x \varepsilon E_k}{\text{l.u.b}}\, f(x)$$

$$m[f, E_k] = \underset{x \varepsilon E_k}{\text{g.l.b}}\, f(x)$$

Definition. By a **measurable partition of E we mean a finite collection** P = $\{E_1, E_2, \ldots, E_n\}$of **measurable** subsets of E such that

$$\bigcup_{k=1}^{n} E_k = E$$

and such that $m(E_j \cap E_k) = 0$ (j, k = 1, …, n , j≠ k).
The sets $E_1, E_2, .., E_n$ are called the **components of P.**

If P and Q are measurable partitions, then Q is called a refinement of P if every component of Q is wholly contained in some component of P.
Thus a measurable partition P is a finite collection of subsets whose union is all of E and whose intersections with one another have measure zero.

Definition. Let f be a bounded function on E and let P = $\{E_1, \ldots, E_n\}$ be any measurable partition E. We define the upper sum U[f, P] as

$$U[f; P] = \sum_{k=1}^{n} M[f; E_k].m\,E_k$$

Similarly, we define the lower sum L[f; P] as

$$L[f; P] = \sum_{k=1}^{n} m[f; E_k].m\,E_k$$

As in the case of Riemann integral, we can see that every upper sum for f is greater than or equal to every lower sum for f.
We then define the Lebesgue upper and lower integrals of a bounded function f on E by

$$\inf_{P} U\,[f; P] \quad \text{and} \sup L[f; P]$$

respectively taken over all measurable position of E. We denote them respectively by

$$\overline{\int_E} f \quad \text{and} \quad \underline{\int_E} f$$

Definition. We say that a bounded function f on E is Lebesgue integrable on E if

$$\overline{\int_E} f = \underline{\int_E} f$$

Also we know that if ψ is a simple function, then

$$\int_E \psi = \sum_{k=1}^{n} a_k mE_k$$

Keeping this in mind, we see that

$$\overline{\int_E} f = \inf \overline{\int_E} \psi\,(x)\,dx$$

for all simple functions $\psi(x) \geq f(x)$. Similarly

$$\underline{\int_E} f = \sup \underline{\int_E} \phi(x)dx$$

for all simple functions $\phi(x) \leq f(x)$.
Now we use the theorem :

"Let f be defined and bounded on a measurable set E with mE finite. In order that

$$\inf_{f \leq \psi} \int_E \psi(x)dx = \sup_{f \geq \phi} \int_E \phi(x)dx$$

for all simple functions ϕ and ψ, it is necessary and sufficient that f is measurable."
And our definition of **Lebesgue integration** takes the form :

"If f is a bounded measurable function defined on a measurable set E with mE finite, we define the (Lebesgue) integral of f over E by

$$\int_E f(x)dx = \inf \int_E \psi(x)dx$$

for all simple functions $\psi \geq f$."

The following theorem shows that the Lebesgue integral is in fact a generalization of the Riemann integral.

Theorem 24. Let f be a bounded function defined on [a, b]. If f is Riemann integrable on [a, b], then it is measurable and

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

Proof. Since f is a bounded function defined on [a, b] and is Riemann integrable, therefore,

$$R\int_a^{\bar{b}} f(x)dx = \inf_{\phi \geq f} \int_a^b \phi(x)dx$$

and

$$R\int_{\underline{a}}^b f(x)dx = \operatorname{suf}_{\psi \leq f} \int_a^b \psi(x)dx$$

for all step functions ϕ and ψ and then

$$R\int_a^{\bar{b}} f(x)dx = R\int_{\underline{a}}^b f(x)dx \Rightarrow \inf_{\phi \geq f} \int_a^b \phi(x)dx = \sup_{\psi \leq f} \int_a^b \psi(x)dx \qquad \text{(i)}$$

Since every step function is a simple function, we have

$$R\int_{\underline{a}}^b f(x)dx = \operatorname{suf}_{\psi \leq f} \int_a^b \psi(x)dx \leq \inf_{\phi \geq f} \int_a^b \phi(x)dx \leq R\int_a^{\bar{b}} f(x)dx$$

Then (i) implies that

$$\sup_{\psi \leq f} \int_a^b \psi(x)dx = \inf_{\phi \geq f} \int_a^b \phi(x)dx$$

and this implies that f is measurable also.

Comparison of Lebesgue and Riemann integration

(1) The most obvious difference is that in Lebesgue's definition we divide up the interval into subsets while in the case of Rimann we divide it into subintervals.

(2) In both Riemann's and Lebesgue's definitions we have upper and lower sums which tend to limits. In the Riemann case the two integrals are not necessarily the same and the function is integrable only if they are the same. In the Lebesgue case the two integrals are necessarily the same, their equality being consequence of the assumption that the function is measurable.

(3) Lebesgues's definition is more general than Riemann. We know that if function is the R-integrable then it is Lebesgue integrable also, but the converse need not be true. For example the characteristic function of the set of irrational points have Lebesgue integral but is not R-integrable.

Let χ be the characteristic function of the irrational numbers in [0,1]. Let E_1 be the set of irrational numbers in [0,1], and let E_2 be the set of rational numbers in [0,1]. Then $P = [E_1, E_2]$ is a measurable partition of (0, 1]. Moreover, χ is identically 1 on E_1 and χ is identically 0 on E_2. Hence $M[\chi, E_1] = m[\chi, E_1] = 1$, while $M[\chi, E_2] = m[\chi, E_2] = 0$. Hence $U[\chi, P] = 1.mE_1 + 0.m\,E_2 = 1$. Similarly $L(\chi, P] = 1.m\,E_1 + 0.\,M\,E_2 = 1$. Therefore, $U\,[\chi, P] = L[\chi, P]$.

Therefore, it is Lebesgue integrable.

For Riemann integration

$$M[\chi, J] = 1, \; m[\chi, J] = 0$$

for any interval $J \subset [0, 1]$

$\therefore$ $U[\chi, J] = 1, \; L[\chi, J] = 0$.

$\therefore$ The function is not Riemann-integrable.

Theorem 25. If f and g are bounded measurable functions defined on a set E of finite measure, then

(i) $$\int_E af = a\int_E f$$

(ii) $$\int_E (f+g) = \int_E f + \int_E g$$

(iii) If $f \le g$ a.e., then

$$\int_E f \le \int_E g$$

(iv) If $f = g$ a.e., then

$$\int_E f = \int_E g$$

(v) If $A \le f(x) \le B$, then

$$AmE \le \int_E f \le BmE.$$

(vi) If A and B are disjoint measurable sets of finite measure, then

$$\int_{A\cup B} f = \int_A g + \int_B f$$

Proof. We know that if ψ is a simple function then so is $a\psi$. Hence

$$\int_E af = \inf_{\psi\ge f} \int_E a\psi = a \inf_{\psi\ge f} \int_E \psi = a\int_E f$$

which proves (i).

To prove (ii) let ε denote any positive real number. There are simple functions $\phi \le f$, $\psi \ge f$, $\xi \le g$ and $\eta \ge g$ satisfying

$$\int_E \phi(x)dx > \int_E f - \in, \qquad \int_E \psi(x)dx < \int_E f + \in,$$

$$\int_E \xi(x)dx > \int_E g - \in, \qquad \int_E \eta(x)dx < \int_E g + \in,$$

Since $\phi + \xi \le f + g \le \psi + \eta$, we have

$$\int_E (f+g) \ge \int_E (\phi+\xi) = \int_E \phi + \int_E \xi > \int_E f + \int g - 2\in$$

$$\int_E (f+g) \le \int_E (\psi+\eta) = \int_E \psi + \int_E \eta < \int_E f + \int_E g + 2\in$$

Since these hold for every $\in > 0$, we have

$$\int_E (f+g) = \int_E f + \int_E g$$

To prove (iii) it suffices to establish

$$\int_E (g-f) \ge 0$$

For every simple function $\psi \ge g-f$, we have $\psi \ge 0$ almost everywhere in E. This means that

$$\int_E \psi \ge 0$$

Hence we obtain

$$\int_E (g-f) = \inf_{\psi\ge(g-f)} \int_E \psi(x)dx \ge 0 \qquad (1)$$

which establishes (iii).

Similarly we can show that

$$\int_E (g-f) = \sup_{\psi \le (g-f)} \int_E \psi(x)dx \le 0 \qquad (2)$$

Therefore, from (1) and (2) the result (iv) follows.

To prove (v) we are given that

$$A \le f(x) \le B$$

Applying (iv) we get

$$\int_E f(x)dx \le \int_E B dx = B\int_E dx$$
$$= BmE$$

That is,

$$\int_E f \le BmE$$

Similarly we can prove that $\int_E f \ge AmE$.

Now we prove (vi).

We know that

$$\chi_{A\cup B} = \chi_A + \chi_B$$

Therefore,

$$\int_{A\cup B} f = \int_{A\cup B} \chi_{A\cup B} f = \int_{A\cup B} f(\chi_A + \chi_B)$$
$$= \int_{A\cup B} f\chi_A + \int_{A\cup B} f\,\chi_B$$
$$= \int_A f + \int_B f$$

which proves the theorem.

Theorem 26 (Lebesgue Bounded Convergence Theorem). Let $< f_n >$ be a sequence of measurable functions defined on a set E of finite measure and suppose that $<f_n >$ is uniformly bounded, that is, there exists a real number M such that $|f_n(x)| \le M$ for all n ε N and all x ε E. If $\lim_{n\to\infty} f_n(x) = f(x)$ for each x in E, then

$$\int_E f = \lim_{n\to\infty} \int_E f_n .$$

Proof. We shall apply Egoroff's theorem to prove this theorem. Accordingly for a given $\in > 0$, there is an N and a measurable set $E_0 \subset E$ such that $mE_0^c < \frac{\in}{4M}$ and for $n \ge N$ and x ε E_0 we have

$$|f_n(x) - f(x)| < \frac{\in}{2m(E)}$$

Then we have

$$\left| \int_E f_n - \int_E f \right| = \left| \int_E (f_n - f) \right| \le \int_E |f_n - f|$$
$$= \int_{E_0} |f_n - f| + \int_{E_0^c} |f_n - f|$$

$$< \frac{\in}{2m(E)}.m(E_0)+\frac{\in}{4M}.2M$$

$$< \frac{\in}{2}+\frac{\in}{2}=\in.$$

Hence

$$\int_E f_n \to \int_E f.$$

The integral of a non-negative function

Definition. If f is a non-negative measurable function defined on a measurable set E, we define

$$\int_E f = \sup_{h\le f}\int_E h,$$

where h is a bounded measurable function such that $m\{x|h(x) \neq 0\}$ is finite.

