MAL-512: M. Sc. Mathematics (Real Analysis)

Lesson No. 1 **Written by Dr. Nawneet Hooda**

Lesson: Sequences and Series of Functions -1 **Vetted by Dr. Pankaj Kumar**

Consider sequences and series whose terms depend on a variable, i.e., those whose terms are real valued functions defined on an interval as domain. The sequences and series are denoted by $\{f_n\}$ and $\sum f_n$ respectively.

Point-wise Convergence

Definition. Let $\{f_n\}$, n = 1, 2, 3,…be a sequence of functions, defined on an interval I, $a \le x \le b$. If there exits a real valued function f with domain I such that

$$f(x) = \lim_{n\to\infty} \{f_n(x)\}, \quad \forall\ x \in I$$

Then the function f is called the limit or the point-wise limit of the sequence $\{f_n\}$ on [a, b], and the sequence $\{f_n\}$ is said to be point-wise convergent to f on [a, b]. Similarly, if the series $\sum f_n$ converges for every point $x \in I$, and we define

$$f(x) = \sum_{n=0}^{\infty} f_n(x), \quad \forall\ x \in [a, b]$$

the function f is called the sum or the point-wise sum of the series $\sum f_n$ on [a, b].

Definition. If a sequence of functions $\{f_n\}$ defined on [a, b], converges poinwise to f , then to each $\in > 0$ and to each $x \in [a, b]$, there corresponds an integer N such that

$$|f_n(x) - f(x)| < \in, \quad \forall\ n \ge N \qquad (1.1)$$

Remark:

1. The limit of differentials may not equal to the differential of the limit.

Consider the sequence $\{f_n\}$, where $f_n(x) = \dfrac{\sin nx}{\sqrt{n}}$, (x real).

It has the limit

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

$\therefore$ $\quad f'(x) = 0$, and so $f'(0) = 0$

But

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

so that

$$f_n'(0) = \sqrt{n} \to \infty \text{ as } n \to \infty$$

Thus at x = 0, the sequence $\{f'_n(x)\}$ diverges whereas the limit function $f'(x) = 0$,

2. Each term of the series may be continuous but the sum f may not.

Consider the series

$$\sum_{n=0}^{\infty} f_n \text{ , where } f_n(x) = \frac{x^2}{(1+x^2)^n} \text{ (x real)}$$

At x = 0, each $f_n(x) = 0$, so that the sum of the series $f(0) = 0$.

For $x \neq 0$, it forms a geometric series with common ratio $1/(1 + x^2)$, so that its sum function $f(x) = 1 + x^2$.

Hence,

$$f(x) = \begin{cases} 1+x^2, & x \neq 0 \\ 0 & , x = 0 \end{cases}$$

3. The limit of integrals is not equal to the integral of the limit.

Consider the sequence $\{f_n\}$, where

$$f_n(x) = nx(1 - x^2)^n, \; 0 \le x \le 1, \; n = 1, 2, 3, \ldots$$

For $0 < x \le 1$, $\lim\limits_{n \to \infty} f_n(x) = 0$

At x = 0, each $f_n(0) = 0$, so that $\lim\limits_{n \to \infty} f_n(0) = 0$

Thus the limit function $f(x) = \lim\limits_{n \to \infty} f_n(x) = 0$, for $0 \le x \le 1$

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

Again,

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

so that

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

Thus,

$$\lim_{n\to\infty}\left\{\int_0^1 f_n\,dx\right\} \neq \int_0^1 f\,dx = \int_0^1 \left[\lim_{n\to\infty}\{f_n\}\right] dx$$

Uniform Convergence

Definition. A sequence of functions $\{f_n\}$ is said to converge uniformly on an interval [a, b] to a function f if for any $\varepsilon > 0$ and for all $x \in [a, b]$ there exists an integer N (independent of x but dependent on ε) such that for all $x \in [a, b]$

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall\, n \geq N \qquad (1)$$

Remark. Every uniformly convergent sequence is pointwise convergent, and the uniform limit function is same as the pointwise limit function. But the converse is not true. However non-pointwise convergence implies non-uniform convergence .

Definition. A series of functions $\sum f_n$ is said to converge uniformly on [a, b] if the sequence $\{S_n\}$ of its partial sums, defined by

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

converges uniformly on [a, b].

Definition. A series of functions $\sum f_n$ converges uniformly to f on [a, b] if for $\in > 0$ and all $x \in [a, b]$ there exists an integer N (independent of x and dependent on ε) such that for all x in [a, b]

$$|f_1(x) + f_2(x) + \ldots + f_n(x) - f(x)| < \varepsilon, \text{ for } n \geq N$$

Theorem (Cauchy's Criterion for Uniform Convergence). The sequence of functions $\{f_n\}$ defined on [a, b] converges uniformly on [a, b] if and only if for every $\varepsilon > 0$ and for all $x \in [a, b]$, there exists an integer N such that

$$|f_{n+p}(x) - f_n(x)| < \varepsilon, \quad \forall\, n \geq N, p \geq 1 \qquad \ldots(1)$$

Proof. Let the sequence $\{f_n\}$ uniformly converge on [a, b] to the limit function f, so that for a given $\varepsilon > 0$, and for all $x \in [a, b]$, there exist integers n_1, n_2 such that

$$|f_n(x) - f(x)| < \varepsilon/2, \quad \forall\, n \geq n_1$$

and

$$|f_{n+p}(x) - f(x)| < \varepsilon/2, \quad \forall\, n \geq n_2, p \geq 1$$

Let $N = \max(n_1, n_2)$.

$$\therefore \quad |f_{n+p}(x) - f_n(x)| \leq |f_{n+p}(x) - f(x)| + |f_n(x) - f(x)|$$

$$< \varepsilon/2 + \varepsilon/2 = \varepsilon, \qquad \forall\, n \geq N, p \geq 1$$

Conversely. Let the given condition hold so by Cauchy's general principle of convergence, $\{f_n\}$ converges for each $x \in [a, b]$ to a limit, say f and so the sequence converges pointwise to f.

For a given $\varepsilon > 0$, let us choose an integer N such that (1) holds. Fix n, and let $p \to \infty$ in (1). Since $f_{n+p} \to f$ as $p \to \infty$, we get

$$|f(x) - f_n(x)| < \varepsilon \qquad \forall\, n \geq N, \text{ all } x \in [a, b]$$

which proves that $f_n(x) \to f(x)$ uniformly on [a, b].

Remark. Other form of this theorem is :

The sequence of functions $\{f_n\}$ defined on [a, b] converges uniformly on [a, b] if and only if for every $\varepsilon > 0$ and for all $x \in [a, b]$, there exists an integer N such that

$$|f_n(x) - f_m(x)| < \varepsilon, \qquad \forall\, n, m \geq N$$

Theorem 2. A series of functions $\sum f_n$ defined on [a, b] converges uniformly on [a, b] if and only if for every $\varepsilon > 0$ and for all $x \in [a, b]$, there exists an integer N such that

$$|f_{n+1}(x) + f_{n+2}(x) + \ldots + f_{n+p}(x)| < \varepsilon, \quad \forall\, n \geq N, p \geq 1 \qquad \ldots(2)$$

Proof. Taking the sequence $\{S_n\}$ of partial sums of functions $\sum f_n$, defined by

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

And applying above theorem, we get the result.

Example . Show that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{nx}{1+n^2x^2}, \text{ for } x \in [a, b] .$$

is not uniformly convergent on any interval [a, b] containing 0.

Solution. The sequence converges pointwise to f, where $f(x) = 0$, $\forall$ real x.

Let $\{f_n\}$ converge uniformly in any interval [a, b], so that the pointwise limit is also the uniform limit. Therefore for given $\varepsilon>0$, there exists an integer N such that for all $x \in [a, b]$, we have

$$\left| \frac{nx}{1+n^2x^2} - 0 \right| < \varepsilon, \quad \forall\ n \geq N$$

If we take $\varepsilon = \frac{1}{3}$, and t an integer greater than N such that $1/t \in [a, b]$, we find on taking $n = t$ and $x = 1/t$, that

$$\frac{nx}{1+n^2x^2} = \frac{1}{2} \not< \frac{1}{3} = \in.$$

which is a contradiction and so the sequence is not uniformly convergent in the interval [a, b], having the point 1/t. But since $1/t \to 0$, the interval [a, b] contains 0. Hence the sequence is not uniformly convergent on any interval [a, b] containing 0.

Example . The sequence $\{f_n\}$, where

$$f_n(x) = x^n$$

is uniformly convergent on [0, k], $k < 1$ and only pointwise convergent on [0, 1].

Solution.

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

Thus the sequence converges pointwise to a discontinuous function on [0, 1]

Let $\varepsilon > 0$ be given.

For $0 < x \le k < 1$, we have

$|f_n(x) - f(x)| = x^n < \varepsilon$

if

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

or if

$$n > \log (1/\varepsilon)/\log(1/x)$$

This number, $\log (1/\varepsilon)/\log (1/x)$ increases with x, its maximum value being $\log (1/\varepsilon)/\log(1/k)$ in $]0, k]$, $k > 0$.

Let N be an integer $\ge \log (1/\varepsilon)/\log(1/k)$.

$\therefore \quad |f_n(x) - f(x)| < \varepsilon, \quad \forall\, n \ge N, 0 < x < 1$

Again at $x = 0$,

$$|f_n(x) - f(x)| = 0 < \varepsilon, \qquad \forall \quad n \ge 1$$

Thus for any $\varepsilon > 0$, $\exists$ N such that for all $x \in [0, k]$, $k < 1$

$$|f_n(x) - f(x)| < \varepsilon, \qquad \forall\, n \ge N$$

Therefore, the sequence $\{f_n\}$ is uniformly convergent in $[0, k]$, $k < 1$.

However, the number $\log (1/\varepsilon)/\log (1/x) \to \infty$ as $x \to 1$ so that it is not possible to find an integer N such that $|f_n(x) - f(x)| < \varepsilon$, for all $n \ge N$ and all x in $[0, 1]$. Hence the sequence is not uniformly convergent on any interval containing 1 and in particular on $[0, 1]$.

Example . Show that the sequence $\{f_n\}$, where

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

is uniformly convergent in any interval $[0, b]$, $b > 0$.

Solution. The limit function is

$$f(x) = \lim_{n\to} f_n(x) = 0 \qquad \forall\, x \in [0, b]$$

so that the sequence converges pointwise to 0.

For any $\varepsilon > 0$,

$$|f_n(x) - f(x)| = \frac{1}{x+n} < \varepsilon$$

if $n > (1/\varepsilon) - x$, which decreases with x, the maximum value being $1/\varepsilon$.

Let N be an integer $\geq 1/\varepsilon$, so that for $\varepsilon > 0$, there exists N such that

$$|f_n(x) - f(x)| < \varepsilon, \qquad \forall\, n \geq N$$

Hence the sequence is uniformly convergent in any interval [0, b], b > 0.

Example . The series Σf_n, whose sum to n terms, $S_n(x) = nxe^{-nx^2}$, is pointwise and not uniformly convergent on any interval [0, k], k > 0.

Solution. The pointwise sum $S(x) = \lim_{n\to\infty} S_n(x) = 0$, for all $x \geq 0$. Thus the series converges pointwise to 0 on [0, k].

Let us suppose, if possible, the series converges uniformly on [0, k], so that for any $\varepsilon > 0$, there exists an integer N such that for all $x \geq 0$,

$$|S_n(x) - S(x)| = nxe^{-nx^2} < \in, \qquad \forall\, n \geq N \qquad \ldots(*)$$

Let N_0 be an integer greater than N and $e^2\varepsilon^2$, then for $x = 1/\sqrt{N_0}$ and $n = N_0$, (*) gives

$$\sqrt{N_0}/e < \varepsilon \;\Rightarrow\; N_0 < e^2\varepsilon^2$$

so we arrive at a contradiction. Hence the series is not uniformly convergent on [0, k].

Note **.** The interval of uniform convergence is always to be a closed interval, that is it must include the end points. But the interval for pointwise or absolute convergence can be of any type.

Theorem 3. Let $\{f_n\}$ be a sequence of functions, such that

$$\lim_{n\to\infty} f_n(x) = f(x),\ x \in [a, b]$$

and let

$$M_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|$$

Then $f_n \to f$ uniformly on [a, b] if and only if $M_n \to 0$ as $n \to \infty$.

Proof. Let $f_n \to f$ uniformly on [a, b], so that for a given $\varepsilon>0$, there exists an integer N such that

$$|f_n(x) - f(x)| < \varepsilon, \qquad \forall\, n \geq N, \ \forall\, x \in [a, b]$$

$$\Rightarrow \qquad M_n = \sup_{x\in[a,b]} |f_n(x) - f(x)| \leq \varepsilon, \qquad \forall\, n \geq N$$

$$\Rightarrow \qquad M_n \to 0, \text{ as } n \to \infty$$

Conversely. Let $M_n \to 0$, as $n \to \infty$, so that for any $\varepsilon > 0$, $\exists$ an integer N such that

$$M_n < \varepsilon, \qquad \forall\, n \geq N$$

$$\Rightarrow \qquad \sup_{x\in[a,b]} |f_n(x) - f(x)| < \varepsilon, \qquad \forall\, n \geq N$$

$$\Rightarrow \qquad |f_n(x) - f(x)| < \varepsilon, \qquad \forall\, n \geq N, \ \forall\, x \in [a, b]$$

$$\Rightarrow \qquad f_n \to f \text{ uniformly on } [a, b].$$

Example. Show that 0 is a point of non-uniform convergence of the sequence $\{f_n\}$, where $f_n(x) = 1-(1 - x^2)^n$.

Solution. We have

$$M_n = \sup \{|f_n(x) - f(x)| : x \in]0, \sqrt{}[\}$$

$$= \sup \{(1-x^2)^n : x \in]\, 0, \sqrt{2}[\}$$

$$\geq \left(1-\frac{1}{n}\right)^n \qquad \left[\text{Taking } x = \frac{1}{\sqrt{n}} \in]0, \sqrt{2}\right]$$

$$\to \frac{1}{e} \text{ as } n\to\infty.$$

Thus M_n cannot tend to zero as $n\to\infty$.

It follows that the sequence is non-uniformly convergent.

Also as $n\to\infty$, $x\to 0$ and consequently 0 is a point of non-uniform convergence.

Example . Prove that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{x}{1+nx^2}, \ x \text{ real}$$

converges uniformly on any closed interval I.

Here pointwise limit,

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

$$M_n = \sup_{x\in I} |f_n(x) - f(x)| = \sup_{x\in I} \left|\frac{x}{1+nx^2}\right|$$

$$= \frac{1}{2\sqrt{n}} \to 0 \quad \text{as} \quad n \to \infty$$

Hence $\{f_n\}$ converges uniformly on I.

$\left[\dfrac{x}{1+nx^2}\right.$ attains the maximum value $\dfrac{1}{2\sqrt{n}}$ at $x = \dfrac{1}{\sqrt{n}}$, i.e. at the origin.$\left.\right]$

Example . Show that the sequence $\{f_n\}$, where

$$f_n(x) = nxe^{-nx^2}, \ \ x \geq 0$$

is not uniformly convergent on [0, k], k > 0

Solution. $\quad f(x) = \lim_{n\to\infty} f_n(x) = 0, \qquad \forall\ x \geq 0$

Also nxe^{-nx^2} attains maximum value $\sqrt{\dfrac{n}{2e}}$ at $x = \dfrac{1}{\sqrt{2n}}$

Now

$$M_n = \sup_{x\in[0,k]} |f_n(x) - f(x)|$$

$$= \sup_{x\in[0,k]} nxe^{-nx^2} = \sqrt{\frac{n}{2e}} \to \infty \text{ as } n\to\infty$$

Therefore the sequence is not uniformly convergent on [0, k].

Example. Prove that the sequence $\{f_n\}$, where $f_n(x) = x^{n-1}(1-x)$ converges uniformly in the interval [0, 1].

Solution. Here $f(x) = \lim_{n\to\infty} x^{n-1}(1-x) = 0 \ \ \forall\ x \in [0, 1]$.

Let $\quad y = |f_n(x) - f(x)| = x^{n-1}(1-x)$

Now y is maximum or minimum when

$$\frac{dy}{dx} = (n-1)x^{n-2}(1-x) - x^{n-1} = 0$$

$$x^{n-2}[(n-1)(1-x) - x] = 0$$

or $$x = 0 \text{ or } \frac{n-1}{n}.$$

A so $\frac{d^2y}{dx^2} = -ve$ when $x = \frac{n-}{n}$

$\therefore$ $M_n = \max y = \left(1+\frac{1}{n}\right)^{n-1}\left(1-\frac{n-1}{n}\right) \to \frac{1}{e} \times 0 = 0$ as $n\to\infty$.

Hence the sequence is uniformly convergent on [0, 1] by M_n-test.

Example. Show that 0 is a point of non-uniform convergence of the sequence $\{f_n\}$, where $f_n(x) = 1-(1 - x^2)^n$.