Theorem 27. If f and g are non-negative measurable functions, then

(i) $$\int_E cf = c\int_E f, \quad c > 0$$

(ii) $$\int_E (f+g) = \int_E f + \int_E g$$

and

(iii) If $f \le g$ a.e., then

$$\int_E f \le \int_E g.$$

Proof. The proof of (i) and (iii) follow directly from the theorem concerning properties of the integrals of bdd functions.

We prove (ii) in detail.

If $h(x) \le f(x)$ and $k(x) \le g(x)$, we have $h(x) + k(x) \le f(x) + g(x)$, and so

$$\int_E (h+k) \le \int_E (f+g)$$

i.e. $$\int_E h + \int_E k \le \int_E (f+g)$$

Taking suprema, we have

(iv) $$\int_E f + \int_E g \le \int_E (f+g)$$

On the other hand, let l **be a bounded measurable function** which vanishes outside a set of finite measure and which is not greater than (f+g). Then we define the functions h and k by setting

$h(x) = \min(f(x), l(x))$

and

$k(x) = l(x) - h(x)$

We have

$h(x) \le f(x)$,

$k(x) \le g(x)$,

while h and k are bounded by the bound l and vanish where l vanishes. Hence

$$\int_E l = \int_E h + \int_E k \le \int_E f + \int_E g$$

and so taking supremum, we have

$$\sup_{l \le f+g} \le \int_E f + \int_E g$$

that is,

$$\text{(v)} \qquad \int_E f + \int_E g \ge \int_E (f+g)$$

From (iv) and (v), we have

$$\int_E (f+g) = \int_E f + \int_E g \,.$$

Fatou's Lemma. If $< f_n >$ is a sequence of non-negative measurable functions and $f_n(x) \to f(x)$ almost everywhere on a set E, then

$$\int_E f \le \underline{\lim} \int_E f_n$$

Proof. Let h be a bounded measurable function which is not greater than f and which vanishes outside a set E′ of finite measure. Define a function h_n by setting

$$h_n(x) = \min\{h(x), f_n(x)\}$$

Then h_n is bounded by the bounds for h and vanishes outside E′. Now $h_n(x) \to h(x)$ for each x in E′. Therefore by "**Bounded Convergence Theorem**" we have

$$\int_E h = \int_{E'} h = \lim \int_{E'} h_n \le \underline{\lim} \int_E f_n$$

Taking the supremum over h, we get

$$\int_E f \le \underline{\lim} \int_E f_n$$

Theorem 28 (Lebesgue Monotone Convergence Theorem). Let $< f_n >$ be an **increasing sequence** of non-negative measurable functions and let $f = \lim f_n$. Then

$$\int f = \lim \int f_n$$

Proof. By Fatou's Lemma we have

$$\int f \le \underline{\lim} \int f_n$$

But for each n we have $f_n \le f$, and so $\int f_n \le \int f$. But this implies

$$\overline{\lim} \int f_n \le \int f$$

Hence

$$\int f = \lim \int f_n$$

Definition. A non-negative measurable function f is called integrable over the measurable set E if

$$\int_E f < \infty$$

Theorem 29. Let f and g be two non-negative measurable functions. If f is integrable over E and $g(x) < f(x)$ on E, then g is also integrable on E, and

$$\int_E (f-g) = \int_E f - \int_E g$$

Proof. Since

$$\int_E f = \int_E (f-g) + \int_E g$$

and the left handside is finite, the term on the right must also be finite and so g is integrable.

Theorem 30. Let f be a non-negative function which is integrable over a set E. Then given $\epsilon > 0$ there is a $\delta > 0$ such that for every set $A \subset E$ with $mA < \delta$ we have

$$\int_A f < \epsilon \,.$$

Proof. If $|f| \leq K$, then

$$\int_A f \leq \int_A K = KmA$$

Set $\delta < \frac{\in}{K}$. Then

$$\int_A f < K.\frac{\in}{K} = \in .$$

Set $f_n(x) = f(x)$ if $f(x) \leq n$ and $f_n(x) = n$ otherwise. Then each f_n is bounded and f_n converges to f at each point. By the monotone convergence theorem there is an N such that $\int_E f_N > \int_E f - \frac{\in}{2}$, and $\int_E (f - f_N) < \frac{\in}{2}$. Choose $\delta < \frac{\in}{2N}$. If $m A < \delta$, we have

$$\int_A f = \int_A (f - f_N) + \int_A f_N$$

$$< \int_E (f - f_N) + NmA \qquad \left(\text{since } \int_A f_N \leq \int_A N = NmA\right)$$

$$< \frac{\in}{2} + \frac{\in}{2} = \in .$$

The General Lebesgue Integral

We have already defined the positive part f^+ and negative part f^- of a function as

$$f^+ = \max(f, 0)$$

$$\bar{f} = \max(-f, 0)$$

Also it was shown that

$$f = f^+ - \bar{f}$$

$$|f| = f^+ + \bar{f}$$

With these notions in mind, we make the following definition.

Definition. A measurable function f is said to be integrable over E if f^+ and $\bar{f}$ are both integrable over E. In this case we define

$$\int_E f = \int_E f^+ - \int_E \bar{f}$$

Theorem 31. Let f and g be integrable over E. Then

(i) The function f+g is integrable over E and

$$\int_E (f + g) = \int_E f + \int_E g$$

(ii) If $f \leq g$ a.e., then $\int_E f \leq \int_E g$

(iii) If A and B are disjoint measurable sets contained in E, then

$$\int_{A \cup B} f = \int_A f + \int_B f$$

Proof. By definition, the functions f^+, $\bar{f}$, g^+, $\bar{g}$ are all integrable. If $h = f+g$, then $h = (f^+ - \bar{f}) + (g^+ - \bar{g})$ and hence $h = (f^+ + g^+) - (\bar{f} + \bar{g})$. Since $f^+ + g^+$ and $\bar{f} + \bar{g}$ are integrable therefore their difference is also integrable. Thus h is integrable.

We then have

$$\int_E h = \int_E ([f^+ + g^+) - (\bar{f} + \bar{g})]$$

$$= \int_E (f^+ + g^+) - \int_E (\bar{f} + \bar{g})$$

$$= \int_E f^+ + \int_E g^+ - \int_E \bar{f} - \int_E \bar{g}$$

$$= (\int_E f^+ - \int_E \bar{f}) + (\int_E g^+ - \int_E \bar{g})$$

That is,

$$\int_E (f+g) = \int_E f + \int_E g$$

Proof of (ii) follows from part (i) and the fact that the integral of a non-negative integrable function is non-negative.

For (iii) we have

$$\int_{A\cup B} f = \int f\chi_{A\cup B}$$

$$= \int f\,\chi_A + \int f\chi_B$$

$$= \int_A f + \int_B f$$

* It should be noted that f+g is not defined at points where $f = \infty$ and $g = -\infty$ and where $f = -\infty$ and $g = \infty$. However, the set of such points must have measure zero, since f and g are integrable. Hence the integrability and the value of $\int(f+g)$ is independent of the choice of values in these ambiguous cases.

Theorem 32. Let f be a measurable function over E. Then f in integrable over E iff $|f|$ is integrable over E. Moreover, if f is integrable, then

$$|\int_E f| \le \int_E |f|$$

Proof. If f is integrable then both f^+ and f^- are integrable. But $|f| = f^+ + f^-$. Hence integrability of f^+ and f^- implies the integrability of $|f|$.

Moreover, if f is integrable, then since $f(x) \le |f(x)| = |f|(x)$, the property which states that if $f \le g$ a.e. , then $\int f \le \int g$ implies that

$$\int f \le \int |f| \qquad \text{(i)}$$

On ther other hand since $-f(x) \le |f(x)|$, we have

$$-\int f \le \int |f| \qquad \text{(ii)}$$

From (i) and (ii) we have

$$|\int f| \le \int |f| .$$

Conversely, suppose f is measurable and suppose $|f|$ is integrable. Since

$$0 \le f^+(x) \le |f(x)|$$

it follows that f^+ is integrable. Similarly f^- is also integrable and hence f is integrable.

Lemma. Let f be integrable. Then given $\in > 0$ there exists $\delta > 0$ such that $|\int_A f| < \in$ whenever A is a measurable subset of E with $mA < \delta$.

Proof. When f is non-negative, the lemma has been proved already. Now for arbitrary measurable function f we have $f = f^+ - f^-$. So by that we have proved already, given $\in > 0$, there exists $\delta_1 > 0$ such that

$$\int_A f^+ < \frac{\in}{2} ,$$

when $mA < \delta_1$. Similarly there exists $\delta_2 > 0$ such that

$$\int_A f^- < \frac{\epsilon}{2},$$

when $mA < \delta_2$. Thus if $mA < \delta = \min(\delta_1, \delta_2)$, we have

$$\left|\int_A f\right| \le \int_A |f| = \int_A f^+ + \int_A f^- < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

This completes the proof.

Theorem 33 (Lebesgue Dominated Convergence Theorem). Let a sequence $< f_n >$, $n \in N$ of measurable functions be dominated by an integrable function g, that is,

$$|f_n(x)| \le g(x)$$

holds for every $n \in N$ and every $x \in E$ and let $< f_n >$ converges pointwise to a function f, that is, $f(x) = \lim\limits_{n\to\infty} f_n(x)$ for almost all x in E. Then

$$\int_E f = \lim_{n\to\infty} \int_E f_n$$

Proof. Since $|f_n| \le g$ for every $n \in N$ and $f(x) = \lim f_n x)$, we have $|f| \le g$. Hence f_n and f are integrable. The function $g - f_n$ is non-negative, therefore by Fatou's Lemma we have

$$\int_E g - \int_E f = \int_E (g - f) \le \underline{\lim} \int_E (g - f_n)$$

$$= \int_E g - \overline{\lim} \int_E f_n$$

whence

$$\int_E f \ge \overline{\lim} \int_E f_n$$

Similarly considering $g + f_n$ we get

$$\int_E f \le \overline{\lim} \int_E f_n$$

Consequently, we have

$$\int_E f = \lim \int_E f_n .$$

5

PART A : "DIFFERENTIATION AND INTEGRATION

The "fundamental theorem of the integral calculus" is that differentiation and integration are inverse processes. This general principle may be interpreted in two different ways.

If f(x) is integrable, the function

$$F(x) = \int_a^x f(t)dt \qquad \text{(i)}$$

is called the indefinite integral of f(x); and the principle asserts that

$$F'(x) = f(x) \qquad \text{(ii)}$$

On the other hand, if F(x) is a given function, and f(x) is defined by (ii), the principle asserts that

$$\int_a^x f(t)dt = F(x) - F(a) \qquad \text{(iii)}$$

The main object of this chapter is to consider in what sense these theorems are true.

From the theory of Riemann integration (ii) follows from (i) if x is a point of continuity of f. For we can choose h_0 so small that $|f(t) - f(x)| < \in$ for $|t-x| \le h_0$; and then

$$\left|\frac{F(x+h - F(x)}{h} - f(x)\right| = \left|\frac{1}{h}\int_x^{x+h} \{f(t) - f(x)\}\, dt\right| \le \in \qquad (|h| < h_0),$$

by the mean-value theorem. This proves (ii).

We shall show that more generally this relation holds almost everywhere. Thus differentiation is the inverse of Lebesgue integration.

The problem of deducing (iii) from (ii) is more difficult and even using Lebesgue integral it is true only for a certain class of functions. We require in the first place that F′(x) should exist at any rate almost everywhere and as we shall see this is not necessarily so. Secondly, if F′(x) exists we require that it should be integrable.

5.1. Differentiation of Monotone Functions

Definition. Let *C* be a collection of intervals. Then we say that *C* covers a set E in the sense of Vitali, if for each $\in > 0$ and x in E there is an interval I ε *C* such that x ε I and $l(I) < \in$.

Now we prove the following lemma which will be utilized in proving a result concerning the differentiation of monotone functions.

Lemma 1 (Vitali). Let E be a set of finite outer measure and *C* a collection of intervals which cover E in the sense of Vitali. Then given $\in > 0$ there is a finite disjoint collection $\{I_1,\dots, I_n\}$ of intervals in *C* such that

$$m^*[E - \bigcup_{n=1}^{N} I_n] < \in .$$

Proof. It suffices to prove the lemma in the case that each interval in *C* is closed, for otherwise we replace each interval by its closure and observe that the set of endpoints of $I_1, I_2,\dots, I_N$ has measure zero.

Let O be an open set of finite measure containing E. Since *C* is a Vitali covering of E, we may suppose without loss of generality that each I of *C* is contained in O. We choose a sequence $< I_n>$ of disjoint intervals of *C* by induction as follows :

Let I_1 be any interval in C and suppose $I_1,\dots, I_n$ have already been chosen. Let k_n be the supremum of the lengths of the intervals of C which do not meet any of the intervals $I_1,\dots,I_n$. Since each I is contained in O, we have $k_n \le m\,O < \infty$. Unless $E \subset \bigcup_{i=1}^{n} I_i$, we can find I_{n+1} in C with $l(I_{n+1}) > \frac{1}{2}\,k_n$ and I_{n+1} disjoint from $I_1, I_2,\dots, I_n$.

Thus we have a sequence $< I_n >$ of disjoint intervals of C, and since $\cup I_n \subset O$, we have $\Sigma\, l(I_n) \le m\,O < \infty$. Hence we can find an integer N such that

$$\sum_{N+1}^{\infty} l(I_n) < \frac{\in}{5}$$

Let

$$R = E - \bigcup_{i=1}^{N} I_n$$

It remains to prove that $m^*R < \in$.

Let x be an arbitrary point of R. Since $\bigcup_{i=1}^{N} I_n$ is a closed set not containing x, we can find an interval I in C which contains x and whose length is so small that I does not meet any of the intervals $I_1, I_2,\dots, I_N$. If now $I \cap I_i = \phi$ for $i \le N$, we must have $l(I) \le k_N < 2l\,(I_{N+1})$. Since $\lim l(I_n) = 0$, the interval I must meet at least one of the intervals I_n. Let n be the smallest integer such that I meets I_n. We have $n > N$, and $l(I) \le k_{n-1} \le 2l(I_n)$. Since x is in I, and I has a point in common with I_n, it follows that the distance from x to the midpoint of I_n is at most $l(I) + \frac{1}{2}\,l(I_n) \le \frac{5}{2}\,l(I_n)$.

Let J_m denote the interval which has the same midpoint as I_m and five times the length of I_m. Then we have $x \in J_m$. This proves

$$R \subset \bigcup_{N+1}^{\infty} J_n$$

Hence

$$m^*R \le \sum_{N+1}^{\infty} l(J_n) = 5 \sum_{N+1}^{\infty} l(J_n) < \in .$$

The Four Derivatives of a Function

Whether the differential coefficients

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

exists or not, the four expressions

$$D^+f(x) = \overline{\lim_{h\to 0+}} \frac{f(x+h) - f(x)}{h}$$

$$D^- f(x) = \overline{\lim_{h\to 0+}} \frac{f(x) - f(x-h)}{h}$$

$$D_+f(x) = \underline{\lim_{h\to 0+}} \frac{f(x+h) - f(x)}{h}$$

$$D_-f(x) = \underline{\lim_{h\to 0+}} \frac{f(x) - f(x-h)}{h}$$

always exist. These derivatives are known as Dini Derivatives of the function f.

$D^+ f(x)$ and $D_+ f(x)$ are called upper and lower derivatives on the right and $D^- f(x)$ and $D_- f(x)$ are **called upper and lower derivatives on the left.** Clearly we have $D^+ f(x) \ge D_+ f(x)$ and $D^- f(x) \ge D_- f(x)$. If $D^+ f(x) = D_+ f(x)$, the function f is said to have a **right hand derivative** and if $D^- f(x) = D_- f(x)$, the function is said to have a **left hand derivative.**

If

$D^+ f(x) = D_+ f(x) = D^- f(x) = D_- f(x) \neq \pm \infty$, we say that f is differentiable at x and define $f'(x)$ to be the common value of the derivatives at x.

Theorem 1. Every non-decreasing function f defined on the interval [a, b] is differentiable almost everywhere in [a, b]. The derivative f' is measurable and

$$\int_a^b f'(x)dx \leq f(b) - f(a).$$

Proof. We shall show first that the points x of the open interval (a, b) at which **not** all of the four Dini-derivatives of f are equal form a subset of measure zero. It suffices to show that the following four subsets of (a, b) are of measure zero :

$A = \{x \,\varepsilon\, (a, b) \mid D_- f(x) < D^+ f(x)\}$,
$B = \{x \,\varepsilon\, (a, b) \mid D_+ f(x) < D^- f(x)\}$,
$C = \{x \,\varepsilon\, (a, b) \mid D_- f(x) < D^- f(x)\}$
$D = \{x \,\varepsilon\, (a, b) \mid D_+ f(x) < D^+ f(x)\}$.

To prove $m^* A = 0$, consider the subsets

$A_{u,v} = \{x \,\varepsilon\, (a, b) \mid D_- f(x) < u < v < D^+ f(x)\}$

of A for all rational numbers u and v satisfying u < v. Since A is the union of this countable family $\{A_{u,v}\}$, it is sufficient to prove $m^*(A_{u,v}) = 0$ for all pairs u, v with u < v .

For this purpose, denote $\alpha = m^*(A_{u,v})$ and let $\in$ be any positive real number. Choose an open set $U \supset A_{u,v}$ with $m^* U < \alpha + \in$. Set x be any point of $A_{u,v}$. Since $D_- f(x) < u$, there are arbitrary small closed intervals of the form [x −h, x] contained in U such that

$$f(x) - f(x-h) < uh.$$

Do this for all $x \,\varepsilon\, A_{u,v}$ and obtain a Vitali cover *C* of $A_{u,v}$. Then by Vitali covering theorem there is a finite subcollection $\{J_1, J_2, \ldots, J_n\}$ of disjoint intervals in *C* such that

$$m^*(A_{u,v} - \bigcup_{i=1}^{n} J_i) < \in$$

Summing over these n intervals, we obtain

$$\sum_{i=1}^{n} [f(x_i) - f(x_i - h_i)] < u \sum_{i=1}^{n} h_i$$
$$< u\, m^* U$$
$$< u(\alpha + \in)$$

Suppose that the interiors of the intervals $J_1, J_2, \ldots, J_n$ cover a subset F of $A_{u,v}$. Now since $D^+ f(y) > v$, there are arbitrarily small closed intervals of the form [y, y+k] contained in some of the intervals J_i (i = 1, 2,…, n) such that

$$f(y+k) - f(y) > vk$$

Do this for all $y \,\varepsilon\, F$ and obtain a Vitali cover *D* of F. Then again by Vitali covering lemma we can select a finite subcollection $[K_1, K_2, \ldots, K_m]$ of disjoint intervals in *D* such that

$$m^*[F - \bigcup_{i=1}^{m} K_i] < \in$$

Since $m^*F > \alpha - \in$, it follows that the measure of the subset H of F which is covered by the intervals is greater than $\alpha - 2\in$. Summing over these intervals and keeping in mind that each K_i is contained in a J_n, we have

$$\sum_{i=1}^{n} \{f(x_i) - f(x_i - h_i)\} \geq \sum_{i=1}^{m} [f(y_i + k_i) - f(y_i)]$$
$$> v \sum_{i=1}^{m} k_i$$
$$> v(\alpha - 2\in)$$

so that

$$v(\alpha - 2\in) < u(\alpha + \in)$$

Since this is true for every $\in > 0$, we must have $v\,\alpha \le u\alpha$. Since $u < v$, this implies that $\alpha = 0$. Hence $m^*A = 0$. Similarly, we can prove that $m^*B = 0$, $m^*C = 0$ and $m^*D = 0$.
This shows that

$$g(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

is defined almost everywhere and that f is differentiable whenever g is finite. If we put

$$g_n(x) = n[f\left(x+\frac{1}{n}\right) - f(x)] \quad \text{for } x \,\varepsilon\, [a,b] ,$$

where we re-define $f(x) = f(b)$ for $x \ge b$. Then $g_n(x) \to g(x)$ for almost all x and so g is measurable since every g_n is measurable. Since f is non-decreasing, we have $g_n \ge 0$. Hence, by Fatou's lemma

$$\int_a^b g \le \underline{\lim}\int_a^b g_n = \underline{\lim}\; n\int_a^b [f(x+\frac{1}{n}) - f(x)]dx$$

$$= \underline{\lim}\, n\left[\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} f(x)dx - \int_a^b f(x)dx\right]$$

$$= \underline{\lim}\, n\left[\int_a^b f(x)dx + \int_b^{b+\frac{1}{n}} f(x)dx - \int_a^{a+\frac{1}{n}} f(x)dx - \int_a^b f(x)dx\right]$$

$$= \underline{\lim}\, n\left[\int_b^{b+\frac{1}{n}} f(x)dx - \int_a^{a+\frac{1}{n}} f(x)dx\right]$$

$$\le f(b) - f(a)$$

(Use of $f(x) = f(b)$ for $x \ge b$ for first interval and f non-decreasing in the 2nd integral).
This shows that g is integrable and hence finite almost everywhere. Thus f is differentiable almost everywhere and $g(x) = f'(x)$ almost everywhere. This proves the theorem.