Solution. Here

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

Suppose, if possible, that the sequence is uniformly convergent in a neighborhood]0, k[of 0 where k is a number such that $0 < k < \sqrt{2}$. There exists therefore a positive integer m such that

$$|f_m(x) - f(x)| < \frac{1}{2}, \text{ taking } \in = \frac{1}{2},$$

i.e. if $$(1 - x^2)^m < \frac{1}{2} \text{ for every } x \in]0, k[.$$

Since $(1-x^2)^m \to 1$ as $x\to 0$, we arrive at a contradiction. Hence 0 is point of non-uniform convergence of the sequence.

Example. Test for uniform convergence the series

$$\sum_{n=0}^{\infty} xe^{-nx} \text{ in the closed interval } [0, 1].$$

Solution. Here $f_n(x) = \sum_{n=1}^{n-1} xe^{-nx} = \frac{x(-11/e^{nx})}{1-1/e^x}$

$$= \frac{xe^x}{e^x-1}\left(1+\frac{1}{e^{nx}}\right)$$

$$\text{Now } f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{where } x = 0 \\ \dfrac{xe^x}{e^x - 1} & \text{when } 0 < x \le 1 \end{cases}$$

We consider $0 < x \le 1$. We have

$$M_n = \sup \{|f_n(x) - f(x) : x \in [0, 1]\}$$

$$= \sup \left\{ \frac{xe^x}{(e^x - 1)e^{nx}} : x \in [0,1] \right\}$$

$$\ge \frac{1/n.e^{1/n}}{(e^{1/n} - 1)e} \qquad \left(\text{Taking } x = \frac{1}{n} \in [0,1] \right)$$

$$\text{Now} \quad \lim_{n\to\infty} \frac{1/n\ e^{1/xn}}{(e^{1/n} - 1)} \qquad \left[\text{Form } \frac{0}{0} \right]$$

$$= \lim_{n\to\infty} \frac{1/ne^{1/n}(-1/n^2) + (-1/n^2)e^{1/n}}{e.e^{1/n} - (-1/n^2)}$$

$$= \lim_{n\to\infty} \frac{(1/n + 1)}{e} = \frac{(0+1)}{e} = \frac{1}{e}.$$

Thus M_n does not tend zero as $n\to\infty$.

Hence the sequence is non-uniformly convergent by M_n-test.

Here 0 is a point of non-uniform convergence.

Example . The sequence $\{f_n\}$, where

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

is not uniformly convergent on any interval containing zero.

Solution. Here

$$\lim_{n\to\infty} f_n(x) = 0, \qquad \forall\ x$$

Now $\dfrac{nx}{1+n^2x^2}$ attains the maximum value $\dfrac{1}{2}$ at $x = \dfrac{1}{n}; \dfrac{1}{n}$ tending to 0 as $n\to\infty$. Let us take an interval [a, b] containing 0.

Thus

$$M_n = \underset{x\in[a,b]}{\text{Sup}} |f_n(x) - f(x)|$$

$$= \underset{x\in[a,b]}{\text{Sup}} \left| \frac{nx}{1+n^2x^2} \right|$$

$$= \frac{1}{2}, \text{ which does not tend to zero as } n\to\infty.$$

Hence the sequence $\{f_n\}$ is not uniformly convergent in any interval containing the origin.

Theorem 4. (Weierstrass's M-test). A series of functions $\sum f_n$ will converge uniformly (and absolutely) on [a, b] if there exists a convergent series $\sum M_n$ of positive numbers such that for all $x \in [a, b]$

$$|f_n(x)| \le M_n, \text{ for all } n$$

Let $\varepsilon > 0$ be a positive number.

Since $\sum M_n$ is convergent, therefore there exists a positive integer N such that

$$|M_{n+1} + M_{n+2} + \ldots + M_{n+p}| < \varepsilon \quad \forall\ n \ge N,\ p \ge 1 \qquad \ldots(1)$$

Hence for all $x\in [a, b]$ and for all $n \ge N$, $p \ge 1$, we have

$$|f_{n+1}(x) + f_{n+2}(x) + \ldots + f_{n+p}(x)| \le |f_{n+1}(x)| + |f_{n+2}(x)| + \ldots + |f_{n+p}(x)| \qquad \ldots(2)$$

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

$$< \varepsilon \qquad \ldots(3)$$

(2) and (3) imply that $\sum f_n$ is uniformly and absolutely convergent on [a, b].

Example. Test for uniform convergence the series.

(i) $\sum \frac{x}{(n+x^2)^2}$, (ii) $\sum \frac{x}{n(1+nx^2)}$

Solution. (i) Here $u_n(x) = \frac{x}{(n+x^2)^2}$.

Now $u_n(x)$ is maximum or minimum when $\frac{du_n(x)}{dx} = 0$

or $$(n + x^2)^2 - 4x^2 (n + x^2) = 0$$

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

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

If will be seen that $\frac{d^2u_n(x)}{dx^2}$ is –ve when $x = \sqrt{\frac{n}{3}}$.

Hence Max $u_n(x) = \frac{\sqrt{\frac{n}{3}}}{\left(n + \frac{n}{3}\right)^2} = \frac{3\sqrt{3}}{16n^{3/2}} = M_n$.

Therefore $|u_n(x)| \le M_n$.

But $\sum M_n$ is convergent.

Hence the given series is uniformly convergent for all values of x by Weierstrass's M test.

(ii) Here $u_n(x)$ is Maximum or minimum when

$$n(1 + nx^2) - 2n^2x^2 = 0 \quad \text{or} \quad x = \pm 1/\sqrt{(n)}.$$

It can be easily shown that $x = \frac{1}{\sqrt{n}}$ makes $u_n(x)$ a maximum.

Hence Max $u_n(x) = \frac{1/\sqrt{n}}{n(1+1)} = \frac{1}{2.n^{3/2}} = M_n$. But $\sum M_n$ is convergent.

Hence the given series is uniformly convergent for all values of x by Weierstrass's M-test.

Example :- Consider $\sum_{n=1}^{\infty} \frac{x}{n(1+nx^2)}$, $x \in \mathbf{R}$.

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 $\quad f_n'(x) = 0$

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

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

Since $\sum \frac{1}{n^{3/2}}$ is convergent, Weierstrass' M-Test implies that $\sum_{n=1}^{\infty} \frac{x}{n(1+nx^2)}$ is uniformly convergent for all $x \varepsilon \mathbf{R}$.

Example :- Consider the series $\sum_{n=1}^{\infty} \frac{x}{(n+x^2)^2}$, $x\ \varepsilon \mathbf{R}$. We have

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

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

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

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

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

Example **.** The series $\sum \frac{x}{n^p + x^2 n^q}$ converges uniformly over any finite interval [a, b], for (i) $p > 1, q \geq 0$ (ii) $0 < p \leq 1,\ p + q > 2$

(i) When $p > 1, q \geq 0$

$$|f_n(x)| = \left| \frac{x}{n^p + x^2 n^q} \right| \leq \frac{\alpha}{n^p}$$

where $\alpha \geq \max \{|a|, |b|\}$.

The series $\sum(\alpha / n^p)$ converges for $p > 1$.

Hence by M-test, the given series converges uniformly over the interval [a, b].

(ii) When $0 < p \leq 1, p + q > 2$.

$|f_n(x)|$ attains the maximum value $\frac{1}{2n^{\frac{1}{2}(p+q)}}$ at the point, where $x^2 n^4 = n^p$

$$\therefore \quad |f_n(x)| \leq \frac{1}{2n^{\frac{1}{2}(p+q)}}$$

The series $\sum \frac{1}{2n^{\frac{1}{2}(p+q)}}$ converges for $p + q > 2$. Hence by M-test, the given series converges uniformly over any finite interval [a, b].

Example . Test for uniform convergence, the series

$$\frac{2x}{1+x^2} + \frac{4x^3}{1+x^4} + \frac{8x^7}{1+x^8} + \ldots, -\frac{1}{2} \leq x \leq \frac{1}{2}$$

Solution. The nth term $f_n(x) = \frac{2^n x^{2^n - 1}}{1 + x^{2^n}}$

$$|f_n(x)| \leq 2^n (\alpha)^{2^n - 1}$$

where $|x| \leq \alpha \leq \frac{1}{2}$.

The series $\sum 2^n (\alpha)^{2^n - 1}$ converges, and hence by M-test the given series converges uniformly on $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

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

$$u_1v_1 + u_2v_2 + \ldots + u_nv_n$$

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

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

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

$$\sum_{i=1}^{n} u_iv_i = u_1 v_1 + u_1v_2 + \ldots + u_nv_n$$

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

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

$$= A[v_1 - v_2 + v_2 - v_3 + \ldots + v_{n-1} - v_n + v_n]$$

Similarly, we can show that

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

Hence the result follows.

Theorem (Abel's test). If $a_n(x)$ is a positive, monotonic decreasing function of n for each fixed value of x in the interval [a, b], and $a_n(x)$ is bounded for all values of n and x , and if the series $\sum u_n(x)$ converges uniformly on [a, b], then $\sum a_n(x)u_n(x)$ also converges uniformaly.

Proof. Since $a_n(x)$ is bounded for all values of n and for x in [a, b], therefore there exists a number K > 0, independent of x and n, such that for all $x \in [a, b]$,

$$0 \le a_n(x) \le K, \quad (\text{for } n = 1, 2, 3, \ldots) \qquad \ldots(1)$$

Again, since $\sum u_n(x)$ converges uniformly on [a, b], therefore for any $\varepsilon > 0$, we can find and integer N such that

$$\left|\sum_{r=n+1}^{n+p} u_r(x)\right| < \frac{\varepsilon}{K}, \qquad \forall\ n \geq N,\ p \geq 1 \qquad \ldots(2)$$

Hence using Abel's lemma we get

$$\left|\sum_{r=n+1}^{n+p} a_r(x)u_r(x)\right| \leq a_{n+1}(x) \qquad \max_{q=1,2,\ldots,p} \left|\sum_{r=n+1}^{n+q} u_r(x)\right|$$

$$< K\ \frac{\varepsilon}{K} = \varepsilon, \text{ for } n \geq N,\ p \geq 1,\ a \leq x \leq b$$

$\Rightarrow$ $\sum a_n(x)\ u_n(x)$ is uniformly convergent on [a, b].

Example . The series $\sum \frac{(-1)^n}{n}\ |x|^n$ is uniformly convergent in $-1 \leq x \leq 1$.

Solution. Since $|x|^n$ is positive, monotonic decreasing and bounded for $-1 \leq x \leq 1$, and the series $\sum \frac{(-1)^n}{n}$ is uniformly convergent, therefore $\sum \frac{(-1)^n}{n}\ |x|^n$ is also so in $-1 \leq x \leq 1$.

Theorem . (Dirchlet's test). If $a_n(x)$ is a monotonic function of n for each fixed value of x in [a, b], and $a_n(x) \to 0$ uniformly for $a \leq x \leq b$, and if there is a number $K > 0$, independent of x and n, such that for all values of x in [a, b],

$$\left|\sum_{r=1}^{n} u_r(x)\right| \leq K, \qquad \forall\ n$$

then , $\sum a_n(x)\ u_n(x)$ converges uniformly on [a, b].

Proof. Since $a_n(x)$ tends uniformly to zero, therefore for any $\varepsilon > 0$, there exists an integer N (independent of x) such that for all $x \in [a, b]$

$$|a_n(x)| < \varepsilon/4K, \quad \text{for all } n \geq N$$

Let $S_n(x) = \sum_{r=1}^{n} u_r(x)$, for all $x \in [a, b]$, and for all n,

$$\therefore \quad \sum_{r=n+1}^{n+p} a_r(x)\, u_r(x) = a_{n+1}(x)\ \{S_{n+1} - S_n\} + a_{n+2}(x)\ \{S_{n+2} - S_{n+1}\} + \ldots$$

$$+ a_{n+p}(x)\ \{S_{n+p} - S_{n+p-1}\}$$

$$= - a_{n+1}(x)\ S_n + \{a_{n+1}(x) - a_{n+2}(x)\}\ S_{n+1} + \dots$$

$$+ \{a_{n+p-1}(x) - a_{n+p}(x)\}\ S_{n+p-1} + a_{n+p}(x)\ S_{n+p}$$

$$= \sum_{r=n+1}^{n+p-1} \{a_r(x) - a_{r+1}(x)\}\ S_r(x) - a_{n+1}(x)\ S_n(x)$$

$$+ a_{n+p}(x)\ S_{n+p}(x)$$

$$\therefore \qquad \left| \sum_{r=n+1}^{n+p} a_r(x) u_r(x) \right| \le \sum_{r=n+1}^{n+p-1} | a_r(x) - a_{r+1}(x)|\ |S_r(x)| + |a_{n+1}(x)|\ |S_n(x)| +$$

$$+ |a_{n+p}(x)|\ |S_{n+p}(x)|$$

Making use of the monotonicity of $a_n(x)$

$$\sum_{r=n+1}^{n+p-1} | a_r(x) - a_{r+1}(x)| = |a_{n+1}(x) - a_{n+p}(x)|, \text{ for } a \le x \le b,$$

and the relation $|S_n(x)| \le K$, for all $x \in [a, b]$ and for all $n = 1, 2, 3, \dots$, we deduce that for all $x \in [a, b]$ and all $p \ge 1$, $n \ge N$

$$\left| \sum_{r=n+1}^{n+p} a_r(x) u_r(x) \right| \le K|a_{n+1}(x) - a_{n+p}(x)| + \frac{\varepsilon}{4K} 2K$$

$$< K \frac{\varepsilon}{2K} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore by Cauchy's criterion, the series $\sum a_n(x) u_n(x)$ converges uniformly on [a, b].

Example. Consider the series $\sum \frac{(-1)^{n-1}}{(n + x^2)}$ for uniform convergence for all values of x.

Solution. Let $u_n = (-1)^{n-1}$, $v_n(x) \frac{1}{n + x^2}$

Since $f_n(x) = \sum_{r=1}^{n} u_r = 0$ or 1 according as n is even or odd, $f_n(x)$ is bounded for all n.

Also $v_n(x)$ is a positive monotonic decreasing sequence, converging to zero for all real values of x.

Hence by Dirichlets test, the given series is uniformly convergent for all real values of x.

Example . Prove that the series $\sum(-1)^n \frac{x^2+n}{n^2}$, converges uniformly in every bounded interval, but does not converge absolutely for any value of x.

Solution. Let the bounded interval be [a, b], so that ∃ a number K such that, for all x in [a, b], $|x| < K$.

Let us take $\sum u_n = \sum(-1)^n$, which oscillates finitely, and

$$a_n = \frac{x^2+n}{n^2} < \frac{K^2+n}{n^2}$$

Clearly a_n is a positive, monotonic decreasing function of n for each x in [a, b], and tends to zero uniformly for $a \le x \le b$.

Hence by Dirchlet's test, the series $\sum(-1)^n \frac{x^2+n}{n^2}$ converges uniformly on [a, b].

Again $\sum\left|(-1)^n \frac{x^2+n}{n^2}\right| = \sum \frac{x^2+n}{n^2} \sim \sum \frac{1}{n}$, which diverges. Hence the given series is not absolutely convergent for any value of x.

Example. Prove that if δ is any fixed positive number less than unity, the series $\sum \frac{x^n}{n+1}$ is uniformly convergent in $[-\delta, \delta]$.

Solution. Let $u_n(x) = x^n$, $v_n = \frac{1}{n+1}$

$|x| \le \delta < 1$, we have

$|f_n(x)| = |x + x^2 + \ldots + x^n|$

$\le |x| + |x|^2 + \ldots + |x|^n$

$$\leq \delta + \delta^2 + \ldots + \delta^n = \frac{\delta(1-\delta^n)}{1-\delta} < \frac{\delta}{1-\delta}.$$

Also $\{v_n\}$ is a monotonic decreasing sequence converging to zero.

Hence the given series is uniformly convergent by Dirichlet's test.

Example. Show that the series $\sum_{n=1}^{\infty}(-1)^{n-1}\ x^n$ converges uniformly in $0 \leq x \leq k < 1$.

Solution. Let $u_n = (-1)^{n-1}$, $v_n(x) = x^n$.

Since $f_n(x) = \sum_{n=1}^{n} u_r = 0$ or 1 according as n is even or odd, $f_n(x)$ is bounded for all n. Also $\{v_n(x)\}$ is a positive monotonic decreasing sequence, converging to zero for all values of x in $0 \leq x \leq k < 1$. Hence by Dirichlet's test, the given series is uniformly convergent in $0 \leq x \leq k < 1$.

Example 14. Prove that the series $\sum \frac{\cos n\theta}{n^p}$ converges uniformly for all values of $p > 0$ in an interval $[\alpha, 2\pi - \alpha[$, where $0 < \alpha < \pi$.

Solution. When $0 < p \leq 1$, the series converges uniformly in any interval $[\alpha, 2\pi - \alpha]$, $\alpha > 0$. Take $a_n = (1/n^p)$ and $u_n = \cos n\theta$ in Dirchlet's test.

Now $(1/n^p)$ is positive monotonic decreasing and tending uniformly to zero for $0 < p \leq 1$, and

$$\left|\sum_{t=1}^{n} u_t\right| = \left|\sum_{t=1}^{n} \cos t\theta\right| = |\cos\theta + \cos 2\theta + \ldots + \cos n\theta|$$

$$= \left|\frac{\cos((n+1)/2)\theta \sin(n/2)\theta}{\sin(\theta/2)}\right| \leq \text{cosec}\,(\alpha/2),\ \forall\ n,$$

$$\text{for } \theta \in [\alpha, 2\pi - \alpha]$$

Now by Dirchlet test , the series $\sum(\cos n\theta/n^p)$ converges uniformly on $[\alpha, 2\pi - \alpha]$ where $0 < \alpha < \pi$. When $p > 1$, Weierstrass's M-test , the series converges uniformly for all real values of θ.