Functions of Bounded Variation

Let f be a real-valued function defined on the interval [a,b] and let $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ be any partition of [a,b].

By the variation of f over the partition $P = \{x_0, x_1,\ldots, x_n\}$ of [a,b], we mean the real number

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

and then

$$V_a^b(f) = \sup \{V(f,P) \text{ for all possible partitions P of } [a,b]\,)$$

$$= \sup_P \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})\,|$$

is called the total variation of f over the interval [a,b]. If $V_a^b(f) < \infty$, then we say that f is a function of bounded variation and we write $f \,\varepsilon\, BV$.

Lemma 2. Every non-decreasing function f defined on the interval [a,b] is of bounded variation with total variation $V_a^b(f) = f(b) - f(a)$.

Prof. For every partition $P = [x_0, x_1, \ldots, x_n\}$ of [a,b] we have

$$V(f,P) = \sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) |$$
$$= \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})$$
$$= f(b) - f(a)$$

This implies the lemma.

Theorem 2 (Jordan Decomposition Theorem). A function f: $[a,b] \to \mathbf{R}$ is of bounded variation if and only if it is the difference of two non-decreasing functions.

Proof. Let f = g–h on [a,b] with g and h increasing. Then for any, subdivision we have

$$\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) | \le \sum_{i=1}^{n} [g(x_i) - g(x_{i-1})] + \sum_{i=1}^{n} [h(x_i) - h(x_{i-1})]$$
$$= g(b) - g(a) + h(b) - h(a)$$

Hence

$$V_a^b(f) \le g(b) + h(b) - g(a) - h(a) ,$$

which proves that f is of bounded variations.

On the other hand, let f be of bounded variation. Define two functions g, h : $[a, b] \to R$ by taking

$$g(x) = V_a^x(f), \;\; h(x) = V_a^x(f) - f(x)$$

for every x ε [a, b]. Then f(x) = g(x) − h(x).

The function g is clearly non-decreasing. On the other hand, for any two real numbers x and y in [a, b] with $x \le y$, we have

$$h(y) - h(x) = [V_a^y(f) - f(y)] - [V_a^x(f) - f(x)]$$
$$= V_x^y(f) - [f(y) - f(x)]$$
$$\ge V_x^y(f) - V_x^y(f) = 0$$

Hence h is also non-decreasing. This completes the proof of the theorem.

Examples. (1) If f is monotonic on [a,b], then f is of bounded variation on [a, b] and V(f) = |f(b) − f(a)| , where V(f) is the total variation.

(2) If f' exists and is bounded on [a, b], then f is of bounded variation. For if $|f'(x)| \le M$ we have

$$\sum_{i=1}^{n} | f(x_i) - f(x_{i-1}) | \le \sum_{i=1}^{n} M(x_i - x_{i-1}) = M(b-a)$$

no matter which partition we choose.

(3) f may be continuous without being of bounded variation. Consider

$$f(x) = \begin{cases} x \sin \dfrac{\pi}{x} & (0 < x \le 2) \\ 0 & (x = 0) \end{cases}$$

Let us choose the partition which consists of the points

$$0, \frac{2}{2^{n-1}}, \frac{2}{2^{n-3}}, \ldots, \frac{2}{5}, \frac{2}{3}, 2$$

Then the sum in the total variation is

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

and this can be made arbitrarily large by taking n large enough, since $\Sigma \dfrac{1}{n}$ diverges.

(4) Since $|f(x) - f(a)| \le V(f)$ for every x on [a,b] it is clear that every function of bounded variation is bounded.

The Differentiation of an Integral

Let f be integrable over [a,b] and let

$$F(x) = \int_a^x f(t)dt$$

If f is positive, h > 0 , then we see that

$$F(x+h) - F(x) = \int_x^{x+h} f(t)dt \geq 0$$

Hence, integral of a positive function is non-decreasing.

We shall show first that F is a function of bounded variation. Then, being function of bounded variation, it will have a finite differential coefficient F' almost everywhere. Our object is to prove that $F'(x) = f(x)$ almost everywhere in [a,b]. We prove the following lemma :

Lemma 3. If f is integrable on [a,b], then the function F defined by

$$F(x) = \int_a^x f(t)dt$$

is a continuous function of bounded variation on [a,b].

Proof. We first prove continuity of F. Let x_0 be an arbitrary point of [a,b]. Then

$$|F(x) - F(x_0)| = \left| \int_{x_0}^{x} f(t)dt \right|$$

$$\leq \int_{x_0}^{x} |f(t)|dt$$

Now the integrability of f implies integrability of |f| over [a,b]. Therefore, given $\in > 0$ there is a $\delta > 0$ such that for every measurable set $A \subset [a, b]$ with measure less than δ, we have $\int_A |f| < \in$.

Hence

$$|F(x) - F(x_0)| < \in \quad \text{whenever} \quad |x-x_0| < \delta_1$$

and so f is continuous.

To show that F is of bounded variation, let $a = x_0 < x_1 < \ldots < x_n = b$ be any partition of [a,b]. Then

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| = \sum_{i=1}^{n} \left| \int_a^{x_i} f(t)dt - \int_a^{x_{i-1}} f(t)dt \right|$$

$$= \sum_{i=1}^{n} \left| \int_{x_{i-1}}^{x_i} f(t)dt \right|$$

$$\leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(t)|\, dt$$

$$= \int_a^b |f(t)|\, dt$$

Thus

$$V_a^b F \le \int_a^b |f(t)|\, dt < \infty$$

Hence F is of bounded variation.

Lemma 9. If f is integrable on [a, b] and

$$\int_a^x f(t)\, dt = 0$$

for all x ε [a,b], then f = 0 almost everywhere in [a,b].

Proof. Suppose f > 0 on a set E of positive measure. Then there is a closed set $F \subset E$ with $mF > 0$. Let O be the open set such that

$$O = (a, b) - F$$

Then either $\int_a^b f \ne 0$ or else

$$0 = \int_a^b f = \int_F f + \int_O f$$

$$= \int_F f + \sum_{n=1}^{\infty} \int_{a_n}^{b_n} f(t)\, dt, \qquad (1)$$

because O is the union of a countable collection $\{(a_n, b_n)\}$ of open intervals.
But, for each n,

$$\int_{a_n}^{b_n} f(t) dt = \int_a^{b_n} f(t)\, dt - \int_a^{a_n} f(t)\, dt$$

$$= F(b_n) - F(a_n) = 0 \text{ (by hypothesis)}$$

Therfore, from (1), we have

$$\int_F f = 0$$

But since f > 0 on F and mF > 0 , we have $\int_F f > 0$.

We thus arrive at a contradiction. Hence f = 0 almost everywhere.

Lemma 5. If f is bounded and measurable on [a, b] and

$$F(x) = \int_F^x f(t) dt + F(a),$$

then $F'(x) = f(x)$ for almost all x in [a,b].

Proof. We know that an integral is of bounded variation over [a,b] and so $F'(x)$ exists for almost all x in [a,b]. Let $|f| \le K$. We set

$$f_n(x) = \frac{F(x+h) - F(x)}{h}$$

with $h = \frac{1}{n}$. Then we have

$$f_n(x) = \frac{1}{h}\left[\int_a^{x+h} f(t) dt - \int_a^x f(t) dt\right]$$

$$= \frac{1}{h}\int_x^{x+h} f(t)dt$$

$$\Rightarrow |f_n(x)| = \left|\frac{1}{h}\int_x^{x+h} f(t)dt\right|$$

$$\leq \frac{1}{h}\int_x^{x+h} |f(t)|\,dt \leq \frac{1}{h}\int_x^{x+h} K\,dt$$

$$= \frac{K}{h}.h = K$$

Moreover,

$$f_n(x) \to F'(x) \text{ a.e.}$$

Hence by the theorem of bounded convergence, we have

$$\int_a^c F'(x)dx = \lim\int_a^c f_n(x)dx = \lim_{h\to 0}\frac{1}{h}\int_a^c [F(x+h) - F(x)]dx$$

$$= \lim_{h\to 0}\left[\frac{1}{h}\int_{a+h}^{c+h} F(x)dx - \frac{1}{h}\int_a^c F(x)dx\right]$$

$$= \lim\left[\frac{1}{h}\int_c^{c+h} F(x)dx - \frac{1}{h}\int_a^{a+h} F(x)dx\right]$$

$$= F(c) - F(a) \quad \text{(since F is continuous)}$$

$$= \int_a^c f(x)dx$$

Hence

$$\int_a^c [F'(x) - f(x)]dx = 0$$

for all $c \in [a,b]$, and so

$$F'(x) = f(x) \text{ a.e.}$$

by using the previous lemma.

Now we extend the above lemma to unbounded functions.

Theorem 3. Let f be an integrable function on [a,b] and suppose that

$$F(x) = F(a) + \int_a^x f(x)dt$$

Then $F'(x) = f(x)$ for almost all x in [a, b].

Proof. Without loss of generality we may assume that $f \geq 0$ (or we may write "From the definition of integral it is sufficient to prove the theorem when $f \geq 0$).

Let f_n be defined by $f_n(x) = f(x)$ if $f(x) \leq n$ and $f_n(x) = n$ if $f(x) > n$. Then $f - f_n \geq 0$ and so

$$G_n(x) = \int_a^x (f - f_n)$$

is an increasing function of x, which must have a derivative almost everywhere and this derivative will be non-negative. Also by the above lemma, since f_n is bounded (by n), we have

$$\frac{d}{dx}\left(\int_a^x f_n\right) = f_n(x) \text{ a.e.} \qquad \text{(i)}$$

Therefore,

$$F'(x) = \frac{d}{dx}\left(\int_a^x f\right) = \frac{d}{dx}\left(G_n + \int_a^x f_n\right)$$

$$= \frac{d}{dx}G_n + \frac{d}{dx}\int_a^x f_n$$

$$\geq f_n(x) \text{ a.e.} \quad \text{(using (i))}$$

Since n is arbitrary, making $n\to\infty$ we see that

$$F'(x) \geq f(x) \text{ a.e.}$$

Consequently,

$$\int_a^b F'(x)dx \geq \int_a^b f(x)dx$$

$$= F(b) - F(a) \quad \text{(using the hypothesis of the theorem)}$$

Also since F(x) is an increasing real valued function on the interval [a,b], we have

$$\int_a^b F'(x)dx \leq F(b) - F(a) = \int_a^b f(x)dx$$

Hence

$$\int_a^b F'(x)dx = F(b) - F(a) = \int_a^b f(x)dx$$

$$\Rightarrow \int_a^b [F'(x) - f(x)dx = 0$$

Since $F'(x) - f(x) \geq 0$, this implies that $F'(x) - f(x) = 0$ a.e. and so $F'(x) = f(x)$ a.e.

Absolute Continuity

Definition. A real-valued function f defined on [a,b] is said to be **absolutely continuous** on [a,b] if, given $\in > 0$ there is a $\delta > 0$ such that

$$\sum_{i=1}^{n} | f(x_i') - f(x_i)| < \in$$

for every finite collection $\{(x_i, x_i')\}$ of non-overlapping intervals with

$$\sum_{i=1}^{n} | x_i' - x_i| < \delta$$

An absolutely continuous function is continuous, since we can take the above sum to consist of one term only.

Moreover, if

$$F(x) = \int_a^x f(t)dt \ ,$$

then

$$\sum_{i=1}^{n} | F(x_i') - F(x_i)| = \sum_{i=1}^{n} \left| \int_a^{x_i'} f(t)dt - \int_a^{x_i} f(t)dt \right|$$

$$= \sum_{i=1}^{n} \left| \int_{x_i}^{x_i'} f(t)dt \right|$$

$$\leq \sum_{i=1}^{n} \int_{x_i}^{x_i'} |f(t)|\,dt = \int_E |f(t)|dt, \text{ where E is the set of intervals } (x, x_i')$$

$$\leq \to 0 \text{ as } \sum_{i=1}^{n} |x_i' - x_i| \to 0 .$$

The last step being the consequence of the result.

"Let $\in > 0$. Then there is a $\delta > 0$ such that for every measurable set $E \subset [a, b]$ with $m E < \delta$, we have

$$\int_A |f| < \in".$$

Hence every indefinite integral is absolutely continuous.

Lemma 6. If f is absolutely continuous on [a,b], then it is of bounded variation on [a,b].

Proof. Let δ be a positive real number which satisfies the condition in the definition for $\in = 1$. Select a natural number

$$n > \frac{b-a}{\delta}$$

Consider the partition $\pi = \{x_0, x_1, \ldots, x_n\}$ of [a,b] defined by

$$x_i = x_0 + \frac{i(b-a)}{n}$$

for every i = 0, 1,…, n. Since $|x_i - x_{i-1}| < \delta$, it follows that

$$V_{x_{i-1}}^{x_i}(f) < 1$$

This implies

$$V_a^b(f) = \sum_{i=1}^{n} V_{x_{i-1}}^{x_i}(f) < n$$

Hence f is of bounded variation.

Cor. If f is absolutely continuous, then f has a derivative almost everywhere.

Lemma 7. If f is absolutely continuous on [a,b] and $f'(x) = 0$ a.e., then f is constant.

Proof. We wish to show that f(a) = f(c) for any c ε [a,b].

Let $E \subset (a,c)$ be the set of measure $c-a$ in which $f'(x) = 0$, and let $\in$ and η be arbitrary positive numbers. To each x in E there is an arbitrarily small interval [x, x+h] contained in [a,c] such that

$$|f(x+h) - f(x)| < \eta h$$

By Vitali Lemma we can find a finite collection $\{[x_k, y_k]\}$ of non-overlapping intervals of this sort which cover all of E except for a set of measure less than δ, where δ is the positive number corresponding to $\in$ in the definition of the absolute continuity of f . If we label the x_k so that $x_k \leq x_{k+1}$, we have (or if we order these intervals so that)

$$a = y_0 \leq x_1 < y_1 \leq x_2 < \ldots . < y_n \leq x_{n+1} = c$$

and

$$\sum_{k=0}^{n} |x_{k+1} - y_k| < \delta$$

Now

$$\sum_{k=0}^{n} |f(y_k) - f(x_k)| \leq \eta \sum_{k=1}^{n} (y_k - x_k)$$

$$< \eta\ (c-a)$$

by the way to intervals $\{[x_k, y_k]\}$ were constructed, and

$$\sum_{k=0}^{n} | f(x_{k+1}) - f(y_k) | < \in$$

by the absolute continuity of f . Thus

$$|f(c) - f(a)| = | \sum_{k=0}^{n} [f(x_{k+1}) - f(y_k)] + \sum_{k=1}^{n} [f(y_k) - f(x_k)] |$$
$$\leq \in + \eta \, (c-a)$$

Since $\in$ and η are arbitrary positive numbers, f(c) − f(a) = 0 and so f(c) = f(a). Hence f is constant.

Theorem 4. A function F is an indefinite integral if and only if it is absolutely continuous.
Proof. We know that if F is an indefinite integral then F is absolutely continuous. Suppose on the other hand that F is absolutely continuous on [a,b]. Then F is of bounded variation and we may write

$$F(x) = F_1(x) - F_2(x),$$

where the functions F_i are monotone increasing. Hence $F'(x)$ exists almost everywhere and

$$|F'(x)| \leq F_1'(x) + F_2'(x)$$

Thus

$$\int |F'(x)| \, dx \leq F_1(b) + F_2(b) - F_1(a) - F_2(a)$$

and $F'(x)$ is integrable. Let

$$G(x) = \int_a^x F'(t) \, dt$$

Then G is absolutely continuous and so is the function f = F−G. But by the above lemma since $f'(x) = F'(x) - G'(x) = 0$ a.e., we have f to be a constant function. That is,

$$F(x) - G(x) = A \text{ (constant)}$$

or

$$F(x) = \int_a^x F'(t) \, dt = A$$

or

$$F(x) = \int_a^x F'(t) \, dt + A$$

Taking x = a, we have A = F(a) and so

$$F(x) = \int_a^x F'(t) \, dt + F(a)$$

Thus F(x) is indefinite integral of $F'(x)$.

Cor. Every absolutely continuous function is the indefinite integral of its derivative.

Convex Functions

Definition. A function ϕ defined an open interval (a, b) is said to be **convex** if for each x, y ε (a, b) and λ, μ such that $\lambda, \mu \geq 0$ and $\lambda + \mu = 1$, we have

$$\phi(\lambda x + \mu y) \leq \lambda\phi(x) + \mu\phi(y)$$

The end points a, b can take the values $-\infty, \infty$ respectively.

If we take $\mu = 1-\lambda$, $\lambda \geq 0$, then $\lambda + \mu = 1$ and so ϕ will be convex if

(5.1.1) $$\phi(\lambda x + (1-\lambda)y) \leq \lambda\phi(x) + (1-\lambda)\phi(y)$$

If we take a < s < t < u < b and

$$\lambda = \frac{t-s}{u-s} , \quad \mu = \frac{u-t}{u-s} , \quad u = x, \; s = y ,$$

then

$$\lambda + \mu = \frac{t-s+u-t}{u-s} = \frac{u-s}{u-s} = 1$$

and so (5.1.1) reduces to

$$\phi\left(\frac{t-s}{u-s}u + \frac{u-t}{u-s}s\right) \le \frac{t-s}{u-s}\phi(u) + \frac{u-t}{u-s}\phi(s)$$

or

(5.1.2) $$\phi(t) \le \frac{t-s}{u-s}\phi(u) + \frac{u-t}{u-s}\phi(u)$$

Thus the segment joining (s, ϕ(s)) and (u, ϕ(n)) is never below the graph of ϕ.
A function ϕ is sometimes said to be convex on (a,b) it for all x, y ε (a, b),

$$f\left(\frac{x+y}{2}\right) \le \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

(Clearly this definition is consequence of major definition taking $\lambda = \mu = \frac{1}{2}$).

If for all positive numbers λ, μ satisfying $\lambda + \mu = 1$, we have

$$\phi(\lambda x + \mu y) < \lambda\phi(x) + \mu\phi(y),$$

then ϕ is said to be **Strictly Convex.**

Theorem 5. Let ϕ be convex on (a,b) and a < s < t < u < b, then

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(u)-\phi(s)}{u-s} \le \frac{\phi(u-\phi(t)}{u-t}$$

If ϕ is strictly convex, equality will not occur.

Proof. Let a < s < t < u < b and suppose ϕ is convex on (a,b). Since

$$\frac{t-s}{u-s} + \frac{u-t}{u-s} = \frac{t-s+u-t}{u-s} = \frac{u-s}{u-s} = 1\ ,$$

therefore, convexity of ϕ yields

$$\phi\left(\frac{t-s}{u-s}u + \frac{u-t}{u-s}s\right) \le \frac{t-s}{u-s}\phi(u) + \frac{u-t}{u-s}\phi(s)$$

or

(5.1.3) $$\phi(t) \le \frac{t-s}{u-s}\phi(u) + \frac{u-t}{u-s}\phi(s)$$

or

$$(u-s)\ \phi(t) \le (t-s)\ \phi(u) + (u-t)\ \phi(s)$$

or

$$(u-s)\ (\phi(t) - \phi(s)) \le (t-s)\ \phi(u) + u\phi(s) - t\phi(s) - u\phi(s) + s\ \phi(s)$$

or

$$(u-s)(\phi(t) - \phi(s)) \le (t-s)\ (\phi(u)-\phi(s))$$

or

(5.1.4) $$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(u)-\phi(s)}{u-s}$$

This proves the first inequality. The second inequality can be proved similarly.
If ϕ is strictly converse, equality shall not be there in (5.1.3) and so it cannot be in (5.1.4). This completes the proof of the theorem.