MAL-512: M. Sc. Mathematics (Real Analysis)

Lesson No. III **Written by Dr. Nawneet Hooda**

Lesson: Power series **Vetted by Dr. Pankaj Kumar**

THE WEIERSTRASS APPROXIMATION THEOREM

Theorem . Let f be a real continuous function defined on a closed interval [a, b] then there exists a sequence of real polynomials $\{P_n\}$ such that $\lim_{n\to\infty} P_n(x) = f(x)$, uniformly on [a, b].

Proof. If a = b, we take $P_n(x)$ to be a constant polynomial, defined by $P_n(x) = f(a)$, for all n and the conclusion follows .

So let a < b.

The linear transformation $x' = (x - a)/(b - a)$ is a continuous mapping of [a, b] onto [0, 1]. So, we take a = 0, b = 1.

The binomial coefficient $\binom{n}{k}$ is defined by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$, for positive integers n and k when $0 \le k \le n$,

The Bernstein polynomials B_n associated with f is defined as

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(k/n), \; n = 1, 2, 3, \ldots, \text{ and } x \in [0, 1]$$

By binomial theorem,

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} = [x + (1 - x)]^n = 1 \qquad \ldots(1)$$

Differentiating with respect to x, we get

$$\sum_{k=0}^{n} \binom{n}{k} [k\, x^{k-1}(1 - x)^{n[k} - (n - k)\, x^k (1 - x)^{n-k-1}] = 0$$

or

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

Now multiply by x(1 − x) , we take

$$\sum_{k=0}^{n}\binom{n}{k}x^{k}(1-x)^{n-k}(k-nx)=0 \qquad \ldots(2)$$

Differentiating with respect to x, we get

$$\sum_{k=0}^{n}\binom{n}{k}[-nx^{k}(1-x)^{n-k}+x^{k-1}(1-x)^{n-k-1}(k-nx)^{2}]=0$$

Using (1), we have

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

and on multiplying by x(1− x), we get

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

or

$$\sum_{k=0}^{n}\binom{n}{k}x^{k}(1-x)^{n-k}(x-k/n)^{2}=\frac{x(1-x)}{n} \qquad \ldots(3)$$

The maximum value of x(1 − x) in [0, 1] being $\frac{1}{4}$.

$$\therefore \qquad \sum_{k=0}^{n}\binom{n}{k}x^{k}(1-x)^{n-k}(x-k/n)^{2}\le\frac{1}{4n} \qquad \ldots(4)$$

Continuity of f on the closed interval [0, 1], implies that f is bounded and uniformly continuous on [0, 1].

Hence there exists K > 0, such that

$$|f(x)| \le K, \qquad \forall\ x \in [0, 1]$$

and for any ε > 0, there exists δ > 0 such that for all x∈[0, 1].

$$|f(x) - f()k/n)| < \tfrac{1}{2}\varepsilon, \text{ when } |x - k/n| < \delta \qquad \ldots(5)$$

For any fixed but arbitrary x in [0, 1], the values 0, 1, 2, 3,…, n of k may be divided into two parts :

Let A be the set of values of k for which $|x - k/n| < \delta$, and B the set of the remaining values, for which $|x - k/n| \geq \delta$.

For $k \in B$, using (4),

$$\sum_{k\in B}\binom{n}{k}x^k(1-x)^{n-k}\delta^2 \leq \sum_{k\in B}\binom{n}{k}x^k(1-x)^{n-k}(x-k/n)^2 \leq \frac{1}{4n}$$

$$\Rightarrow \qquad \sum_{k\in B}\binom{n}{k}x^k(1-x)^{n-k} \leq \frac{1}{4n\delta^2} \qquad \text{…(6)}$$

Using (1), we see that for this fixed x in [0, 1],

$$|f(x) - B_n(x)| = \left|\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}[f(x)-f(k/n)]\right|$$

$$\leq \sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}|f(x) - f(k/n)|$$

Thus summation on the right may be split into two parts, according as $|x - k/n| < \delta$ or $|x - k/n| \geq \delta$. Thus

$$|f(x) - B_n(x)| \leq \sum_{k\in A}\binom{n}{k}x^k(1-x)^{n-k}|f(x) - f(k/n)|$$

$$+ \sum_{k\in B}\binom{n}{k}x^k(1-x)^{n-k}|f(x) - f(k/n)|$$

$$< \frac{\varepsilon}{2}\sum_{k\in A}\binom{n}{k}x^k(1-x)^{n-k} + 2K\sum_{k\in B}\binom{n}{k}x^k(1-x)^{n-k}$$

$$\leq \varepsilon/2 + 2K/4n\delta^2 < \varepsilon, \text{ using (1), (5) and (6),}$$

for values of n greater than $K/\varepsilon\delta^2$.

Thus $\{B_n(x)\}$ converges uniformly to f(x) on [0, 1].

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

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

then show that $f(x) = 0$ on $[0, 1]$.

Solution. From (1), it follows that, the integral of the product of f with any polynomial is zero.

Now, since f is continuous on [0, 1], therefore, by 'Weierstrass approximation theorem', there exists a sequence $\{P_n\}$ of polynomials, such that $P_n \to f$ uniformly on [0, 1]. And so $P_n f \to f^2$ uniformly on [0, 1], since f, being continuous, is bounded on [0, 1]. Therefore,

$$\int_0^1 f^2 dx = \lim_{n\to\infty} \int_0^1 P_n f \, dx = 0, \text{ using (1)}$$

$\therefore$ $f^2 = 0$ on [0, 1]. Hence $f = 0$ on [0, 1].

POWER SERIES

Consider series of the form

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots \equiv \sum_{n=0}^{\infty} a_n x^n$$

This is called power series (in x) and the numbers a_n(dependent on n but not on x) it's coefficients.

If a power series converges for no value of x other than $x = 0$, we then say that it is nowhere convergent. If it converges for all values of x, it is called everywhere convergent.

Thus if $\sum a_n x^n$ is a power series which does not converge everywhere or nowhere, then a definite positive number R exists such that $\sum a_n x^n$ converges (absolutely) for every $|x| < R$ but diverges for every $|x| > R$. The number R, which is associated with every power series, is called the radius of convergence and the interval, $(-R, R)$, the interval of convergence, of the given power series.

Theorem. 1. If $\overline{\lim} |a_n|^{1/n} = \frac{1}{R}$, then the series $\sum a_n x^n$ is convergent (absolutely) for $|x| < R$ and divergent for $|x| > R$.

Proof.

Now

$$\overline{\lim_{n\to\infty}} |a_n x^n|^{1/n} = \frac{|x|}{R}$$

Hence by Cauchy's root test, the series $\sum a_n x^n$ is absolutely convergent and therefore convergent for $|x| < R$ and divergent for $|x| > R$.

Definition. The radius of convergence R of a power series is defined to be equal to

$$\frac{1}{\overline{\lim} |a_n|^{1/n}}, \text{ when } \overline{\lim} |a_n|^{1/n} > 0$$

$$\infty \ , \qquad \text{when } \overline{\lim} |a_n|^{1/n} = 0$$

$$0 \ , \qquad \text{when } \overline{\lim} |a_n|^{1/n} = \infty$$

Thus for a nowhere convergent power series, $R = 0$, while for an everywhere convergent power series, $R = \infty$.

Theorem . If a power series $\sum a_n x^n$ converges for $x = x_0$ then it is absolutely convergent for every $x = x_1$, when $|x_1| < |x_0|$.

Solution. Since the series $\sum a_n x_0^n$ is convergent, therefore $a_n x_0^n \to 0$, as $n\to\infty$.

Thus, for $\varepsilon = \frac{1}{2}$ (say), there exists an integer N such that

$$|a_n x_0^n| < \tfrac{1}{2}, \text{ for } n \geq N, \text{ and so}$$

$$|a_n x_1^n| = |a_n x_0^n| \cdot \left|\frac{x_1}{x_0}\right|^n < \frac{1}{2} \left|\frac{x_1}{x_0}\right|^n, \text{ for } n \geq N$$

But $\sum \left|\frac{x_1}{x_0}\right|^n$ is a convergent geometric series with common ratio $\left|\frac{x_1}{x_0}\right| < 1$.

Therefore, by comparison test, the series $\sum |a_n x_1^n|$ converges.

Hence $\sum a_n x^n$ is absolutely convergent for every $x = x_1$, when $|x_1| < |x_0|$.

Theorem . If a power series $\sum a_n x^n$ diverges for $x = x'$, then it diverges for every $x = x''$, where $|x''| > |x'|$.

Proof. If the series was convergent for $x = x''$ then it would have to converge for all x with $|x| < |x''|$, and in particular at x', which contradicts the hypothesis. Hence the theorem.

Example . Find the radius of convergence of the series

(i) $x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$ (ii) $1 + x + 2!\, x^2 + 3! + 4!\, x^4 + \ldots$

Solution. (i) The radius of convergence $R = \lim \left|\frac{a_n}{a_{n+1}}\right| = \lim \frac{(n+1)!}{n!} = \infty.$

Therefore the series converges absolutely for all values of x.

(ii) The radius of convergence $R = \lim_{n\to\infty} \frac{n!}{(n+1)!} = 0$. Therefore the series converges for no value of x, of course other than zero.

Example . Find the interval of absolute convergence for the series $\sum_{n=1}^{\infty} x^n / n^n$.

Solution. It is a power series and will therefore be absolutely convergent within its interval of convergence. Now, the radius of convergence

$$R = \frac{1}{\overline{\lim} |a_n|^{1/n}} = \frac{1}{\lim \left|\frac{1}{n^n}\right|^{1/n}} = \infty$$

Hence the series converges absolutely for all x.

Theorem . If a power series $\sum a_n x^n$ converge for $|x| < R$, and let

$$f(x) = \sum a_n x^n,\ |x| < R.$$

then $\sum a_n x^n$ converges uniformly on $[-R + \varepsilon, R-\varepsilon]$, where $\varepsilon > 0$ and that the function f is continuous and differentiable on$(-R, R)$ and

$$f'(x) = \sum n a_n x^{-1},\ |x| < R \qquad \ldots(1)$$

Proof. Let $\varepsilon > 0$ be any number given.

For $|x| \le R - \varepsilon$, we have

$$|a_n x^n| \le |a_n|(R - \varepsilon)^n.$$

But since $\sum a_n(R-\varepsilon)^n$, converges absolutely , therefore by Weierstrass's M-test, the series $\sum a_n x^n$ converges uniformly on $[-R + \varepsilon, R-\varepsilon]$.

Again, since every term of the series $\sum a_n x^n$ is continuous and differentiable on $(-R, R)$, and $\sum a_n x^n$ is uniformly convergent on $[-R + \varepsilon, R -\varepsilon]$, therefore its sum function f is also continuous and differentiable on $(-R, R)$.

Also

$$\overline{\lim_{n\to\infty}} |na_n|^{1/n} = \overline{\lim_{n\to\infty}} (n^{1/n})|a_n|^{1/n} = 1/R$$

Hence the differentiated series $\sum na_n x^{n-1}$ is also a power series and has the same radius of convergence R as $\sum a_n x^n$. Therefore $\sum na_n x^{n-1}$ is uniformly convergent in $[-R + \varepsilon, R -\varepsilon]$.

Hence

$$f'(x) = \sum na_n x^{n-1}, \quad |x| < R$$

Observations. 1 By above theorem, f has derivatives of all orders in $(-R, R)$, which are given by

$$f^{(m)}(x) = \sum_{n=m}^{\infty} n(n-1)...(n-m+1)a_n x^{n-m}, \qquad \text{...(2)}$$

and in particular,

$$f^{(m)}(0) = m!\, a_m, \ (m = 0, 1, 2,\dots) \qquad \text{...(3)}$$

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

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

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

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

Putting x = 0, this yields

$f^{(k)}(0) = \lfloor k \, a_k$ and $f^{k}(0) = \lfloor k \; b_k$

for all $k \in N$. Hence

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

This completes the proof or the theorem.

Abel's Theorem (First form). If a power series $\sum_{n=0}^{\infty} a_n x^n$ converges at x = R of the interval of convergence (−R, R), then it is uniformly convergent in the closed interval [0, R].

Proof. Let $S_{n,p} = a_{n+1} R^{n+1} + a_{n+2} R^{n+2} + \ldots + a_{n+p} R^{n+p}$, p = 1, 2,…

Then ,

$$a_{n+1}R^{n+1} = S_{n,1}$$

$$a_{n+2}R^{n+2} = S_{n,2} - S_{n,1}$$

$$\vdots$$

$$a_{n+p}R^{n+p} = S_{n,p} - S_{n,p-1} \qquad \ldots(1)$$

Let ε > 0 be given.

Since the number series $\sum_{n=0}^{\infty} a_n R^n$ is convergent, therefore by Cauchy's general principle of convergence, there exists an integer N such that for $n \geq N$,

$$|S_{n,q}| < \varepsilon, \text{ for all } q = 1, 2, 3, \ldots \qquad \ldots(2)$$

Note that

$$\left(\frac{x}{R}\right)^{n+p} \leq \left(\frac{x}{R}\right)^{n+p-1} \leq \ldots \leq \left(\frac{x}{R}\right)^{n+1} \leq 1, \text{ for } 0 \leq x \leq R$$

and using (1) and (2) we have for $n \geq N$

$|a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \ldots + a_{n+p}x^{n+p}|$

$$= \left| a_{n+1}R^{n+1}\left(\frac{x}{R}\right)^{n+1} + a_{n+2}R^{n+2}\left(\frac{x}{R}\right)^{n+2} + \ldots a_{n+p}R^{n+p}\left(\frac{x}{R}\right)^{n+p} \right|$$

$$= \left| S_{n,1}\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2}\right\} + S_{n,2}\left\{\left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3}\right\} + \ldots \right.$$

$$+ S_{n,p-1}\left\{\left(\frac{x}{R}\right)^{n+p-1} - \left(\frac{x}{R}\right)^{n+p}\right\} + S_{n,p}\left(\frac{x}{R}\right)^{n+p}\Bigg|$$

$$\leq \quad |S_{n,1}|\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2}\right\} + |S_{n,2}|\left\{\left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3}\right\} + \ldots$$

$$+ |S_{n,p-1}|\left\{\left(\frac{x}{R}\right)^{n+p-1} - \left(\frac{x}{R}\right)^{n+p}\right\} + |S_{n,p}|\left(\frac{x}{R}\right)^{n+p}$$

$$< \quad \varepsilon\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2} + \left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3} + \ldots\right.$$

$$\left. - \left(\frac{x}{R}\right)^{n+p} + \left(\frac{x}{R}\right)^{n+p}\right\}$$

$$= \varepsilon\left(\frac{x}{R}\right)^{n+1} \leq \varepsilon \text{ for all } n \geq N,\ p \geq 1, \text{ and for all } x \in [0, R].$$

Hence by Cauchy's criterion, the series converges uniformly on [0, R].

Abel's Theorem (Second form). Let R be the radius of convergence of a power series $\sum a_n x^n$ and let $f(x) = \sum a_n x^n$, $-R < x < R$. If the series $\sum a_n R^n$ converges, then

$$\lim_{x \to R-0} f(x) = \sum a_n R^n$$

Proof. Taking $x = Ry$, we get

$$\sum a_n x^n = \sum a_n R^n y^n = \sum b_n y^n, \text{ where } b_n = a_n R^n.$$

It is a power series with radius of convergence R', where

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

So, there is no loss of generality in taking R = 1.

Let $\sum_{0}^{\infty} a_n x^n$ be a power series with unit radius of convergence and let

$f(x) = \sum_{0}^{\infty} a_n x^n$, $-1 < x < 1$. If the series $\sum a_n$ converges, then

$$\lim_{x\to 1-0} f(x) = \sum_0^\infty a_n$$

Let $S_n = a_0 + a_1 + a_2 + \ldots + a_n$, $S_{-1} = 0$, and let $\sum_{n=0}^{\infty} a_n = S$, then

$$\sum_{n=0}^{m} a_n x^n = \sum_{n=0}^{m} (S_n - S_{n-1}) x^n = \sum_{n=0}^{m-1} S_n x^n + S^m x^m - \sum_{n=0}^{m} S_{n-1} x^n$$

$$= \sum_{n=0}^{m-1} S_n x^n - x \sum_{n=0}^{m} S_{n-1} x^{n-1} + S_n x^m$$

$$= (1 - x) \sum_{n=0}^{m-1} S_n x^n + S_m x^m$$

For $|x| < 1$, when $m \to \infty$, since $S_m \to S$, and $x^m \to 0$, we get

$$f(x) = (1 - x) \sum_{n=0}^{\infty} S_n x^n, \quad \text{for } 0 < x < 1.$$

Again, since $S_n \to S$, for $\varepsilon > 0$, there exists N such that

$$|S_n - S| < \varepsilon/2, \text{ for all } n \geq N$$

Also

$$(1-x) \sum_{n=0}^{\infty} x^n = 1, \; |x| < 1 \qquad \ldots(3)$$

Hence for $n \geq N$, we have, for $0 < x < 1$,

$$|f(x) - S| = \left| (1-x) \sum_{n=0}^{\infty} S_n x^n - S \right| \qquad \text{[by 1]}$$

$$= \left| (1-x) \sum_{n=0}^{\infty} (S_n - S) x^n \right| \qquad \text{[by 3]}$$

$$\leq (1 - x) \sum_{n=0}^{N} |S_n - S| x^n + \frac{\varepsilon}{2} (1-x) \sum_{n=N+1}^{\infty} x^n \qquad \text{[by 2]}$$

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

For a fixed N, $(1-x)\sum_{n=0}^{N}|S_n - S|\,x^n$ is a positive continuous function of x, and vanishes at x = 1 . Therefore, there exists $\delta > 0$, such that for $1-\delta < x < 1$,

$$(1-x)\sum_{n=0}^{N}|S_n - S|\,x^n < \varepsilon/2.$$

$$\therefore\ |f(x) - S| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \text{ when } 1-\delta < x < 1$$

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

Example . Prove that

$$\frac{1}{2}(\tan^{-1} x)^2 = \frac{x^2}{2} - \frac{x^4}{4}\left(1+\frac{1}{3}\right) + \frac{x^6}{6}\left(1+\frac{1}{3}+\frac{1}{5}\right) + \ldots, \quad -1 < x \le 1.$$

Proof. We have

$$\tan^{-1}x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots, -1 \le x \le 1$$

and

$$(1+x^2)^{-1} = 1 - x^2 + x^4 - x^6 + \ldots, -1 < x < 1$$

Both the series are absolutely convergent in $(-1, 1)$, therefore their Cauchy product will converge absolutely to the product of their sums, $(1+x^2)^{-1}\tan^{-1}x$ in $(-1, 1)$.

$$\therefore \quad (1+x^2)^{-1}\tan^{-1}x = x - \left(1+\frac{1}{3}\right)x^3 + \left(1+\frac{1}{3}+\frac{1}{5}\right)x^5 - \ldots, \quad -1 < x < 1$$

Integrating,

$$\frac{1}{2}(\tan^{-1} x)^2 = \frac{x^2}{2} - \frac{x^4}{4}\left(1+\frac{1}{3}\right) + \frac{x^6}{6}\left(1+\frac{1}{3}+\frac{1}{5}\right) - \ldots, \quad -1 < x < 1$$

the constant of integration vanishes.

The power series on the right converges at x = 1 also, so that by Abel's theorem,

$$\frac{1}{2}(\tan^{-1} x)^2 = \frac{x^2}{2} - \frac{x^4}{4}\left(1+\frac{1}{3}\right) + \frac{x^6}{6}\left(1+\frac{1}{3}+\frac{1}{5}\right) - \ldots, -1 < x \le 1$$

Example . Show that

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, \; -1 < x \le 1,$$

and deduce that

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

Solution. We know

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \ldots, \quad -1 < x < 1$$

Integrating,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, \quad -1 < x < 1$$

the constant of integration vanishes as can be verified by putting $x = 0$.

The power series on the right converges at $x = 1$ also. Therefore by Abel's theorem

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, \quad -1 < x \le 1$$

At $x = 1$, we get, by Abel's theorem (second form),

$$\log 2 = \lim_{x \to 1-} \log(1+x) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

Example **.** Show that

$$\frac{1}{2}[\log(1+x)]^2 = \frac{x^2}{2} - \frac{x^3}{3}(1+\frac{1}{2}) + \frac{x^4}{4}(1+|\frac{1}{2}+\frac{1}{3}) - \ldots, \quad -1 < x \le 1$$

We know

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots, \quad -1 < x \le 1$$

and

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - \ldots, \; -1 < x < 1$$

Both the series are absolutely convergent in $]-1, 1[$, therefore their Cauchy product will converge to $(1+x)^{-1} \log(1+x)$. Thus

$$(1+x)^{-1}\log(1+x) = x - x^2\left(1+\frac{1}{2}\right) + x^3\left(1+\frac{1}{2}+\frac{1}{3}\right) - \ldots,\quad -1 < x < 1$$

Integrating,

$$\frac{1}{2}[\log(1+x)]^2 = \frac{x^2}{2} - \frac{x^3}{3}(1+\frac{1}{2}) + \frac{x^4}{4}(1+\frac{1}{2}+\frac{1}{3}) - \ldots, -1 < x < 1$$

the constant of integration vanishes.

Since the series on the right converges at x = 1 also, therefore by Abel's. Theorem, we have

$$\frac{1}{2}[\log(1+x)]^2 = \frac{x^2}{2} - \frac{x^3}{3}(1+\frac{1}{2}) + \frac{x^4}{4}(1+\frac{1}{2}+\frac{1}{3}) - \ldots, -1 < x \le 1$$

Linear transformations.

Definitions. (i) Let X be a subset of R^n. Then X is said to be a vector space if $x \in X$, $y \in X \Rightarrow x+y \in X$ and $cx \in X$ for all

(ii) If $x_1, x_2, \ldots, x_m \in R^n$ and $c_1, c_2, \ldots. C_m$ are scalars, then the vector $c_1x_1 + c_2x_2 + \ldots + c_mx_m$ is called a linear combination of $x_1, x_2, \ldots, x_m$.

(iii) If SCR^n and if A is the set of linear combinations of elements of S, then are say that S spans A or that A is the span of S.

(iv) We say that the set of vectors $\{x_1, x_2, \ldots, x_n\}$ is independent if $c_1x_1 + c_2x_2 + \ldots + c_mx_m = 0 \Rightarrow c_1 = c_2 \ldots = c_m = 0$.

(v) A vector space X is said to have dimension k if X contains an independent set of k vectors but no independent set of (k + 1) vectors. We then write dim X = k.

The set consists g of 0(zero vector) alone is a vector space. Its dimension is 0.

(vi) A subset B of a vector space X is said to be a basis of X if B is independent and spans X.

Remark (i) Any set consisting of the 0 vector is dependent or equivalently, no independent set contains the zero vector.

(ii) If $B = \{x_1, x_2, \ldots x_m\}$ is a basis of a vector space X, then every $x \in X$ can be expressed unequally in the form

$x = \sum_{i=1}^{m} c_i x_i$. The numbers $c_1, c_2, \ldots, c_m$ are called coordinates of x w.r.t. the basis B. The following theorems of vector-space will be used.

Theorem. If a vector space X is spanned by a set or k vectors, then dim X ≤ k.

Corollary. Dim $k^n = n$.

Theorem. Let X be a vector space with dim X = n. Then

(i) a set A of n vectors in X spans X ⇔ A is independent.

(ii) X has a basis, and every basis consists n-vectors

(iii) If $1 \le r \le n$ and $\{y_1, y_2, \ldots, y_r\}$ is an independent set in X. Then X has a basis containing $\{y_1, y_2, \ldots, y_r\}$.

Definition. Let X and Y be two vector spaces. A mapping $T : X \to Y$ is said to be linear transformation if $T(x_1 + x_2) = Tx_1 + Tx_2$, $T(cx) = cTx \ \forall \ x_1, x_2, x \in X$ and all scalars C.

Linear transformations of X into X will be called linear operators on X. Observe that T(0) = 0 for any Linear transformation T. We say that a Linear transformations T on X is invertible if T is one-one and onto. T is invertible, then we can define an operator T^{-1} on X by setting $T^{-1}(Tx) = x \ \forall \ x \in X$. Also in this cae, we have $T(T^{-1}x) = x \ \forall \ x \in X$ and that T^{-1} is linear.

Theorem. A linear transformation T on a finite dimensional vector space X is one-one ⇔ the range of T is all of X. i.e. T is onto.

Definition. (i) Let X and Y be vector spaces. Denote by L(X, Y) the set of all linear transformations of X into Y. If $T_1, T_2 \in L(X, Y)$ and if c_1, c_2 are scalars, we define

$$(c_1T_1 + c_2T_2)x = c_1T_1x + c_2T_2x \ \forall \ x \in X. \text{ Also } c_1T_1 + c_2T_2 \in L(X, Y)$$

(ii) Let X, Y, Z be vector spaces and let

$T \in L(X, Y)$, $U \in L(Y, Z)$ the product UT is defined by $(UT)x = U(Tx)$, $x \in X$. Then $UT \in L(X, Z)$. Observe that UT need not be the same as TU even if X = Y = Z.

(iii) For $T \in L(R^n, R^m)$, we define the norm $||T||$ of T to be the l.u.b $|Tx|$, where x ranges over all vectors R^n with $|x| \leq 1$.

Observe that the inequality $|Tx| \leq ||T|| \; |x|$ holds for all $x \in R^n$. Also if λ is such that $|Tx| \leq \lambda|x| \; \forall \; x \in R^n$, then $||T|| \leq \lambda$.

Theorem . Let $T \in L(R^n, R^m)$. Then $||T|| < \infty$ and T is a uniformly continuous mapping of R^n into R^m.

Proof. Let $E = \{e_1, \ldots e_n\}$ be the basis of R^n and let $x \in R^n$ with $|x| \leq 1$. Since E spans R^n, $\exists$ scalars $c_1, \ldots, c_n$ s. t. $x = \sum c_i e_i$ so that $|\sum c_i e_i| = |x| \leq 1$.

$\Rightarrow \qquad |c_i| \leq 1$ for $i = 1, \ldots .n$.

Then $|Tx| = |\sum c_i \, Te_i| \leq \sum |c_i| \; |Te_i| \leq \sum |Te_i|$.

It follows that $||T|| \leq \sum_{i=1}^{n} |Te_i| < \infty$.

For uniform continuity of T, observe that

$$|Tx - Ty| = |T(x - y)| \leq ||T|| \; |x - y|, \; (x, y, \in R^n)$$

So for $\varepsilon > 0$, we can choose $\delta = \dfrac{\varepsilon}{||T||}$ s. t.

$$|x - y| < \delta \;\; \Rightarrow |Tx - Ty| < ||T|| \; \frac{\varepsilon}{||T||} = \varepsilon \, .$$

Therefore T is uniformly continuous.

Theorem. If T, $U \in L(R^n, R^m)$ and c is a scalar, then

$$||T + U|| \leq ||T|| + ||U||, \; ||cT|| = |c| \;\; ||T||$$

and $L(R^n, R^m)$ is a metric space

Proof.

We have $\qquad |(T + U)x| = |Tx + Ux| \leq |Tx| + |Ux|$

$$\leq (||T|| + ||U||) \;\; |x|$$

Hence $||T + U|| \leq ||T|| + ||U||$

Similarly we can prove that $||cT|| = |c| \;\; ||T||$

To prove that $L(R^n, R^m)$ is a metric space, let U, V, W, $\in L(R^n, R^m)$, then

$$||U - W|| = ||(U - V) + (V - W)|| \leq ||U-V|| + ||V-W||$$

which is the triangle inequality.

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

$$||UT|| \leq ||U|| \quad ||T||$$

Proof. We have $|(UT)x| = |U(Tx)| \leq ||U|| \ |Tx| \leq ||U|| \ \ ||T|| \ \ |x|$

$\Rightarrow$ $\qquad ||UT|| \leq ||U|| \ \ ||T||.$

Theorem. Let *C* denote the set of all invertible linear operators on R^n.

(i) If $T \in C$, $||T^{-1}|| = \frac{1}{\alpha}$, $U \in L(R^n)$ and $||U - T|| = B < \alpha$ then $U \in C$

(ii) C is an open subset of $L(R^n)$ and the mapping $f : C \to C$ defined by $f(T) = T^{-1} \ \forall \ T \in \forall$ is continuous

Proof. For all $x \in R^n$, we have

$$|x| = |T^{-1}Tx| \leq ||T^{-1}|| \ \ |Tx| = \frac{1}{\alpha} \ |Tx| \text{ so that}$$

$$(\alpha-\beta) |x| = \alpha|x| - \beta|x| \leq |Tx| - ||U - T|| \ \ |x|$$

$$\leq |Tx| \ \ |(U-T)x| = |Tx| - |Ux, -Tx|$$

$$= |Tx| - |Tx - Ux|$$

$$\leq |Tx| - (|Tx| - |Ux|) \quad = |Ux|$$

Thus $|Ux| \geq (\alpha-\beta) |x| \ \forall \ x \in R^n$ …(1)

Now $\quad U_x = U_y \ \Rightarrow U_x - U_y = 0 \ \Rightarrow \ U(x - y) = 0$

$\Rightarrow$ $\qquad |U(x - y)| = 0 \ \Rightarrow |(\alpha-\beta)| \ \ |x - y| = 0$ by (1)

$\Rightarrow$ $\qquad |x - y| = 0 \quad \Rightarrow \ \alpha - y = 0 \quad \Rightarrow \ x = y.$

This shows that is one-one.

Also, U is also onto. Hence U is an invertible operator so that $U \in C$.

(ii) As shown in (i), if $T \in C$, then $\alpha = \frac{1}{\|T^{-1}\|}$ is s. t. every U with $\|U-T\| < \alpha$ belongs to C. Thus to show that C is open, replacing x in (1) by $U^{-1}y$, we have

$$(\alpha-\beta)\,|U^{-1}\,y| \le |UU^{-1}\,y| = |y|$$

So that $(\alpha-\beta)\,\|U^{-1}\|\ \ |y| \le |y|$ or $\|U^{-1}\| \le (\alpha-\beta)^{-1}$.

Now $|f(U) - f(T)| = |U^{-1} - T^{-1}| = |U^{-1}(T-U)\,T^{-1}|$

$$\le \|U^{-1}\|\ \|T-U\|\ \ \|T^{-1}\|$$

$$\le (\alpha-\beta)^{-1}\,\beta\,\frac{1}{\alpha}$$

This shows that f is continuous since $\beta\to 0$ as $U\to T$.

Differentiation in R^n

Definition. Let A be an open subset of R^n, $x\in A$ and f a mapping of A into R^m. If there exists a linear transformation T of R^n into R^m such that

$$\lim_{h\to 0}\frac{|f(x+h)[f(x)-Th|}{|h|} = 0 \qquad \ldots(1)$$

Then f is said to differentiable at x and we rite

$$f'(x) = T \qquad \ldots(2)$$

Uniqueness of the derivatives

Theorem. Let A be an open subset of R^n, $x \in A$ and f a mapping of A into R^m. If f is differentiable with $T = T_1$ and $T = T_2$, where $T_1, T_2 \in L(R^n, R^m)$, then $T_1 = T_2$.

Proof. Let $U = T_1 - T_2$ we have

$$|Uh| = |(T_1 - T_2)h| = |(T_1h - T_2h)|$$

$$= |T_1h - f(x+h) + f(x) + f(x+h) - f(x) - T_2h|$$

$$= |T_1h - f(x+h) + f(x)| + |f(x+h) - f(x) - T_2h|$$

$$= |f(x+h) - f(x) - T_1h| + |f(x+h) - f(x) - T_2h|$$

$$\therefore \qquad \left|\frac{Uh}{h}\right| \le \frac{|f(x+h)--f-(x)-T_1h|}{|h|} + \frac{|f(x+h)-f(x-T_2h|}{|h|}$$

$\rightarrow 0$ as $h \to 0$ by (1) differentiability of f.

For fixed h = 0, it follows that

$$\frac{|U(U_n)|}{|t_n|} \rightarrow 0 \text{ as } t \to 0 \qquad \ldots(1)$$

Linearity of U shows that U(th) = tUh so that the left hand side of (1) is independent of t.

Thus for all $h \in R^n$, we have Uh = 0 $\Rightarrow (T_1 - T_2)h = 0 \Rightarrow T_1h - T_2h = 0$

$\Rightarrow \quad T_1h = T_2h \Rightarrow T_1 = T_2$

The chain Rule

Theorem. Suppose E is an open subset of R^n, f maps E into R^m, f is differentiable at $x_0 \in E$, g maps an open set containing f(E) into R^k, and g is differentiable at $f(x_0)$. Then the mapping F of E into R^k defined by

$F(x) = g(f(x))$ is differentiable at x_0 and

$F'(x_0) = g'(f(x_0))\, f'(x_0)$ product of two linear transformations.

Proof. Let $y_0 = f(x_0)$, $T = f'(x_0)$, $U = f'(y_0)$ and define

$$u(x) = f(x) - f(x_0) - T(x - x_0)$$

$$u(y) = g(y) - g(y_0) - U(y - y_0)$$

$$r(x) = F(x) - F(x_0) - UT\,(x - x_0)$$

We want to prove that $F'(x_0) = Ut$, that is,

$$\lim_{x \to x_0} \frac{|r(x)|}{|x - x_0|} = 0$$

The definition of F, r at y_0 show that

$$r(x) = g(f(x)) - g(y_0) - UT(x - x_0)$$

Now $\quad UT\,(x - x_0) = U(T(x - x_0)) = U(f(x) - y_0 - f(x) + f(x_0) + T(x - x_0))$

$= U(f(x) - y_0) - U(f(x) - f(x_0) - T(x - x_0))$ by linearity of U.