Theorem 6. A differentiable function ϕ is convex on (a,b) if and only if ϕ' is a monotonically increasing function. If ϕ'' exists on (a,b), then ϕ is convex if and only if $\phi'' \ge 0$ on (a, b) and strictly convex if $\psi'' > 0$ on (a,b).

Proof. Suppose first that ϕ is differentiable and convex and let $a < s < t < u < v < b$. Then applying Theorem 5 to $a < s < t < u$, we get

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(u)-\phi(s)}{u-s} \le \frac{\phi(u)-\phi(t)}{u-t}$$

and applying Theorem 5 to $a < t < u < v$, we get

$$\frac{\phi(u)-\phi(t)}{u-t} \le \frac{\phi(v)-\phi(t)}{v-t} \le \frac{\phi(v)-\phi(v)}{v-u}$$

Hence

$$\frac{\phi(t)-\phi(s)}{t-s} \le \frac{\phi(v)-\phi(u)}{v-u}$$

If $t \to s$, $\frac{\phi(t)-\phi(s)}{t-s}$ decreases to $\phi'(s)$ and if $u \to v$, $\frac{\phi(v)-\phi(u)}{v-u}$ increases to $\phi'(v)$. Hence $\phi'(v) \ge \phi'(s)$ for all $s < v$ and so ϕ' is monotonically increasing function.
Further, if ϕ'' exists, it can never be negative due to monotonicity of ϕ'.
Conversely, let $\psi'' \ge 0$. Our aim is to show that ψ is convex. Suppose, on the contrary, that ϕ is not convex on (a, b). Therefore, there are points $a < s < t < u < b$ such that

$$\frac{\phi(t)-\phi(s)}{t-s} > \frac{\phi(u)-\phi(t)}{u-t}$$

that is, slope of chord over (s,t) is larger than the slope of the chord over (t,u). But slope of the chord over (s,t) is equal to $\phi'(\alpha)$, for some $\alpha \in (s, t)$ and slope of the chord over (t,u) is $\phi'(\beta)$, $\beta \in (t,u)$. But $\phi'(\alpha) > \phi'(\beta)$ implies ϕ' is not monotone increasing and so ψ'' cannot be greater than zero. We thus arrive at a contradiction. Hence ϕ is convex.
If $\phi'' > 0$, then ϕ is strictly convex, for otherwise there would exist collinear points of the graph of ϕ and we would have $\phi'(\alpha) = \phi'(\beta)$ for appropriate α and β with $\alpha < \beta$. But then $\phi'' = 0$ at some point between α and β which is a contradiction to $\phi'' > 0$. This completes the proof.

Theorem 7. If ϕ is convex on (a,b), then ϕ is absolutely continuous on each closed subinterval of (a,b).

Proof. Let $[c,d] \subset (a,b)$. If $x, y \in [c, d]$, then we have $a < c \le x \le y \le d < b$ and so by Theorem 5, we have

$$\frac{\phi(c)-\phi(a)}{c-a} \le \frac{\phi(y)-\phi(x)}{y-x} \le \frac{\phi(b)-\phi(d)}{b-d}$$

Thus

$$|\phi(y) - \phi(x)| \le M|x-y| \ , \ x, y \in [c, d]$$

and so ϕ is absolutely continuous there.

Theorem 8. Every convex function on an open interval is continuous.

Proof. If $a < x_1 < x < x_2 < b$, the convexity of a function ϕ implies

(5.1.5) $$\phi(x) \le \frac{x_2 - x}{x_2 - x_1}\phi(x_1) + \frac{x - x_1}{x_2 - x_1}\phi(x_2)$$

If we make $x \to x_1$ in (5.1.5), we obtain $\phi(x_1+ 0) \le \phi(x_1)$; and if we take $x_2 \to x$ we obtain $\phi(x) \le \phi(x + 0)$. Hence $\phi(x) = \phi(x+0)$ for all values of x in (a,b). Similarly $\phi(x-0) = \phi(x)$ for all values of x. Hence

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

and so ϕ is continuous.

Definition. Let ϕ be a convex function on (a,b) and $x_0 \in (a,b)$. The line

(5.1.6) $y = m(x-x_0) + \phi(x_0)$

through $(x_0, \phi(x_0))$ is called a **Supporting Line** at x_0 if it always lie below the graph of ϕ, that is, if

(5.1.7) $\phi(x) \ge m(x-x_0) + \phi(x_0)$

The line (5.1.6) is a supporting line if and only if its slope m lies between the left and right hand derivatives at x_0. Thus, in particular, there is at least one supporting line at each point.

Theorem 9 (**Jensen Inequality**). Let ϕ be a convex function on $(-\infty, \infty)$ and let f be an integrable function on [0,1]. Then

$\int\phi(f(t))dt \geq \phi[\int f(t)dt]$

Proof. Put

$$\alpha = \int_0^1 f(t)dt$$

Let $y = m(x-\alpha) + \phi(\alpha)$ be the equation of supporting line at α. Then (by (….) above),

$$\phi(f(t)) \geq m(f(t)-\alpha) + \phi(\alpha)$$

Integrating both sides with respect to t over [0, 1], we have

$$\int_0^1 \phi(f(t))dt \geq m[\int f(t)dt - \int f(t)dt] + \int_0^1 \phi(\alpha)dt$$

$$= 0 + \phi(\alpha)\int_0^1 dt$$

$$= \phi(\alpha) = \phi[\int_0^1 f(t)dt] .$$

L^p – space

Let p be a positive real number. A measurable function f defined on [0,1] is said to belong to the space L^p if $\int |f|^p < \infty$. Thus L^1 consists precisely of Lebesgue integrable functions on [0,1]. Since

$$|f+g|^p \leq 2^p (|f|^p + |g|^p) ,$$

we have

$$\int |f+g|^p \leq 2^p \int|f|^p + 2^p |f|^p$$

and so if f, g ε L^p, it follows that f+g ε L^p . Further, if α is a scalar and f ε L^p, then clearly αf belongs to L^p. Hence $\alpha f + \beta g$ ε L^p whenever f, g ε L^p and α, β are scalars.

We shall study these spaces in detail in Course On Functional Analysis.

PART B : MEASURE SPACE

5.2. We recall that a σ–algebra β is a family of subsets of a given set X which contains ϕ and is closed with respect to complements and with respect to countable unions. By a set function μ we mean a function which assigns an extended real number to certain sets. With this in mind we make the following definitions :

Definition. By a **measurable space** we mean a couple (X, β) consisting of a set X and a σ- algebra β of subsets of X.
A subset A of X is called measurable (or measurable with respect to B) if A ε β.

Definition. By a measure μ on a measurable space (X, β) we mean a non- negative set function defined for all sets of β and satisfying μ(ϕ) = 0 and

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu E_i \qquad (*)$$

for any sequence E_i of disjoint measurable sets.

By a **measure space** (X, β, μ) we mean a measurable space (X, β) together with a measure μ defined on β.

The property (*) of μ is referred to by saying that μ is countably additive.
An example of the **measure space** is (**R,** *m*, m) where **R** is the set of real numbers, *m* the Lebesgue measurable sets of real numbers and m the Lebesgue measure.

Theorem 10. If A ε β, B ε β, and A ⊂ B, then

$$\mu A \le \mu B$$

Proof. Since

$$B = A \cup [B \setminus A]$$

is a disjoint union, we have

$$\begin{aligned} \mu B &= \mu[A \cup (B \setminus A)] \\ &= \mu(A) + \mu(B \setminus A) \\ &\ge \mu A . \end{aligned}$$

Theorem 11. If E_i ε β, μ $E_1 < \infty$ and $E_i \supset E_{i+1}$, then

$$\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu E_n$$

Proof. Let

$$E = \bigcap_{i=1}^{\infty} E_i$$

Then

$$E_1 = E \cup \bigcup_{i=1}^{\infty} (E_i - E_{i+1}),$$

and this is a disjoint union. Hence

$$\mu(E_1) = \mu(E) + \sum_{i=1}^{\infty} \mu(E_i - E_{i+1})$$

Since

$$E_i = E_{i+1} \cup (E_i - E_{i+1})$$

is a disjoint union, we have

$$mE_i = mE_{i+1} + m(E_i \setminus E_{i+1})$$

i.e. $\mu(E_i - E_{i+1}) = \mu(E_i) - \mu E_{i+1}$

Hence

$$\mu(E_1) = \mu(E) + \sum_{i=1}^{\infty} (\mu E_i - \mu E_{i+1})$$

$$= \mu(E) + \lim_{n\to\infty} \sum_{i=1}^{n-1} (\mu E_i - \mu E_{i+1})$$

$$= \mu(E) + \mu E_1 - \lim_{n\to\infty} \mu E_n,$$

whence $\mu E_1 < \infty$ implies

$$\mu(E) = \lim_{n\to\infty} \mu E_n .$$

Theorem 12. If $E_i \varepsilon \beta$, then

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu E_i \quad .$$

Proof. Let

$$G_n = E_n - \bigcup_{i=1}^{n-1} E_i$$

Then $G_n \subset E_n$ and the sets G_n are disjoint. Hence

$$\mu(G_n) \leq \mu E_n,$$

while

$$\mu(\cup E_i) = \sum_{i=1}^{\infty} \mu G_n \leq \sum_{i=1}^{\infty} \mu E_n$$

Theorem 13. If $E_i \in \beta$ and $E_i \subset E_{i+1}$, then

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu E_n$$

Proof. Let

$$E = \bigcup_{i=1}^{\infty} E_i$$

Then

$$E = E_1 \cup (E_2 - E_1) \cup (E_3 - E_2) \cup \dots$$
$$= E_1 \cup \left\{\bigcup_{i=1}^{\infty} (E_{i+1} - E_i)\right\}$$

and this is a disjoint union. Hence

$$\mu(E) = \mu E_1 + \sum_{i=1}^{\infty} \mu(E_{i+1} \setminus E_i)$$
$$= \mu E_1 + \lim_{n\to\infty} \sum_{i=1}^{\infty} \mu\,(E_{i+1} - E_i)$$
$$= \mu E_1 + \lim_{n\to\infty} [\mu E_n - \mu E_1]$$
$$= \mu E_1 - \mu E_1 + \lim_{n\to\infty} \mu E_n$$
$$= \lim_{n\to\infty} \mu E_n .$$

Definition. A measure μ is called **finite** if $\mu(X) < \infty$. It is called **σ-finite** if there exists a sequence (X_n) of sets in β such that

$$X = \bigcup_{n=1}^{\infty} X_n$$

and $\mu\, X_n < \infty$.