Hence $r(x) = [g(f(x)] - g(y_0) - U(f(x) - y_0)] + [U(f(x) - f(x_0) - T(x - x_0))]$

$= u(f(x)) + U_n(x)$

By definition of U and T, we have

We have $\frac{|u(y)|}{|y-y_0|} \to 0$ as $y \to y_0$ and $\frac{|u(x)|}{|x-x_0|} \to 0$ as $x \to x_0$.

This means that for a given ε>0, we can find η > 0 and

δ > 0 such that $|f(x) - f(x_0)| < \eta$, $|u(x)| \leq \varepsilon|x - x_0|$ if $|x - x_0| < \delta$.

It follows that

$$|uf(x)| \leq \varepsilon\, |f(x) - f(x_0)| = \varepsilon\, |u(x) + T(x - x_0)|$$

$$\leq \varepsilon\, u(x)| + \varepsilon\, |T(x - x_0)|$$

$$\leq \varepsilon^2\, |x - x_0| + \varepsilon\, ||T||\ \ |x-x_0| \qquad \ldots(2)$$

and

$$|Un(x)| \leq ||U||\ \ |u(x)| \leq \varepsilon||U||\, |x - x_0| \qquad \ldots(3)$$

$$\text{if } |x - x_0| < \delta$$

Hence $\frac{|r(x)|}{|x-x_0|} = \frac{|f(x) - U u(x)|}{|x-x_0|}$

$$\leq \frac{|uf(x)|}{|x-x_0|} + \frac{|U_n(x)|}{|x-x_0|}$$

$\leq \varepsilon^2 + \varepsilon||T|| + \varepsilon||U||$ by (2) and (3).

It follows that $\frac{|r(x)|}{|x-x_0|} \to 0$ as $x \to x_0$.

Proof. Put $y_0 = f(x_0)$, $A = f'(x_0)$, $B = g'(y_0)$ and define

$$u(h) = f(x_0 + h) - f(x_0) - Ah,$$

$$V(k) = g(y_0 + k) - g(y_0) - Bk \qquad \forall\ h \in R^n,\ k \in R^m$$

for which $f(x_0 + h)$ and $g(y_0 + k)$ are defined.

Then $|u(h)| = \varepsilon|h|$, $|v(k)| = \eta|k|$ where $\varepsilon \to 0$ as $h \to 0$ and $\eta \to 0$ as $k \to 0$ (same fid g are differentiable)

Given h, put $k = f(x_0 + h) - f(x_0)$. Then

$$|k| = |Ah + u(h)| \le [\|A\| + \varepsilon]\ |h| \qquad \ldots(2)$$

(by definition of F(x))

and

$$F(x_0 + h) - F(x_0) - Bah = g(y_0 + k) - g(y_0) - Bah$$

$$= B(k - Ah) + v\,(k)$$

Hence (1) and (2) imply that for $h \neq 0$,

$$\frac{|F(x_0 + h) - F(x_0) - BAh|}{|h|} \le \|B\|\ \varepsilon + [\|A\| + \varepsilon]\eta$$

Let $h \to 0$, then $\varepsilon \to 0$. Also $k \to 0 \ \Rightarrow\ \eta \to 0$. It follows that $F'(x_0) = BA$.

MAL-512: M. Sc. Mathematics (Real Analysis)

Lesson No. IV **Written by Dr. Nawneet Hooda**

Lesson: Functions of several variables **Vetted by Dr. Pankaj Kumar**

PARTIAL DERIVATIVES

Let f be a function of several variables ,then the ordinary derivative of f with respect to one of the independent variables, keeping all other independent variables constant is called the *partial derivative* . Partial derivative of f(x, y) with respect to x is generally denoted by $\partial f/\partial x$ or f_x or $f_x(x, y)$.Similarly those with respect to y are denoted by $\partial f/\partial y$ or f_y or $f_y(x, y)$.

$$\therefore \quad \frac{\partial f}{\partial x} = \lim_{\delta x \to 0} \frac{f(x+\delta x, y) - f(x, y)}{\delta x}$$

when these limits exist.

The partial derivatives at a particular point (a, b) are definined as (in case the limits exist)

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

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

Example . For the function f(x,y),where

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & (x, y) \neq (0,0) \\ 0 \quad , & (x, y) = (0,0) \end{cases}$$

both the partial derivatives exist at (0, 0) but the function is not continuous at(0, 0) .

Solution. Putting y = mx, we see that

$$\lim_{x \to 0} f(x, y) = \frac{m}{1+m^2}$$

so that the limit depends on the value of m, i.e. on the path of approach and is different for the different paths followed and therefore does not exist. Hence the function f(x, y) is not continuous at (0, 0). Again

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

$$f_y(0, 0) = \lim_{k \to 0} \frac{f(0.0+k) - f(0,0)}{k}$$

$$= \lim_{k \to 0} \frac{0}{k} = 0$$

Definition. Let (x, y), (x + δx, y + δy) be two neighbouring points and let

$$\delta f = f(x + \delta x, y + \delta y) - f(x, y)$$

The function f is said to be differentiable at (x, y) if

$$\delta f = A\ \delta x + B\ \delta y + \delta x\ \phi(\delta x, \delta y) + \delta y\ \psi(\delta x, \delta y) \qquad \ldots(1)$$

where A and B are constants independent of δx, δy and φ, ψ are functions of δx, δy tending to zero as δx, δy tend to zero simultaneously.

Also, A δx + B δy is then called the differential of f at (x, y) and is denoted by df. Thus

$$df = A\ \delta x + B\ \delta y$$

From (1) when (δx, δy) → (0, 0), we get

$$f(x + \delta x, y + \delta y) - f(x, y) \to 0$$

or

$$f(x + \delta x, y + \delta y) \to f(x, y)$$

⇒ The function f is continuous at (x, y)

Again, from (1), when δy = 0 (i.e., y remains constant)

$$\delta f = A\ \delta x + \delta x\ \phi(\delta x, 0)$$

Dividing by δx and proceeding to limits as δx → 0, we get

$$\frac{\partial f}{\partial x} = A$$

Similarly

$$\frac{\partial f}{\partial y} = B$$

Thus the constants A and B are respectively the partial derivatives of f with respect to x and y.

Hence a function which is differentiable at a point possesses the first order partial derivatives thereat. Again the differential of f is given by

$$df = A\delta x + B\ \delta y = \frac{\partial f}{\partial x}\partial x + \frac{\partial f}{\partial y}\delta y$$

Taking f = x, we get dx = δx.

Similarly taking f = y, we obtain dy = δy.

Thus the differentials dx, dy of x, y are respectively δx and δy, and

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = f_x\,dx + f_y\,dy \qquad \ldots(2)$$

is the differential of f at (x, y).

Note 1. If we replace δx, δy, by h, k in equation (1) we say that the function is differentiable at a point (a, b) of the domain of definition if df can be expressed as

$$df = f(a + h, b + k) - f(a, b)$$

$$= Ah + Bk + h\phi(h, k) + k\psi(h, k) \qquad \ldots(3)$$

where $A = f_x$, $B = f_y$ and φ, ψ are functions of h, k tending to zero as h, k tend to zero simultaneously.

Example 12. Let

$$f(x, y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2}, & (x, y) \neq (0,0) \\ 0, & (x, y) = (0,0) \end{cases}$$

Put x = r cos θ, y = r sin θ,

$$\therefore \qquad \left|\frac{x^3 - y^3}{x^2 + y^2}\right| = |r(\cos^3\theta - \sin^3\theta)| \leq 2|r| = 2\sqrt{x^2 + y^2} < \varepsilon,$$

if

$$x^2 < \frac{\varepsilon^2}{8}, \; y^2 < \frac{\varepsilon^2}{8}$$

or, if

$$|x| < \frac{\varepsilon}{2\sqrt{2}} \, | y |< \frac{\varepsilon}{2\sqrt{2}}$$

$$\therefore \qquad \left| \frac{x^3 - y^3}{x^2 + y^2} - 0 \right| < \varepsilon, \text{ when } | x |< \frac{\varepsilon}{2\sqrt{2}}, | y |< \frac{\varepsilon}{2\sqrt{2}}$$

$$\Rightarrow \qquad \lim_{(x,y)\to(0,0)} \frac{x^3 - y^3}{x^2 + y^2} = 0$$

$$\Rightarrow \qquad \lim_{(x,y)\to(0,0)} f(x, y) = f(0, 0)$$

Hence the function is continuous at (0, 0).

Again,

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

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

Thus the function possesses partial derivatives at (0, 0).

If the function is differentiable at (0, 0), then by definition

$$df = f(h, k) - f(0, 0) = Ah + Bk + h\phi + k\psi \qquad \ldots(1)$$

when A and B are constants ($A = f_x(0, 0) = 1$, $B = f_y(0, 0) = -1$) and ϕ, ψ tend to zero as $(h, k) \to (0, 0)$.

Putting $h = \rho \cos \theta$, $k = \rho \sin \theta$, and dividing by ρ, we get

$$\cos^3\theta - \sin^3\theta = \cos\theta + \phi \cos \theta + \psi \sin \theta \qquad \ldots(2)$$

For arbitrary $\theta = \tan^{-1} (h/k)$, $\rho\to 0$ implies that $(h, k) \to (0, 0)$. Thus we get the limit,

$$\cos^3\theta - \sin^3\theta = \cos \theta - \sin \theta$$

or

$$\cos \theta \sin \theta(\cos \theta - \sin \theta) = 0$$

which is plainly impossible for arbitrary θ.

Thus the function is not differentiable at the origin.

Example 13. Show that the function f, where

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

is continuous, possesses partial derivatives but is not differentiable at the origin.

Solution. It may be easily shown that f is continuous at the origin and

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

If the function is differentiable at the origin then by definition

$$df = f(h, k) - f(0, 0) = Ah + Bk + h\phi + k\psi \qquad \ldots(1)$$

where $A = f_x(0, 0) = 0$; $B = f_y(0, 0) = 0$, and ϕ, ψ tend to zero as $(h, k) \to (0, 0)$.

$$\therefore \qquad \frac{hk}{\sqrt{h^2 + k^2}} = h\phi + k\psi \qquad \ldots(2)$$

Putting k = mh and letting $h \to 0$, we get

$$\frac{m}{\sqrt{1 + m^2}} = \lim_{h \to 0} (\phi + m\psi) = 0$$

which is impossible for arbitrary m.

Hence the function is not differentiable at (0, 0).

Example II. Prove that the function

$$f(x, y) = \sqrt{|xy|}$$

is not differentiable at the point (0, 0), but f_x and f_y both exist at the origin .

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

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

If the function is differentiable at (0, 0), then by definition

$$f(h, k) - f(0, 0) = 0h + 0k + h\phi + k\psi$$

where ϕ and ψ are functions of h, k and tend to zero as $(h, k) \to (0, 0)$.

Putting $h = \rho \cos \theta$, $k = \rho \sin \theta$ and dividing by ρ, we get

$$|\cos\theta \sin\theta|^{1/2} = \phi\,\theta + \psi \sin\theta$$

Now for arbitrary θ, $\rho \to 0$ implies that $(h, k) \to (0, 0)$.

Taking the limit as $\rho \to 0$, we get

$$|\cos\theta \sin\theta|^{1/2} = 0.$$

which is impossible for all arbitrary θ.

Example 14. The function f, where

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

is differentiable at the origin.

Solution. It is easy to show that

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

Also when $x^2 + y^2 \neq 0$,

$$|f_x| = \frac{|x^4y + 4x^2y^3 - y^5|}{(x^2 + y^2)^2} \leq \frac{6(x^2 + y^2)^{5/2}}{(x^2 + y^2)^2} = 6(x^2 + y^2)^{1/2}$$

Evidently

$$\lim_{(x,y)\to(0,0)} f_x(x, y) = 0 = f_x(0, 0)$$

Thus f_x is continuous at (0, 0) and $f_y(0, 0)$ exists,

$\Rightarrow$ f is differentiable at (0, 0)

PARTIAL DERIVATIVES OF HIGHER ORDER

If a function f has partial derivatives of the first order at each point (x, y) of a certain region, then f_x, f_y are themselves functions of x, y and may also possess partial derivatives. These are called second order partial derivatives of f and are denoted by

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx} = f_{x^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy} = f_{y^2}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)=\frac{\partial^2 f}{\partial x \partial y} = f_{xy}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)=\frac{\partial^2 f}{\partial y \partial x} = f_{yx}$$

Thus(in case the limits exist)

$$f_{xx}(a, b) = \lim_{h\to 0}\frac{f_x(a+h,b)-f_x(a,b)}{h}$$

$$f_{xy}(a, b) = \lim_{h\to 0}\frac{f_y(a+h,b)-f_y(a,b)}{h}$$

$$f_{yx}(a, b) = \lim_{k\to 0}\frac{f_x(a,b+k)-f_x(a,b)}{k}$$

$$f_{yy}(a, b) = \lim_{k\to 0}\frac{f_y(a,b+k)-f_y(a,b)}{k}$$

Change in the Order of Partial Derivation

Consider an example to show that f_{xy} may be different from f_{yx}.

Example . Let

$$f(x, y) = \frac{xy(x^2-y^2)}{x^2+y^2}, (x, y) \neq (0, 0),$$

$f(0, 0) = 0$, then at the origin $f_{xy} \neq f_{yx}$.

Solution. Now

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

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

$$f_y(h, 0) = \lim_{k\to 0}\frac{f(h,k)-f(h,0)}{k} = \lim_{k\to 0}\frac{hk(h^2-k^2)}{k\cdot(h^2+k^2)} = h$$

$$\therefore \quad f_{xy} = \lim_{h\to 0}\frac{h-0}{h} = 1$$

Again

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

But

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

$$f_x(0, k) = \lim_{h\to 0}\frac{f(h,k) - f(0,k)}{h} = \lim_{h\to 0}\frac{hk(h^2 - k^2)}{h(h^2 + k^2)} = -k$$

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

$$\therefore \quad f_{xy}(0, 0) \neq f_{yx}(0, 0)$$

Sufficient Conditions for the Equality of f_{xy} and f_{yx}

We have two theorems to show that $f_{xy} = f_{yx}$ at a point.

Theorem 3 (**Young's theorem**). If f_x and f_y are both differentiable at a point (a, b) of the domain of definition of a function f, then

$$f_{xy}(a, b) = f_{yx}(a, b)$$

Proof. We prove the theorem by taking equal increment h both for x and y and calculating $\phi(h, h)$ in two different ways.

Let (a + h, b + h) be a point of this neighbourhood. Consider

$$\phi(h, h) = f(a + h, b + h) - f(a + h, b) - f(a, b + h) + f(a, b)$$

$$G(x) = f(x, b + h) - f(x, b)$$

so that

$$\phi(h, h) = G(a + h) - G(a) \qquad \ldots(1)$$

Since f_x exists in a neighbourhood of (a, b), the function G(x) is derivable in (a, a + h) and therefore by Lagrange's mean value theorem, we get from (1),

$$\phi(h, h) = hG'(a + \theta h), \quad 0 < \theta < 1$$

$$= h\{f_x(a + \theta h, b + h) - f_x(a + \theta h, b)\} \qquad \ldots(2)$$

Again, since f_x is differentiable at (a, b), we have

$$f_x(a + \theta h, b + h) - f_x(a, b) = \theta h_{xx}(a, b) + hf_{yx}(a, b)$$
$$+ \theta h\phi_1(h, h) + h\psi_1(h, h) \qquad \ldots(3)$$

and

$$f_x(a + \theta h, b) - f_x(a, b) = \theta h f_{xx}(a, b) + \theta h \phi_2(h, h) \qquad \ldots(4)$$

where ϕ_1, ψ_1, ϕ_2 all tend to zero as $h \to 0$.

From (2), (3), (4), we get

$$\phi(h, h)/h^2 = f_{yx}(a, b) + \theta\phi_1(h, h) + \psi_1(h, h) - \theta\phi_2(h, h) \qquad \ldots(5)$$

Similarly, taking

$$H(y) = f(a + h, y) - f(a, y)$$

we can show that

$$\phi(h, h)/h^2 = f_{xy}(a, b) + \phi_3(h, h) + \theta'\psi_2(h, h) - \theta'\psi_2(h, h) \qquad \ldots(6)$$

where ϕ_3, ψ_2, ψ_3 all tend to zero as $h \to 0$.

On taking the limit as $h \to 0$, we obtain from (5) and (6)

$$\lim_{h \to 0} \frac{\varphi(h, h)}{h^2} = f_{xy}(a, b) = f_{yx}(a, b)$$

Theorem. (**Schwarz's theorem**). If f_y exists in a certain neighbourhood of a point (a, b) of the domain of definition of a function f, and f_{yx} is continuous at (a, b), then $f_{xy}(a, b)$ exists, and is equal to $f_{yx}(a, b)$.

Proof. Let (a + h, b + k) be a point of this neighbourhood of (a, b).

Take

$$\phi(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)$$

$$G(x) = f(x, b + k) - f(x, b)$$

so that

$$\phi(h, k) = G(a + h) - g(a) \qquad \ldots(1)$$

Since f_x exists in a neighbourhood of (a, b), the function g(x) is derivable in (a, a + h), and therefore by Lagrange's mean value theorem, we get from (1)

$$\phi(h, k) = hG'(a + \theta h), \qquad 0 < \theta < 1$$

$$= h\{f_x(a + \theta h, b + k) - f_x(a + \theta h, b)\} \qquad \ldots(2)$$

Again, since f_{yx} exists in a neighbourhood of (a, b), the function f_x is derivable with respect to y in (b, b + k), and therefore by Lagrange's mean value theorem, we get from (2)

$$\phi(h, k) = hkf_{yx}(a + \theta h, b + \theta' k), \qquad 0 < \theta' < 1$$

or

$$\frac{1}{h}\left\{\frac{f(a+h,b+k)-f(a+h,b)}{k}-\frac{f(a,b+k)-f(a,b)}{k}\right\}$$

$$= f_{yx}(a + \theta h, b + \theta' k)$$

Taking limits when $k \to 0$, since f_y and f_{yx} exist in a neighbourhood of (a, b), we get

$$\frac{f_y(a+h,b)-f_y(a,b)}{h} = \lim_{k\to 0} f_{yx}(a + \theta h, b + \theta' k)$$

Again, taking limits as $h \to 0$, since f_{yx} is continuous at (a, b), we get

$$f_{xy}(a, b) = \lim_{h\to 0}\lim_{k\to 0} f_{yx}(a + \theta h, b + \theta' k) = f_{yx}(a, b)$$

2. If the conditions of Young's or Schwarz's theorem are satisfied then $f_{xy} = f_{yx}$ at a point (a, b). But if the conditions are not satisfied, we cannot draw any conclusion regarding the equality of f_{xy} and f_{yx}. Thus the conditions are sufficient but not necessary.

Example . Show that for the function

$$f(x, y) = \begin{cases} \frac{x^2y^2}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

Solution. Here $f_{xy}(0, 0) = f_{yx}(0, 0)$ since

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

Similarly, $f_y(0, 0) = 0$.

Also, for $(x, y) \neq (0, 0)$.

$$f_x(x, y) = \frac{(x^2+y^2).2xy^2 - x^2y^2.2x}{(x^2+y^2)^2} = \frac{2xy^4}{(x^2+y^2)^2}$$

$$f_y(x, y) = \frac{2x^4y}{(x^2+y^2)^2}$$

Again

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

and

$$f_{xy}(0, 0) = 0, \text{ so that } f_{xy}(0, 0) = f_{yx}(0, 0)$$

For $(x, y) \neq (0, 0)$, we have

$$f_{yx}(x, y) = \frac{8xy^3(x^2 + y^2)^2 - 2xy^4.4y(x^2 + y^2)}{(x^2 + y^2)^4}$$

$$= \frac{8x^3y^3}{(x^2 + y^2)^3}$$

and it may be easily shown (by putting $y = mx$) that

$$\lim_{(x,y)\to(0,0)} f_{yx}(x, y) \neq 0 = f_{yx}(0, 0)$$

so that f_{yx} is not continuous at (0, 0), i.e., the conditions of Schwarz's theorem are not satisfied.

We now show that the conditions of Young's theorem are also not satisfied.

$$f_{xx}(0, 0) = \lim_{x \to 0} \frac{f_x(x,0) - f_x(0,0)}{x} = 0$$

Also f_x is differentiable at (0, 0) if

$$f_x(h, k) - f_x(0, 0) = f_{xx}(0, 0).\ H + f_{yx}(0, 0).\ K + h\phi + k\psi$$

or

$$\frac{2hk^4}{(h^2 + k^2)^2} = h\phi + k\psi$$

where ϕ, ψ tend to zero as $(h, k) \to (0, 0)$.

Putting $h = \rho \cos \theta$ and $k = \rho \sin \theta$, and dividing by ρ, we get

$$2 \cos \theta \sin^4 \theta = \cos \theta.\phi + \sin \theta\psi$$

and $(h, k)\to(0, 0)$ is same thing as $\rho\to 0$ and θ is arbitrary. Thus proceeding to limits, we get

$$2 \cos \theta \sin^4\theta = 0$$

which is impossible for arbitrary θ,

$\Rightarrow$ f_x is not differentiable at (0, 0)

Similarly, it may be shown that f_y is not differentiable at (0, 0).

Thus the conditions of Young's theorem are also not satisfied but, as shown above,

$$f_{xy}(0, 0) = f_{yx}(0, 0)$$

TAYLOR'S THEOREM

Theorem. If f(x, y) is a function possessing continuous partial derivatives of order n in any domain of a point (a, b), then there exists a positive number, $0 < \theta < 1$, such that

$$f(a+h,b+k)=f(a,b)+\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)f(a,b)+\frac{1}{2!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^2 f(a, b)$$

$$+\ldots+\frac{1}{(n-1)!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^{n-1} f(a, b) + R_{n,}$$

where $R_n = \frac{1}{n!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^n f(a + \theta h, b + \theta k), 0 < \theta < 1.$

Proof. Let $x = a + th$, $y = b + tk$, where $0 \le t \le 1$ is a parameter and

$$f(x, y) = f(a + th, b + tk) = \phi(t)$$

Since the partial derivatives of f(x, y) of order n are continuous in the domain under consideration, $\phi^x(t)$ is continuous in [0, 1]and also

$$\phi'(t) = \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt} = h\frac{\partial f}{\partial x}+k\frac{\partial f}{\partial y} = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)f$$

$$\phi''(t) = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^n f$$

$$\phi^{(n)}(t) = \left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^n f$$

therefore by Maclaurin's theorem

$$\phi(t) = \phi(0) + t\phi'(0) + \frac{t^2}{2!}\varphi''(0) + \ldots + \frac{t^{n-1}}{(n-1)!}\varphi^{(n-1)}(0) + \frac{t^n}{n}\varphi^{(n)}(\theta t),$$

where $0<\theta<1$.

Now putting t = 1, we get

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2!}\phi''(0) + \ldots + \frac{t^{n-1}}{(n-1)!}\varphi^{(n-1)}(0) + \frac{t^n}{n!}\varphi^{(n)}(0)$$

But $\phi(1) = f(a + h, b + k)$, and $\phi(0) = (a, b)$

$$\phi'(0) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a, b)$$

$$\phi''(0) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b)$$

$$\phi^{(n)}(\theta) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a + \theta h, b + \theta k)$$

$$\therefore f(a + h, b + k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a, b)$$

$$+ \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \ldots..$$

$$+ \frac{1}{(n-1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n-1} f(a, b) + R_n,$$

where $R_n = \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a + \theta h, b + \theta k), 0 < \theta < 1.$

R_n is called the remainder after n terms and theorem, Taylor's theorem with remainder or Taylor's expansion about the point (a, b)

If we put $a = b = 0$; $h = x$, $k = y$, we get

$$f(x, y) = f(0, 0) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(0, 0)$$

$$+ \frac{1}{2!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f(0, 0) + \ldots.$$

$$+ \frac{1}{(n-1)!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^{n-1} f(0, 0) + R_n$$

Where $R_n = \frac{1}{n!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^n f(\theta x, \theta y), 0 < \theta < 1.$

Note. This theorem can be stated in another from,

$$f(x, y) = f(a, b) + \left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right] f(a, b)$$

$$+\frac{1}{2!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^2 f(a, b) + \ldots$$

$$+\frac{1}{(n-1)!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^{n-1} f(a, b) + R_n,$$

where $R_n = \frac{1}{n!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^n f(a + (x-a)\theta, b + (y-b)\theta),$

$0 < \theta < 1$. It is called the Taylor's expansion of f (x, y) about the point (a, b) in powers of x – a and y – b,

Example 30. Expand $x^2y + 3y - 2$ in powers of x – 1and y + 2.In Taylor's expansion take a = 1, b = – 2. Then

$f(x, y) = x^2y + 3y - 2,$	$f(1,-2) = -10$
$f_x(x, y) = 2xy,$	$f_x(1,-2) = -4$
$f_y(x, y) = x^2 + 3,$	$f_y(1,-2) = 4$
$f_{xx}(x, y) = 2y,$	$f_{xx}(1,-2) = 4$
$f_{xy}(x, y) = 2x$,	$f_{xy}(1,-2) = 2$
$f_{yy}(x, y) = 0,$	$f_{yy}(1,-2) = 0$
$f_{xxx}(x, y) = 0 = f_{yyy}(x, y),$	$f_{yxx}(1,-2) = 2 = f_{xxy}(1, -2)$

All higher derivatives are zero.

$$\therefore x^2y + 3y - 2 = -2 = -10 - 4(x-1) + 4(y+2) + \tfrac{1}{2}[-4(x-1)^2$$

$$+4(x-1)(y+2)] + \frac{1}{3!}\, 3(x-1)^2 (y+2)(2) + 0$$

$$= -10 - 4(x-1) + 4(y+2) - 2(x-1)^2$$

$$+ 2(x-1)(y+2) + (x-1)^2 (y+2)$$

Example 31. If f(x, y) = $\sqrt{|xy|}$ prove that Taylor's expansion about the point (x, y) is not valid in any domain which includes the origin.

Solution.

$f_x(x, y) = 0 = f_y(0, 0)$

$$f_x(x, y) = \begin{cases} \frac{1}{2}\sqrt{|y \backslash x|}, & x > 0 \\ -\frac{1}{2}\sqrt{|y \backslash x|}, & x < 0 \end{cases}$$

$$f_x(x, y) = \begin{cases} \frac{1}{2}\sqrt{|x \backslash y|}, & y > 0 \\ -\frac{1}{2}\sqrt{|x \backslash y|}, & y < 0 \end{cases}$$

$$\therefore f_x(x, x) = f_y(x, x) = \begin{cases} \frac{1}{2}, & x > 0 \\ -\frac{1}{2}, & x < 0 \end{cases}$$

Now Taylor's expansion about (x, x) for n = 1, is

$f(x + h, x + h) = f(x, x) + h\,[(f_x(x + \theta h, x + \theta h) + f_y(x + \theta h, x + \theta h)]$

or

$$|x + h| = \begin{cases} |x| + h, & \text{if } x + \theta h > 0 \\ |x| - h, & \text{if } x + \theta h < 0 \\ |x|\ , & \text{if } x + \theta h = 0 \end{cases} \qquad (1)$$

If the domain (x, x; x + h, x + h) includes the origin, then x and x + h must be of opposite signs, that is either

$|x + h| = x + h,\ |x| = -x$

or

$|x + h| = -(x + h),\ |x| = x$

But under these conditions none of the in equalities (1) holds. Hence the expansion is not valid.

Definition. Let f be differentiable mapping of an open subset A of R^n into R^m. Then f is said to be continuously differentiable in A if f' is a continuous mapping of A into $L(R^n, R^m)$ and write

$$f \in C'(A)$$

To be precise f is a C' mapping in A if to each $x\in A$ and each $\in>0$, there exists $\delta>0$ such that

$$y\in A, |y-x| < \delta \;\Rightarrow\; \|f'(y) - f'(x)\| < \in.$$

The Inverse function theorem

This theorem asserts, roughly speaking that if f is a C' mapping, then f is invertible in a neighbourhood of any point x at which the linear transformation $f'_2(x)$ is invertible.

Theorem. Suppose A is an open subset of R^n, f is a C' mapping of A into R^n, $f'(a)$ is invertible for some $a\in A$ and $b = f(a)$. Then (i) there exists G and H in R^n such that

$$a \in G, b \in H.$$

f is one-one on G and f(G) = H.

(ii) if g is the inverse of f (which exists by (i)) defined on H by

$$g(f(x)) = x \qquad (x\in G),$$

then $\qquad g\in C'(H).$

Proof. (i) Let $f'(a) = T$ let λ be so chosen that

$$4\lambda \|T^{-1}\| = 1.$$

Since f' is a continuous mapping of A into $L(R^n, R^m)$, there exists an open ball G with centre a such that

$$x \in G \;\Rightarrow\; \|f'(x) - T\| < 2\lambda. \qquad \ldots(1)$$

Suppose $x\in G$ and $x+h \in G$. Define

$$F(t) = f(x + th) - t\,Th \;(0 \le t \le 1) \qquad \ldots(2)$$

Since G is convex (see example 2 of § 2, ch. 11), we have

$x + th \in G$ if $0 \le t \le 1$.

Also $\quad |F'(t)| = |f'(x + th)h - Th| = [|f'(x + th) - T \mid h]$

$$\le \|f'(x + th) - T\| \, |h| < 2\lambda \, |h| \text{ by } (1) \qquad \ldots(3)$$

Since T is invertible, we have

$$|h| = |T^{-1}\, Th| \le \|T^{-1}\| \, |Th| = \frac{1}{4\lambda} f| \, Th| \qquad \ldots(4)$$

$[\because \; 4\lambda \, \|T^{-1}\| = 1]$

From (3) and (4), we have

$$|F'(t)| < \frac{1}{2} \, |Th|, (0 \le t \le 1| \quad (0 \le t \le 1) \qquad \ldots(5)$$

Also,

$$|F(1) - F(0)| \le (1 - 0) \, |F'(t_0)| \text{ for some } t_0' \in (0, 1)$$

$$\le \frac{1}{2} \, |Th| \text{ by } (5) \qquad \ldots(6)$$

Now (2) and (6) give

$$|f(x + h) - f(x) - Th| \le \frac{1}{2} \, |Th| \qquad \ldots(7)$$

Now $\quad \frac{1}{2}|Th| \ge |Th - (f(x + h) - f(x))|$

$$\ge |Th| - |f(x + h) - f(x)|$$

or $\quad |f(x + h) - f(x)| \ge \frac{1}{2} \, |Th| \ge 2\lambda \, |h|$ by (4) $\qquad \ldots(8)$

Also, (7) and (8) hold whenever $x \in G$ and $x + h \in G$.

In particular, it follows from (8) that f is one-one on G. For if $x, y \in G$, then

$$f(x) = f(y) \;\Rightarrow\; f(x) - f(y) = 0$$

$$\Rightarrow\; 0 = |f(x) - f(y)| \ge 2\lambda \, |x - y| \text{ by } (8)$$

$$\Rightarrow\; |x - y| = 0$$

[Θ $2\lambda|x - y|$ cannot be negative]

$\Rightarrow x - y = 0 \Rightarrow x = y.$

We now prove that f[G] is an open subset of R^n. Let y_0 be an arbitrary point of f[G]. Then $y_0 = f(x_0)$ for some $x_0 \in G$. Let S be an open ball with centre x_0 and radius $r > 0$ such that $\overline{S} \subset G$. Then $(x_0) \in f[X] \subset f[G]$. We shall show that f[S] contains the open ball with centre at $f(x_0)$ and radius λr. This will prove that f[G] contains a neighbourhood of $f(x_0)$ and this in turn will prove that f[G] is open.

Fix y so that $|y - f(x_0)| < \lambda r$ and define

$$\phi(x) = |y - f(x)| \ (x \in \overline{S}).$$

$|x - x_0| = r$, then (8) shows that

$$2\lambda r \leq |f(x) - f(x_0)| = |f(x) - y + y - f(x_0)|$$

$$\leq |f(x) - y| + |y - f(x_0)| = \phi(x) + \phi(x_0) < \phi(x_0) + \lambda r$$

This shows that

$$\phi(x_0) < \lambda r < \phi(x) \ (|x - x|) = r \qquad \ldots(9)$$

Since ϕ is continuous and S is compact, there exists $x^* \in \overline{S}$ such that

$$\phi(x^*) < \phi(x) \ 0 \text{for } x \in \overline{S} \qquad \ldots(10)$$

By (9), $x^* \in S$.

Put $w = y -(x^*)$. Since T is invertible, there exists $h \in R^*$ such that $Th = W$. Let $t \in [0, 1)$ be chosen so small that

$$x^* + th \in S$$

Then $$|f(x_0^*) - y + Tth| = |-w + tTh|$$

$$= |-w + tw| = (1 - t)|w| \qquad \ldots(11)$$

Also (7) shows that

$$|f(x^* + th) - f(x^*) - Tth| \leq \frac{1}{2}|Tth|$$

$$= \frac{1}{2}|tTh| \frac{1}{2}|tw|. \qquad \ldots(12)$$

Now $\phi(x^* + th) = |y - f(x^* + th)| = |f(x^* + th) - y|$

$$= |f(x^* + th) - f(x^*) - Tth + f(x^*) - y + Tth|$$

$$\le |f(x^* + th) - f)(x^*) - Tth| + | f(x^*) - y + Tth|$$

$$\le (1 - t) | w | + \frac{1}{2} |tw| \text{ by (1) and (12)}$$

$$= (1 - \frac{1}{2}t)| w | = (1 - \frac{1}{2}t) \phi (x^*) \qquad \ldots(13)$$

Definition of ϕ shows that $\phi(x) \ge 0$. We claim that $\phi(x) \ge 0$. We claim that $\phi(x^*) > 0$ ruled out. For if $\phi(x^*) > 0$, then (13) shows that

$$\phi(x^* + th)) < \phi(x^*), \text{ since } 0 < t < 1.$$

But this contradicts (10). Hence we must have $\phi(x^*) = 0$ which implies that $f(x^*) = y$ so that $y \in f[S]$ since $x^* \in S$. This shows that the open sphere with centre at $f(x_0)$ and radius λr is contained in f(S).