By virtue of a lemma proved earlier in Chapter 3, we may always take $\{X_n\}$ to be a disjoint sequence of sets. Lebesgue measure on [0,1] is an example of a finite measure while Lebesgue measure on $(-\infty, \infty)$ is an example of a σ-finite measure.

Definition. A set E is said to be of finite measure if $E \in \beta$ and $\mu\, E < \infty$.

A set E is said to be of **σ-finite measure** if E is the union of a countable collection of measurable sets of finite measure.

Any measurable set contained in a set of σ-finite measure is itself of σ-finite measure, and the union of a countable collection of sets of σ-finite measure is again of σ-finite measure.

Definition. A measure space (X, β, μ) is said to be **complete** if β contains all subsets of sets of **measure zero**, that is, if $B \varepsilon \beta$, $\mu B = 0$ and $A \subset B$ imply $A \varepsilon \beta$.
For example Lebesgue measure is complete, while Lebesgue measure restricted to the σ-algebra of Borel sets is not complete.

Definition. If (X, β, μ) is a measure space, we say that a subset E of X is **locally measurable** if $E \cap B \varepsilon \beta$ for each $B \varepsilon \beta$ with $\mu B < \infty$.

The collection *C* of all locally measurable sets is a σ-algebra containing β.
The measure μ is called **saturated** if every locally measurable set is measurable, i.e., is in β.
For example every σ-finite measure is saturated.

Example. Show that $\mu(E_1 \Delta E_2) = 0$ implies $\mu E_1 = \mu E_2$ provided that

E_1 and $E_2 \varepsilon \beta$.

Solution. Since $E_1, E_2 \varepsilon \beta$, we have $E_1 \setminus E_2$ and $E_2 \setminus E_1$ in β and so $E_1 \Delta E_2 \varepsilon \beta$. Moreover,

$$\begin{aligned}\mu(E_1 \Delta E_2) &= \mu [(E_1 \setminus E_2) \cup (E_2 \setminus E_1)] \\ &= \mu (E_1 \setminus E_2) + \mu(E_2 \setminus E_1)\end{aligned}$$

But, by hypothesis, $\mu (E_1 \Delta E_2) = 0$. Therefore,

$$\mu(E_1 \setminus E_2) = 0 \text{ and } \mu(E_2 \setminus E_1) = 0 .$$

Also, we can write

$$E_2 = [E_1 \cup (E_2 - E_1)] - (E_1 - E_2)$$

Then

$$\mu E_2 = \mu E_1 + 0 - 0 = \mu E_1 .$$

5.3. Measure and Outer Measure

In case of Lebesgue measure we defined measure for open sets and used this to define outer measure, from which we obtain the notion of measurable set and Lebesgue measure.

Definition. By an outer measure μ^* we mean an extended real valued set function defined on all subsets of a space X and having the following properties :

(i) $\mu^* \phi = 0$

(ii) $A \subset B \Rightarrow \mu^*A \leq \mu^*B$ (monotonicity)

(iii) $E \subset \sum_{i=1}^{\infty} E_i \Rightarrow \mu^* E \leq \sum_{i=1}^{\infty} \mu^* E_i$ (subadditivity)

Because of (ii), property (iii) can be replaced by

(iii)′ $E = \bigcup_{i=1}^{\infty} E_i$, E_i disjoint $\Rightarrow \mu^* E \leq \sum_{i=1}^{\infty} \mu^* E_i$

The outer measure μ^* is called finite if $\mu^* X < \infty$.

By analogy with the case of Lebesgue measure we define a set E to be **measurable** with respect to μ^* if for every set A we have

$$\mu^*A = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Since μ^* is subadditive, it is only necessary to show that

$$\mu^* A \geq \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

for every A in order to show that E is measurable.

This inequality is trivially true when $\mu^* A = \infty$ and so we need only establish it for sets A with μ^*A finite.

Theorem 14. The class β of μ^*-measurable sets is a σ-algebra. If $\bar{\mu}$ is restricted to β, then $\bar{\mu}$ is a complete measure on β.

Proof. It is obvious that the empty set is measurable. The symmetry of the definition of measurability in E and E^c shows that E^c is measurable whenever E is measurable.

Let E_1 and E_2 be measurable sets. From the measurability of E_2 ,

$$\mu^*A = \mu^*(A \cap E_2) + \mu^*(A \cap E_2^c)$$

and

$$\mu^*A = \mu^*(A \cap E_2) + \mu^*(A \cap E_2^c \cap E_1) + \mu^*(A \cap E_1^c \cap E_2^c)$$

by the measurability of E_1. Since

$$A \cap [E_1 \cup E_2] = [A \cap E_2] \cup [A \cap E_1 \cap E_2^c]$$

we have

$$\mu^*(A \cap [E_1 \cup E_2]) \le \mu^*(A \cap E_2) + \mu^*(A \cap E_2^c \cap E_1)$$

by subadditivity, and so

$$\mu^*A \ge \mu^*(A \cap [E_1 \cup E_2]) + \mu^*(A \cap E_1^c \cap E_2^c)$$

This means that $E_1 \cup E_1$ is measurable. Thus the union of two measurable sets is measurable, and by induction the union of any finite number of measurable sets is measurable, showing that β is an algebra of sets.

Assume that $E = \cup E_i$, where $< E_i >$ is a disjoint sequence of measurable set, and set

$$G_n = \bigcup_{i=1}^{n} E_i$$

Then (by what we have proved above) G_n is measurable, and

$$\mu^* A = \mu^*(A \cap G_n) + \mu^*(A \cap G_n^c) \ge \mu^*(A \cap G_n) + \mu^*(A \cap E^c)$$

because $E^c \subset G_n^c$

Now $G_n \cap E_n = E_n$ and $G_n \cap E_n^c = G_{n-1}$, and by the measurability of E_n, we have

$$\mu^*(A \cap G_n) = \mu^*(A \cap E_n) + \mu^*(A \cap G_{n-1})$$

By induction (as above, $\mu^*(A \cap G_{n-1}) = \mu^*(A \cap E_{n+1}) + \mu^*(A \cap E_{n-2}$ and so on)

$$\mu^*(A \cap G_n) = \sum_{i=1}^{n} \mu^*(A \cap E_i)$$

and so

$$\mu^* A \ge \mu^*(A \cap E^c) + \sum_{i=1}^{\infty} \mu^*(A \cap E_i)$$

$$\ge \mu^*(A \cap E^c) + \mu^*(A \cap E),$$

since $\quad A \cap E \subset \bigcup_{i=1}^{\infty}(A \cap E_i)$

Thus E is measurable. Since the union of any sequence of sets in an algebra can be replaced by a disjoint union of sets in an algebra, it follows that B is a σ-algebra.

We now show that $\bar{\mu}$ is finitely additive. Let E_1 and E_2 be disjoint measurable sets. Then the measurability of E_2 implies that

$$\begin{aligned}\bar{\mu}(E_1 \cup E_2) &= \mu^*(E_1 \cup E_2) \\ &= \mu^*([E_1 \cup E_2] \cap E_2) + \mu^*([E_1 \cup E_2] \cap E_2^c) \\ &= \mu^* E_2 + \mu^* E_1\end{aligned}$$

Finite additivity now follows by induction.

If E is the disjoint union of the measurable sets $\{E_i\}$, then

$$\bar{\mu}_i E \geq \bar{\mu}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \bar{\mu} E_i$$

and so

$$\bar{\mu} E \geq \sum_{i=1}^{\infty} \bar{\mu} E_i$$

But $\bar{\mu} E \leq \sum_{i=1}^{\infty} \bar{\mu} E_i$, by the subadditivity of μ^*. Hence $\bar{\mu}$ is **countably additive** and hence a measure since it is non-negative and $\bar{\mu}\phi = \mu^*\phi = 0$.

Measure on an Algebra

By a **measure on an algebra** we mean a non-negative extended real valued set function μ defined on an algebra **A** of sets such that

(i) $\mu \phi = 0$

(ii) If $< A_i >$ is a disjoint sequence of sets in **A whose union is also in A**, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu A_i$$

Thus a measure on an algebra **A** is a measure $\Leftrightarrow$ **A** is a σ-algebra.

We construct an outer measure μ^* and show that the measure $\bar{\mu}$ induced by μ^* is an extension of μ (measure defined on an algebra).

We define

$$\mu^* E = \inf \sum_{i=1}^{\infty} \mu A_i \ ,$$

where $< A_i >$ ranges over all sequence from **A** such that

$$E \subset \bigcup_{i=1}^{\infty} A_i \ .$$

Lemma 8. If A ε **A** and if $< A_i >$ is any sequence of sets in **A** such that $A \subset \bigcup_{i=1}^{\infty} A_i$, then $\mu A \leq \sum_{i=1}^{\infty} \mu A_i$.

Proof. Set

$$B_n = A \cap A_n \cap A^c_{n-1} \cap \ \ldots . \cap A_i^c$$

Then B_n ε **A** and $B_n \subset A_n$. But A is the disjoint union of the sequence $< B_n >$ and so by countable additivity

$$\mu A = \sum_{n=1}^{\infty} \mu B_n \leq \sum_{n=1}^{\infty} \mu A_n$$

Corollary. If A ε **A**, $\mu^* A = \mu A$.

In fact, we have, from above

$$\mu A \leq \sum_{n=1}^{\infty} \mu A_n < \mu^* A + \in \ ,$$

that is,

$$\mu A \leq \mu^* A + \in$$

Since $\in$ is arbitrary, we have

$$\mu A \leq \mu^* A$$

Also, by definition,

$$\mu^* A \leq \mu A$$

Hence

$$\mu^* A = \mu A \ .$$

Lemma 9. The set function μ^* is an outer measure.

Proof. μ^*, by definition, is a monotone non-negative set function defined for all sets and $\mu^* \phi = O$. We have only to show that it is countably subadditive. Let $E \subseteq \bigcup_{i=1}^{\infty} E_i$. If $\mu^* E_i = \infty$ for any i, we have $\mu^* E \leq \Sigma \mu^* E_i = \infty$. If not, given $\in > 0$, there is for each i a sequence $< A_{ij} >_{j=1}^{\infty}$ of sets in **A** such that $E_i \subset \bigcup_{j=1}^{\infty} A_{ij}$ and

$$\sum_{j=1}^{\infty} \mu A_{ij} < \mu^* E_i + \frac{\in}{2^i}$$

Then

$$\mu^* E \leq \sum_{i,j} \mu A_{ij} < \sum_{i=1}^{\infty} \mu^* E_i + \in$$

Since $\in$ was an arbitrary positive number,

$$\mu^* E \leq \sum_{i=1}^{\infty} \mu^* E_i$$

which proves that μ^* is subadditive.