We have thus proved that every point of f[G] has a neighbourhood contained in f[G] and consequently f[G] is an open-subset of R^n By setting H = f[G], part (i) of the theorem is proved.

(ii) Take $y \in H$, $y + k \in \in H$ and put

$$x = g(y),\ h = g(y + k) - g(y)$$

By hypothesis, $T = f'(a)$'s invertible and $f'(x) \in L(R^n)$. Also by (1).

$$||f'(x) - T|| < 2\lambda < 4\lambda = \frac{1}{|| T^{-1} ||} \text{ (see the choice of } \lambda)$$

Hence , $f'(x)$ has an inverse, say U

Now $\quad k = f(x + h) - f(x) = f'(x)\, h + r(h) \qquad \ldots(14)$

where $\quad |r(h)| / |h| \to 0$ as $h \to 0$.

Applying U to (14), we obtain

$$Uk = U(f'(x)\,h + r(h)) = Uf'(x)h + Ur(h) = h + Ur(h)$$

[Θ U is the inverse of f′(x) implies Uf′(x) h = h]

or $$h = Uk - U(r(h))$$

or $$g(y + k) - g(y) = Uk - U(r(h)) \qquad \ldots(15)$$

By (8), $2\lambda\,|h| \le |k|$. Hence $h\to 0$ if $k \to 0$

(which shows, incidentally, that g is continuous at y),

and $$\frac{|U(r(h))|}{|k|} \le \frac{\|U\|\ \ |r(h)|}{2\lambda\,|h|} \to 0 \text{ as } k\to 0 \qquad \ldots(16)$$

Comparing (15) and (16), we see that g is differentiable at y and that

$$g'(y) = U = [f'(x)]^{-1} = [f'(g(y))]^{-1}, \quad (y \in H) \qquad \ldots(17)$$

Also g is a continuous mapping of H onto G, f′ is continuous mapping of G into the set *C* of all invertible elements of $L(R^n)$, and inversion is a continuous mapping of *C* onto *C* , These facts combined with (17) imply that $g \in C'(H)$.

Example. Let f(x) = x, $g(x) = x^2$.

Evaluate $\int_0^1 f dg$.

Solution. Since f is continuous and g is non-increasing on [0, 1], it follows that $\int_0^1 f dg$ exists.

Now, we consider the partition

$$P = \{0, 1/n, 2/n,\ldots, r/n,\ldots, n/n = 1\}$$

and the intermediate partition, $Q = \left\{\frac{1}{n}, \frac{2}{n}, \ldots \frac{n}{n} = 1\right\}$

Then *RS* (P, Q, f, g) = $\sum_{r=1}^{n} f\left(\frac{r}{n}\right)\left[g\left(\frac{r}{n}\right) - g\left(\frac{r-1}{n}\right)\right]$

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

$$= \frac{2}{n^3}\sum_{r=1}^{n} r^2 - \frac{1}{n^2}\sum_{r=1}^{n} r$$

$$= \frac{2}{n^3}.\frac{n(n+1)(2n+1)}{6} - \frac{1}{n^2}.\frac{n(n+1)}{2}$$

$$= \frac{1}{6n^2}[2(2n^2 + 3n+1) - 3n - 3] = \frac{4n^2 + 3n - 1}{6n^2}$$

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

Hence

$$\int_0^1 f\ dg = \lim_{||P|| \to 0} RS(P, Q, f, g)$$

$$= \lim_{n \to \infty} \frac{1}{6}\left[4 + \frac{1}{2n} - \frac{1}{6n^2}\right] = \frac{1}{6}[4 + 0 - 0] = \frac{2}{3}.$$

MAL-512: M. Sc. Mathematics (Real Analysis)

Lesson No. V **Written by Dr. Nawneet Hooda**

Lesson: Jacobians and extreme value problems **Vetted by Dr. Pankaj Kumar**

IMPLICIT FUNCTIONS

1.1 Existence theorem (Case of two variables)

Let f(x, y) be a function of two variables x and y and let (a,b) be a point in its domain of definition such that

(i) f(a, b) = 0 the partial derivatives f_x and f_y exist, and are continuous in a certain neighbourhood of (a, b) and

(ii) $f_y(a, b) \neq 0$,

then there exists a rectangle (a − h, a + h; b − k, b + k) about (a, b) such that for every value of x in the interval [a − h, a + h], the equation f(x, y) = 0 determines one and only one value y = $\phi(x)$, lying in the interval [b − k, b + k], with the following properties:

(1) b = $\phi(a)$,

(2) $f[x, \phi(x)] = 0$, for every x in [a − h, a + h], and

(3) $\phi(x)$ is derivable, and both $\phi(x)$ and $\phi'(x)$ are continuous in [a − h, a + h].

Proof. (Existence). Let f_x, f_y be continuous in a neighbourhood

$$R_1: (a - h_1, a + h_1; b - k_1, b + k_1) \text{ of } (a, b)$$

Since f_x, f_y exist and are continuous in R_1, therefore f is differentiable and hence continuous in R_1.

Again, since f_y is continuous, and $f_y(a, b) \neq 0$, there exists a rectangle

$$R_2:(a - h_2, a + h_2; b - k_2, b + k_2),\ h_2<h_1, k_2<k_1$$

(R_2 contained in R_1) such that for every point of this rectangle, $f_y \neq 0$

Since f = 0 and $f_y \neq 0$(it is therefore either positive or negative) at the point (a, b), a positive number $k<k_2$ can be found such that

f(a, b − k), (a, b + k)

are of opposite signs, for, f is either an increasing or a decreasing function of y, when y = b.

Again, since f is continuous, a positive number $h<h_2$ can be found such that for all x in [a – h, a + h],

$f(x, b - k), f(x, b + k),$

respectively, may be as near as we please to f(a, b – k), f(a, b + k) and therefore have opposite signs.

Thus, for all x in [a – h, a + h], f is a continuous function of y and changes sign as y changes from b – k to b + k. therefore it vanishes for some y in [b – k, b + k].

Thus, for each x in [a – h, a + h], there is a y in [b – k, b + k] for which f(x, y) = 0; this y is a function of x, say $\phi(x)$ such that properties (1) and (2) are true.

Uniqueness. We, now, show that $y = \phi(x)$ is a unique solution of f(x, y) = 0 in R_3: (a – h, a + h; b – k, b + k); that is f(x, y) cannot be zero for more than one value of y in [b – k, b + k].

Let, if possible, there be two such values y_1, y_2 in [b – k, b + k] so that $f(x, y_1) = 0$, $f(x, y_2) = 0$. Also f(x, y) considered as a function of a single variable y is derivable in [b – k, b + k], so that by Roll's theorem, $f_y = 0$ for a value of y between y_1 and y_2, which contradicts the fact that $f_y \neq 0$ in $R_2 \supset R_3$. hence our supposition is wrong and there cannot be more than one such y.

Let (x, y), $(x + \delta x, y + \delta y)$ be two points in R_3: (a – h, a + h; b – k, b + k) such that

$$y = \phi(x),\ y + \delta y = \phi(x + \delta x)$$

and

$$f(x, y) = 0,\ f(x + \delta x, y + \delta y) = 0$$

Since f is differentiable in R_1 and consequently in R_3 (contained in R_1),

$$\therefore \quad 0 = f(x + \delta x, y + \delta y) - f(x, y)$$

$$= \delta x\, f_x + \delta y\, f_y + \delta x\, \psi_1 + \delta y\, \psi_2$$

Where ψ_1, ψ_2 are functions of δx and δy, and tend to 0 as

$(\delta x, \delta y) \to (0,0)$

or

$$\frac{\delta y}{\delta x} = -\frac{f_x}{f_y} - \frac{\psi_1}{f_y} - \frac{\delta y}{\delta x}\frac{\psi_2}{f_y} \qquad (f_y \neq 0 \text{ in } R_3)$$

Proceeding to limits as $(\delta x, \delta y) \to (0, 0)$, we get

$$\phi'(x) = \frac{dy}{dx} = -\frac{f_x}{f_y}$$

Thus $\phi(x)$ is derivable and hence continuous in R_3. Also $\phi'(x)$, being a quotient of two continuous functions, is itself continuous in R_3.

JACOBIANS

If $u_1, u_2, \ldots, u_n$ be n differentiable of n variables $x_1, x_2, \ldots, x_n$, then the determinant

$$\begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots & \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \cdots & \frac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial u_n}{\partial x_1} & \frac{\partial u_n}{\partial x_2} & \cdots & \frac{\partial}{\partial} \end{vmatrix}$$

is called the Jacobian of the functions $u_1, u_2, \ldots, u_n$ with respect to $x_1, x_2, \ldots, x_n$ and denoted by

$$\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} \text{ or } J\left(\frac{u_1, u_2, \ldots, u_n}{x_1, x_2, \ldots, x_n}\right)$$

Example 3. If $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$, then show that

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} = r^2\sin\theta$$

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)} = \begin{vmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$

$$= r^2\sin\theta \begin{vmatrix} \sin\theta\cos\varphi & \cos\theta\cos\varphi & -\sin\theta \\ \sin\theta\sin\varphi & \cos\theta\sin\varphi & \cos\varphi \\ \cos\theta & -\sin\theta & 0 \end{vmatrix}$$

Adding $(\cos\phi)$ R_1 to $(\sin\phi)$ R_2,

$$= \frac{r^2\sin\theta}{\sin\varphi} \begin{vmatrix} \sin\theta\cos\varphi & \cos\theta\cos\varphi & -\sin\varphi \\ \sin\theta & \cos\theta & 0 \\ \cos\theta & -\sin\theta & 0 \end{vmatrix}$$

$$= r^2\sin\theta$$

Example 5. the roots of the equation in λ

$$(\lambda - x)^3 + (\lambda - y)^3 + (\lambda - z)^3 = 0$$

are u, v, w. Prove that

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = -2\frac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$$

Here u, v, w are roots of the equation

$$\lambda^3 - (x + y + z)\lambda^2 + (x^2 + y^2 + z^2)\lambda - \frac{1}{3}(x^3 + y^3 + z^3) = 0$$

$$\text{Let } x + y + z = \xi,\ x^2 + y^2 + z^2 = \eta,\ \frac{1}{2}(x^3 + y^3 + z^3) = \zeta \qquad (1)$$

and

$$u + v + w = \xi,\ vw + wu + uv = \zeta \qquad (2)$$

Hence from (1),

$$\frac{\partial(\xi, \eta, \zeta)}{\partial(x, y, z)} = \begin{vmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \\ x^2 & y^2 & z^2 \end{vmatrix}$$

$$= 2(y - z)(z - x)(x - y) \qquad (3)$$

and from (2),

$$\frac{\partial(\xi, \eta, \zeta)}{\partial(u, v, w)} = \begin{vmatrix} 1 & 1 & 1 \\ v+w & w+u & u+v \\ vw & wu & uv \end{vmatrix}$$

$$= -(v - w)(w - u)(u - v) \qquad (4)$$

Hence from (3) and (4) and using theorem 1, we get

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \frac{\partial(u,v,w)}{\partial(\xi,\eta,\zeta)}\cdot\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)} = -2\frac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$$

Example 7. If $u = \frac{x^2+y^2+z^2}{x}$, $v = \frac{x^2+y^2+z^2}{y}$, and $w = \frac{x^2+y^2+z^2}{z}$ find $\frac{\partial(x,y,z)}{\partial(u,v,w)}$.

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} 1-\frac{y^2+z^2}{x^2} & \frac{2y}{x} & \frac{2z}{x} \\ \frac{2x}{y} & 1-\frac{x^2+z^2}{y^2} & \frac{2z}{y} \\ \frac{2x}{z} & \frac{2y}{z} & 1-\frac{x^2+y^2}{z^2} \end{vmatrix}$$

Applying $C_1 \to C_1 + \frac{y}{x}C_2 + \frac{z}{x}C_3$

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{x^2+y^2+z^2}{x^2} & \frac{2y}{x} & \frac{2z}{x} \\ \frac{x^2+y^2+z^2}{xy} & 1-\frac{x^2+z^2}{y^2} & \frac{2z}{y} \\ \frac{x^2+y^2+z^2}{xz} & \frac{2z}{z} & 1-\frac{x^2+y^2}{z^2} \end{vmatrix}$$

$$= \frac{(x^2+y^2+z^2)}{x^2.xy.xz}\begin{vmatrix} 1 & 2xy & 2xz \\ 1 & xy-\frac{x}{y}(x^2+z^2) & 2xz \\ 1 & 2yz & xz-\frac{x}{z}(x^2+y^2) \end{vmatrix}$$

$$= \frac{(x^2+y^2+z^2)}{x^4yz}\begin{vmatrix} 1 & 2xy & 2xz \\ 0 & -\frac{x(x^2+y^2+z^2)}{y} & 0 \\ 0 & 0 & -\frac{x}{z}(x^2+y^2+z^2) \end{vmatrix}$$

$$= \frac{(x^2+y^2+z^2)^3}{x^2y^2z^2}$$

$$\therefore \qquad \frac{\partial(x,y,z)}{\partial(y,v,w)} = \frac{x^2y^2z^2}{(x^2+y^2+z^2)^3}$$

Example 15. If $u = x + 2y + z,\ v = x - 2y + 3z$

$w = 2xy - xz + 4yz - 2z^2$,

prove that $\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0$, and find a relation between u, v, w.

Solution. We have

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \\ 2y-z & 2x+4z & -x+4y-4z \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 1 & -4 & 2 \\ 2y-z & 2x+6z-4y & -x+2y-3z \end{vmatrix} \text{ Performing } c_2-2c_1 \text{ and } c_3-c_1$$

$$= \begin{vmatrix} -4 & 2 \\ 2x+6y-4y & -x+2y-3z \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ 0 & -x+2y-3z \end{vmatrix} \text{ Performing } c_1+2c_2$$

Hence a relation between u, v and w exists

Now,

$u = v = 2x + 4z$

$u - v = 4y - 2z$

$w = x(2y - z) + 2z(2y - 3)$

$= (x + 2z)(2y - z)$

$\Rightarrow \qquad 4w = (u + v)(u - v)$

$\Rightarrow \qquad 4w = u^2 - v^2$

Which is the required relation.

Example. 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

$$\frac{\partial u}{\partial x} = p, \frac{\partial u}{\partial y} = q, \frac{\partial u}{\partial z} = r$$

$$\frac{\partial v}{\partial x} = p' \frac{\partial y}{\partial y} = q', \frac{\partial v}{\partial z} = r'$$

$$\frac{\partial w}{\partial x} = 2ax + 2hy + 2gz$$

$$\frac{\partial w}{\partial y} = 2hx + 2by + 2fz$$

$$\frac{\partial w}{\partial z} = 2gx + 2fy + 2cz$$

Therefore

$$\begin{vmatrix} p & q & r \\ p' & q' & r' \\ 2ax+2hy+2gz & 1hx+2hy+2fz & 2gx+2fy+2cz \end{vmatrix}$$

$$\Rightarrow \quad \begin{vmatrix} p & q & r \\ p' & q' & r' \\ a & h & g \end{vmatrix} = 0, \begin{vmatrix} p & q & r \\ p' & q' & r' \\ h & b & f \end{vmatrix} = 0, \begin{vmatrix} p & q & r \\ p' & q' & r' \\ g & f & c \end{vmatrix} = 0$$

which is the required condition.

Example. 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 $\quad J(u, v) = \begin{vmatrix} \frac{\partial u}{\partial x}, & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$

$$= \begin{vmatrix} \frac{1}{1+x^2} & \frac{1}{1+y^2} \\ \frac{1+y^2}{(1-xy)^2} & \frac{1+x^2}{(1+xy)^2} \end{vmatrix} = 0$$

Therefore u and v are connected by a functional relation

Let $\quad u = \phi(v)$, that is,

$$f(x) + f(y) = \phi\left(\frac{x+y}{1-xy}\right)$$

Putting y = 0, we get

$$f(x) + f(0) = \phi(x)$$

$$\Rightarrow \qquad f(x) + 0 = \phi(x) \qquad \because f(0) = 0$$

Hence $$f(x) + f(y) = f\left(\frac{x+y}{1-xy}\right)$$

Example. 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. If $u = x^2 + y^2 + z^2$, $v = x + y + z$, $w = xy + yz + zx$, show that the Jacobian $\frac{\partial(u,v,w)}{\partial(x,y,z)}$ vanishes identically.

Solution. $$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \\ y+z & z+x & x+y \end{vmatrix}$$

$$= 2\begin{vmatrix} x & y & z \\ 1 & 1 & 1 \\ x+z & z+x & x+y \end{vmatrix}$$

$$= 2\begin{vmatrix} x+y+z & x+y+z & x+y+z \\ 1 & 1 & 1 \\ y+z & z+x & x+y \end{vmatrix} \quad \text{[Adding } R_3 \text{ to } R_1\text{]}$$

$$= 2(x + y + z)\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ y+z & z+x & x+y \end{vmatrix}$$

Now we find relation between u, v, w. We have

$$v^2 = (x + y + z)^2 = x^2 + y^2 + z^2 + 2(xy + yz + zx) = uu + 2w$$

or $$v^2 = u + 2w.$$

Example. If $u = x + y + z + t$, $v = x + y - z - t$, $w = xy - zt$,

$$r = x^2 + y^2 - z^2 - t^2,$$

show that $\frac{\partial(u,v,w,r)}{\partial(x,y,z,t)} = 0$

and hence find a relation between x, y, z and t.