Lemma 10. If A ε **A**, then A is measurable with respect to μ^*.

Proof. Let E be an arbitrary set of finite outer measure and ε a positive number. Then there is a sequence $< A_i >$ from **A** such that $E \subset \cup A_i$ and

$$\Sigma \mu A_i < \mu^* E + \in$$

By the additivity of μ on **A**, we have

$$\mu(A_i) = \mu(A_i \cap A) + \mu(A_i \cap A^c)$$

Hence

$$\mu^* E + \in > \sum_{i=1}^{\infty} \mu(A_i \cap A) + \sum_{i=1}^{\infty} \mu(A_i \cap A^c)$$

$$> \mu^* (E \cap A) + \mu^* (E \cap A^c)$$

because

$$E \cap A \subset \cup (A_i \cap A)$$

and

$$E \cap A^c \subset \cup (A_i \cap A^c)$$

Since $\in$ was an arbitrary positive number,

$$\mu^* E \geq \mu^* (\in \cap A) + \mu^* (E \cap A^c)$$

and thus A is μ^* - measurable.

Remark. The outer measure μ^* which we have defined above is called the **outer measure induced by μ**.

Notation. For a given algebra $\mathbf{A}$ of sets we use $\mathbf{A}_\sigma$ to denote those sets which are countable unions of sets of $\mathbf{A}$ and use $\mathbf{A}_{\sigma\delta}$ to denote these sets which are countable intersection of sets in $\mathbf{A}_\sigma$**.**

Theorem 15. Let μ be a measure on an algebra $\mathbf{A}$, μ^* the outer measure induced by μ, and E any set. Then for $\varepsilon > 0$, there is a set A ε $\mathbf{A}_\sigma$ with $E \subset A$ and

$$\mu^* A \le \mu^* E + \in$$

There is also a set B ε $\mathbf{A}_{\sigma\delta}$ with $E \subset B$ and $\mu^* E = \mu^* B$.

Proof. By the definition of μ^* there is a sequence $< A_i >$ from $\mathbf{A}$ such that $E \subset \cup A_i$ and

$$\sum_{i=1}^{\infty} \mu A_i \le \mu^* E + \in \qquad (1)$$

Set

$$A = \cup A_i$$

Then $\mu^* A \le \Sigma \mu^* A_i = \Sigma \mu A_i$ (2)

because μ^* and μ agree on members of $\mathbf{A}$ by the corollary.
Hence (1) and (2) imply

$$\mu^* A \le \mu^* E + \in$$

which proves the first part.

To prove the second statement, we note that for **each positive integer n there is a set** A_n in $\mathbf{A}_\sigma$ such that $E \subset A_n$ and

$$\mu^* A_n < \mu^* E + \frac{1}{n} \quad \text{(from first part proved above)}$$

Let $B = \cap A_n$. Then B ε $\mathbf{A}_{\sigma\delta}$ and $E \subset B$. Since $B \subset A_n$,

$$\mu^* B < \mu^* A_n \le \mu^* E + \frac{1}{n}$$

Since n is arbitrary, $\mu^* B \le \mu^* E$. But $E \subset B$ implies $\mu^* B \ge \mu^* E$ by monotonicity. Hence $\mu^* B = \mu^* E$.

Theorem 16 (**Cartheodory Extension Theorem**). Let μ be a measure on an algebra **A** and let μ^* be the outer measure induced by μ. Then the following properties hold :

(a) μ^* is an outer measure

(b) $A \subset$ **A** implies $\mu(A) = \mu^* A$

(c) $A \subset$ **A** implies A is μ^* - measurable.

(d) The restriction $\bar{\mu}$ of μ^* to the μ^*-measurable sets is an extension of μ to a σ-algebra containing **A.**

(e) μ is finite (or σ-finite) implies that $\bar{\mu}$ is finite (or σ-finite).

(f) If μ is σ-finite, then $\bar{\mu}$ is the only measure on the smallest σ-algebra containing **A** which is an extension of μ.

Proof. We have already proved (a), (b) and (c). The fact that $\bar{\mu}$ is an extension of μ from **A** to be a measure on a σ-algebra containing **A** follows from (b), (c) and the result. "The class β of μ^*-measure sets is a σ-algebra. If $\bar{\mu}$ is μ^* restricted to β, then $\bar{\mu}$ is a measure on β." Also $\bar{\mu}$ is finite or σ-finite whenever μ is finite or σ-finite.
We establish (f) now.

Let β be the smallest σ-algebra containing **A** and μ_1 be another measure on β such that $\mu_1(E) = \mu(E)$ for $E \in$ **A**. We need to show the following :

$$\mu_1(A) = \bar{\mu}(A) \quad \text{for } A \in \beta \tag{1}$$

Since μ is σ-finite, we can write $X = \bigcup_{i=1}^{\infty} E_i$, $E_i \in$ **A**, $E_i \cap E_j = \phi (i \neq j)$ and $\mu(E_i) < \infty$, $1 \leq i < \infty$. For $A \in \beta$, then $A = \cup (A \cap E_i)$ and we have

$$\bar{\mu}(A) = \sum_{i=1}^{\infty} \bar{\mu}(A \cap E_i)$$

and

$$\mu_1(A) = \sum_{i=1}^{\infty} \mu_1 (A \cap E_i)$$

So, to prove (1) it is sufficient to show the following :

$$\mu_1 A = \bar{\mu}(A) \text{ for } A \in \beta \text{ whenever } \bar{\mu}(A) < \infty$$

Let $A \varepsilon \beta$ with $\bar{\mu} A < \infty$. Given $\in > 0$, there are $E_i \varepsilon$ **A**, $1 \leq i < \infty$, $A \subset \bigcup_{i=1}^{\infty} E_i$ and

$$\bar{\mu}\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu(E_i) < \bar{\mu}(A) + \in$$

Since

$$\mu_1(A) \leq \mu_1\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu_1(E_i) = \sum_{i=1}^{\infty} \mu(E_i)$$

Thus, (2) implies

$$\mu_1(A) \leq \bar{\mu}(A) + \in \qquad (2)$$

Since this is true for all > 0, it follows that

$$\mu_1(A) \leq \bar{\mu} A \qquad (3)$$

Now considering the sets E_i from inequality (2), $F = \bigcup_{i=1}^{\infty} E_i \ \varepsilon \beta$ and so F is μ^*-measurable. Since $A \subset F$,

$$\bar{\mu}(F) = \bar{\mu}(A) + \bar{\mu}(F - A)$$

or

$$\bar{\mu}(F \setminus A) = \bar{\mu}(F) - \bar{\mu}(A) < \varepsilon \quad \text{(from (2))}$$

Since $\mu_1(E) = \bar{\mu}(E)$ for each $E \varepsilon$ **A**, we have $\mu_1(F) = \bar{\mu}(F)$. Then

$$\bar{\mu}(A) \leq \bar{\mu}(F) = \mu_1(F) = \mu_1(A) + \mu_1(F \setminus A)$$
$$\leq \mu(A) + \bar{\mu}(F \setminus A)$$

(by inequality (3) because (3) is true if A is replaced by any set in β with finite $\bar{\mu}$-measure). The relation (4) then yields

$$\bar{\mu}(A) \leq \mu_1(A) + \in$$

Since this true for all $\in > 0$, we have

$$\bar{\mu}(A) \leq \mu_1(A) \qquad (5)$$

The relations (3) and (5) then yield

$$\bar{\mu} A = \mu_1(A)$$

which completes the proof of the theorem.

Definition. Let f be a non-negative extended real valued measurable function on the measure space (X, β, μ). Then $\int f \, d\mu$ is the supremum of the integrals $\int\phi\ d\mu$ as ϕ ranges over all simple functions with $0 \leq \phi \leq f$.

Lemma 1 (Fatou's Lemma). Let $< f_n >$ be a sequence of non-negative measurable functions which converge almost everywhere on a set E to a function f. Then

$$\int_E f \leq \underline{\lim} \int_E f_n$$

Proof. Without loss of generality we may assume that $f_n(x) \to f(x)$ for each $x \in E$. From the definition of $\int f$ it is sufficient to show that , if ϕ is any non-negative simple function with $\phi \leq f$, then $\int_E \phi \geq \underline{\lim} \int f_n$

If $\int_E \phi = \infty$, then there is a measurable set $A \subset E$ with $\mu A = \infty$ such that $\phi >$ a 0 on A. Set

$$A_n = \{x \in E :\ f_k(x) > a \text{ for all } k \geq n\}$$

Then $A_n \subseteq A_{n+1}$. Thus $< A_n >$ is an increasing sequence of measurable sets whose union contains A, since $\phi \leq \lim f_n$. Thus $\lim \mu A_n = \infty$. Since $\int_E f_n \geq a\,\mu A_n$, we have

$$\lim \int_E f_n = \infty = \int_E \phi$$

If $\int_E \phi < \infty$, then there is a measurable set $A \subset E$ with $\mu A < \infty$ such that ϕ vanishes identically on $E \setminus A$. Let M be the maximum of ϕ, let $\in$ be a given positive number, and set

$$A_n = [x \in E \quad \text{if } f_h(x) > (1-\in)\,\phi(x) \text{ for all } k \geq n]$$

Then $< A_n >$ is an increasing sequence of sets whose union contains A, and so $(A \setminus A_n)$ is a decreasing sequence of sets whose intersection is empty. Therefore, (by a proposition proved already) $\lim \mu\,(A — A_n) = 0$ and so we can find an n such that $\mu(A — A_k) < \in$ for all $k \geq n$. Thus for $k \geq n$

$$\int_E f_k \geq \int_{A_k} f_k \geq (1-\in) \int_{A_k} \phi$$

$$\geq (1-\in) \left[\int_A \phi - \int_{A-A_k} \phi \right]$$

$$\geq (1-\in) \int \phi - \int_{A-A_k} \phi$$

$$\geq \int \phi - \in \int \phi - \in M$$

Since c is arbitrary.

$$\underline{\lim} \int f_k \geq \int \phi \quad .$$