Solution. $$\frac{\partial(u,v,w,r)}{\partial(x,y,z,t)} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ y & x & -t & -z \\ 2x & 2y & -2z & -2t \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 2 & 0 & 0 \\ 1 & 1 & -1 & -1 \\ y & x & -t & -z \\ 2x & 2y & -2z & -2t \end{vmatrix} \quad \text{[Adding } R_2 \text{ to } R_1]$$

$$\begin{vmatrix} 2 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 \\ y & x-y & -t & -z \\ 2x & 2y-2x & -2z & -2t \end{vmatrix} \quad \text{[Subtracting } C_1 \text{ from } C_2]$$

$$= 2\begin{vmatrix} 0 & -1 & -1 \\ x-y & -t & -z \\ -2(x-y) & -2z & -2t \end{vmatrix}$$

$$= 2\begin{vmatrix} 0 & 0 & -1 \\ x-y & z-t & -z \\ -2(x-y) & 2(t-z) & -2t \end{vmatrix} \quad \text{[Operating } C_2 - C_3]$$

$$= 2(x-y)(z-t)\begin{vmatrix} 0 & 0 & -1 \\ 1 & 1 & -z \\ -2 & -2 & -2t \end{vmatrix} \quad [\because C_1 \text{ and } C_2 \text{ are identical}]$$

Hence the functions u, v, w, r are not independent.

Now we find a relation between u, v, w, r. We have

$$uv = (x + y + z + t)(x + y - z - t) = (x + y)^2 - (z + t)^2$$

$$= (x^2 + y^2 - z^2 - t^2) + 2(xy - zt) = r + 2w$$

Thus uv = r + 2w,

which is the required relation.

EXTREME VALUES: MAXIMA MINIMA

Let (a, b) be a point of the domain of definition of a function f. The f (a, b) is an extreme value of f, if for every point (x, y) of some neighbourhood of (a, b) the difference f(x,y)- f(a,b)keeps the same sign. (1)

The extreme value f (a, b) is called a maximum or minimum value according as the sign of (1) is negative or positive.

10.1 A Necessary Condition

A necessary condition for f(x, y) to have an extreme value at (a, b) is that f_x (a, b) = 0, f_y(a, b) = 0; provided these partial derivatives exist.

Points at which $f_x = 0$, $f_y = 0$ are called Stationary points.

10.2 Sufficient Conditions for f(x, y) **to have extreme value at** (a, b)

Let $f_x(a, b) = 0 = f_y(a, b)$. Further, let us suppose that f(x, y) possesses continuous second order partial derivatives in a neighbourhood of (a, b) and that these derivatives at (a, b) viz. $f_{xx}(a, b)$, $f_{xy}(a, b)$, $f_{yy}(a, b)$ are not all zero.

Let (a + h, b+ k) be a point of this neighbourhood.

Let us write

$$r = f_{xx}(a, b),\ s = f_{xy}(a, b), tC = f_{yy}(a, b)$$

(1) If $rs - t^2 > 0$, then f(a, b) is a maximum value if $r < 0$, and a minimum value if $r > 0$.

(2) If $rt - s^2 < 0$, f(a, b) is not an extreme value.

(3) If $rt - s^2 = 0$,

Thus is the doubtful case and requires further investigation.

Example 32. Find the maxima and minima of the function

$$f(x, y) = x^3 + y^3 - 3x - 12y + 20$$

Solution. We have

$$f_x(x, y) = 3y^2 - 3 = 0, \text{ when } x = \pm 1$$

$$f_y(x, y) = 3y^2 - 12 = 0, \text{ when } y = \pm 2$$

Thus the function has four stationary points:

(1, 2), (–1, 2), (1, –2), (–1, –2),

Now

$$f_{xx}(x, y) = 6x f_{xy}(x, y) = 0,\ f_{yy}(x, y) = 6y$$

At (1, 2),

$$f_{xx} = 6 > 0, \text{ and } f_{xx}f_{yy} - (f_{xy})^2 = 72 > 0$$

Hence (1, 2) is a point of minima of the function.

At (1, –2),

$$f_{xx} = -6, \text{ and } f_{xx}f_{yy} - (f_{xy})^2 = -72 < 0$$

Hence the function has neither maximum nor minimum at (1, −2).

At (−1, −2),

$$f_{xx} = -6, \text{ and } f_{xx}f_{yy} - (f_{xy})^2 = 72 > 0$$

Hence the function has a maximum value at (−1, −2).

Note. Stationary points like (−1, 2), (1, −2) which are not extreme points are called the saddle points.

Example 33. Show that the function

$$f(x, y) = 2x^4 - 3x^2y + y^2$$

has neither a maximum nor minimum at (0, 0)where

$$f_{xx}f_{yy} - (f_{xy})^2 = 0$$

Now

$$f_x(x, y) = 8x^3 - 6xy,\ f_y(x, y) = -3x^2 + 2y$$

$$\therefore\ f_x(0, 0) = 0 = f_y(0, 0)$$

Also

$$f_{xx}(x, y) = 24x^2 - 6y = 0, \text{ at } (0, 0)$$

$$f_{xy}(x, y) = -6x = 0 \text{ at } (0, 0)$$

$$f_{yy}(x, y) = 2, \text{ at } (0, 0)$$

Thus at (0, 0), $f_{xx}(0, 0).f_{yy}(0, 0) - [f_{xy}(0, 0)]^2 = 0$.

So that it is a doubtful case, and so requires further examination.

Again

$$f(x, y) = (x^2 - y)(2x^2 - y),\ f(0, 0) = 0$$

or

$$f(x, y) - f(0, 0) = (x^2 - y)\ (2x^2 - y)$$

$$>0, \text{ for } y< 0 \text{ or } x^2>y>0$$

$$<0, \text{ for } y> x^2>\frac{y}{2}>0$$

Thus f(x, y) − f(0, 0) does not keep the same sign near the origin. Hence f has neither a maximum nor minimum value at the origin.

Example 34. Show that

$$f(x, y) = y^2 + x^2y + x^4, \text{ has a minimum at } (0, 0).$$

It can be easily verified that the origin.

$f_x = 0$, $f_y = 0$, $f_{xx} = 0$, $f_{xy} = 0$, $f_{yy} = 2$.

Thus at the origin $f_{xx}f_{yy} - (f_{xy})^2 = 0$, so that it is a doubtful case and requires further investigation.

But we can write

$$f(x, y) = (y + \frac{1}{2}x^2)^2 + \frac{3}{4}x^4$$

and

$$f(x, y) - f(0, 0) = (y + \frac{1}{2}x^2)^2 + \frac{3}{4}x^4$$

which is greater than zero for all values of (x, y). hence f has a minimum value at the origin.

Example. 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

$$y^2 + x^2y + x^4 = \left(y + \frac{1}{2}x^2\right)^2 + \frac{3x^4}{4}$$

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. Let

$$f(x, y) = 2x^4 - 3x^2 y + y^2$$

Then $\frac{\partial f}{\partial x} = 8x^3 - 6xy \qquad \Rightarrow \frac{\partial f}{\partial x}(0,0) = 0; \frac{\partial f}{\partial y} = -3x^2 + 2y \quad \Rightarrow \frac{\partial f}{\partial y}(0,0) = 0$

$$r = \frac{\partial^2 f}{\partial x^2} 24x^2 - 6y = 0 \text{ at } (0, 0), S = \frac{\partial^2 f}{\partial x \partial y} = -6x = 0 \text{ at } (0, 0)$$

$$t = \frac{\partial^2 f}{\partial y^2} = 2. \text{ Thus } rt - S^2 = 0. \text{ Thus it is a doubtful case}$$

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).

Lagrange's Undetermined Multipliers

3.2 Lagrange's Method of Multipliers

To find the stationary points of the function

$$f(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m) \qquad \ldots(1)$$

of n + m variables which are connected by the equations

$$\phi_r(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m) = 0, \; r = 1, 2, \ldots, m \qquad \ldots(2)$$

For stationary values, df = 0

$$\therefore \qquad 0 = df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \ldots + \frac{\partial f}{\partial x_n} dx_n$$

$$+ \frac{\partial f}{\partial u_1} du_1 + \ldots + \frac{df}{du_m} du_m \qquad \ldots(3)$$

Differentiating equations (2) we get

$$\left.\begin{aligned}
&\frac{\partial \varphi_1}{\partial x_1}dx_1+\ldots+\frac{\partial \varphi_1}{\partial x_n}dx_n+\frac{\partial \varphi_1}{\partial u_1}du_1+\ldots+\frac{\partial \varphi_1}{\partial u_m}du_m=0\\
&\frac{\partial \varphi_2}{\partial x_1}dx_1+\ldots+\frac{\partial \varphi_2}{\partial x_n}dx_n+\frac{\partial \varphi_2}{\partial u_1}du_1+\ldots+\frac{\partial \varphi_2}{\partial u_m}du_m=0\\
&\vdots\\
&\frac{\partial \varphi_m}{\partial x_1}dx_1+\ldots+\frac{\partial \varphi_m}{\partial x_n}dx_n+\frac{\partial \varphi_m}{\partial u_1}du_1+\ldots+\frac{\partial \varphi_m}{\partial u_m}du_m=0
\end{aligned}\right\} \quad \ldots(4)$$

Multiplying the equations (4) by $\lambda_1, \lambda_2, \ldots, \lambda_m$ respectively and adding to the equation (3) we get

$$0 = df = \left(\frac{\partial f}{\partial x_1}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial x_1}\right)dx_1+\ldots+\left(\frac{\partial f}{\partial x_n}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial x_n}\right)dx_n$$

$$+\left(\frac{\partial f}{\partial u_1}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial u_1}\right)du_1+\ldots+\left(\frac{\partial f}{\partial u_m}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial u_m}\right)du_m \quad \ldots(5)$$

Let the m multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ be so chosen that the coefficients of the m differentials $du_1, du_2, \ldots, du_m$ all vanish, i.e.,

$$\frac{\partial f}{\partial u_1}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial u_1}=0,\ldots,\frac{\partial f}{\partial u_m}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial u_m}=0 \quad \ldots(6)$$

Then (5) becomes

$$0 = df = \left(\frac{\partial f}{\partial x_1}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial x_1}\right)dx_1+\ldots+\left(\frac{\partial f}{\partial x_n}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial x_n}\right)dx_n$$

so that the differential df is expressed in terms of the differentials of independent variables only. Hence

$$\frac{\partial f}{\partial x_1}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial x_1}=0,\ldots.,\frac{\partial f}{\partial x_n}+\Sigma\lambda_r\frac{\partial \varphi_r}{\partial x_n}=0 \quad \ldots(7)$$

Equations (2), (6), (7) form a system of n + 2m equations which may be simultaneously solved to determine the m multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the n + m coordinates $x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m$ of the stationary points of f.

An Important Rule. For practical purposes,

Define a function

$$F = f + \lambda_1\phi_1 + \lambda_2\phi_2 + \ldots + \lambda_m\phi_m$$

At a stationary point of F, dF = 0. Therefore

$$0 = dF = \frac{\partial F}{\partial x_1}dx_1 + \frac{\partial F}{\partial x_2}dx_2 + \ldots + \frac{\partial F}{\partial x_n}dx_n + \frac{\partial F}{\partial u_1}du_1 + \ldots + \frac{\partial F}{\partial u_m}du_m$$

$$\therefore \quad \frac{\partial F}{\partial x_1} = 0, \ldots, \frac{\partial F}{\partial x_n} + 0, \frac{\partial F}{\partial u_1} = 0, \ldots, \frac{\partial F}{\partial u_m} = 0$$

which are same as equations (7) and (6).

Thus the stationary points of f may be found by determining the stationary points of the function F, where

$$F = f + \lambda_1\phi_1 + \lambda_2\phi_2 + \ldots + \lambda_m\phi_m$$

and considering all the variables as independent variables.

A stationary point will be an extreme point of f if d^2F keeps the same sign, and will be a maxima or minima according as d^2F is negative or positive.

Example 9. Find the shortest distance from the origin to the hyperbola

$$x^2 + 8xy + 7y^2 = 225, = 0$$

Solution. We have to find the minimum value of $x^2 + y^2$ subject to the constraint

$$x^2 + 8xy + 7y^2 = 225$$

Consider the function

$$F = x^2 + y^2 + \lambda(x^2 + 8xy + 7y^2 - 225)$$

where x, y are independent variables and λ a constant.

$$dF = (2x + 2x\lambda + 8y\lambda)\, dx + (2y + 8x\lambda + 14y\lambda)dy$$

$$\therefore \left.\begin{aligned}(1+\lambda)x + 4\lambda y = 0\\ 4\lambda x + (1+7\lambda)y = 0\end{aligned}\right\} \qquad \therefore \quad \lambda = 1, -\frac{1}{9}$$

For λ = 1, x = –2y, and substitution in $x^2 + 8xy + 7y^2 = 225$, gives $y^2 = -45$, for which no real solution exists.

For $\lambda = -\frac{1}{9}$, y = 2x and substitution in $x^2 + 8xy + 7y^2 = 225$, gives $x^2 = 5$, $y^2 = 20$, and so $x^2 + y^2 = 25$.

$$d^2F = 2(1 + \lambda)\, dx^2 + 16\lambda\, dx\, dy + 2(1 + 7\lambda)\, dy^2$$

$$= \frac{16}{9}dx^2 - \frac{16}{9}dx\, dy + \frac{4}{9}dy^2, \text{ at } \lambda = -\frac{1}{9}$$

$$= \tfrac{4}{9}(2dx - dy)^2$$

> 0, and cannot vanish because $(dx, dy) \neq (0, 0)$.

Hence the function $x^2 + y^2$ has a minimum value 25.

Example 10. Find the maximum and minimum values of $x^2 + y^2 + z^2$ subject to the conditions $\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} = 1$, and $z = x + y$.

Solution. Let us consider a function F of independent variables x, y, z where

$$F = x^2 + y^2 + z^2 + \lambda_1 \left(\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} - 1\right) + \lambda_2 (x + y - z)$$

$$\therefore \qquad dF = \left(2x + \frac{x}{2}\lambda_1 + \lambda_2\right)dx + \left(2y + \frac{2y}{5}\lambda_1 + \lambda_2\right)dy + \left(2z + \frac{2z}{25}\lambda_1 - \lambda_2\right)dz$$

As x, y, z are independent variables, we get

$$2x + \frac{x}{2}\lambda_1 + \lambda_2 = 0$$

$$2y + \frac{2y}{5}\lambda_1 + \lambda_2 = 0$$

$$2z + \frac{2z}{25}\lambda_1 - \lambda_2 = 0$$

$$\therefore \qquad x = \frac{-2\lambda_2}{\lambda_1 + 4}, \qquad y = \frac{-5\lambda_2}{2\lambda_1 + 10}, \qquad z = \frac{25\lambda_2}{2\lambda_2 + 50}$$

Substituting in $x + y = z$, we get

$$\frac{2}{\lambda_1 + 4} + \frac{5}{2\lambda_1 + 10} + \frac{25}{2\lambda_1 + 50} = 0, \quad \lambda_2 \neq 0 \qquad \ldots(1)$$

for if, $\lambda_2 = 0$, $x = y = z = 0$, but $(0, 0, 0)$ does not satisfy the other condition of constraint.

Hence from (1), $17\lambda_1^2 + 245\lambda_1 + 750 = 0$, so that $\lambda_1 = -10, -75/17$.

For $\lambda_1 = -10$,

$$x = \tfrac{1}{3}\lambda_2, \; y = \tfrac{1}{2}\lambda_2, \; z = \tfrac{5}{6}\lambda_2$$

Substituting in $\frac{x^2}{4}+\frac{y^2}{5}+\frac{z^2}{25}=1,$ we get

$$\lambda_2^2 = 180/19 \quad \text{or} \quad \lambda_2 = \pm 6\sqrt{5/19}$$

The corresponding stationary points are

$$(2\sqrt{5/19}, 3\sqrt{5/19}, 5\sqrt{5/19}), (-2\sqrt{5/19}, -3\sqrt{5/19}, -5\sqrt{5/19})$$

The value of $x^2 + y^2 + z^2$ corresponding to these points is 10.

For $\lambda_1 = -75/17$,

$$x = \frac{34}{7}\lambda_2, \ y = -\frac{17}{4}\lambda_2, \ z = \frac{17}{28}\lambda_2,$$

which on substitution in $\frac{x^2}{4}+\frac{y^2}{5}+\frac{z^2}{25} = 1$ give

$$\lambda_2 = \pm 140/(17\sqrt{646})$$

The corresponding stationary points are

$$(40/\sqrt{646}, -35/\sqrt{646}, 5/\sqrt{646}), (-40/\sqrt{646}, 35/\sqrt{646}, -5/\sqrt{646})$$

The value of $x^2 + y^2 + z^2$ corresponding to these points is 75/17.

Thus the maximum value is 10 and the minimum 75/17.

Example 11. Prove that the volume of the greatest rectangular parallelepiped that can be inscribed in the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = 1$, is $\frac{8abc}{3\sqrt{3}}$

Solution. We have to find the greatest value of 8xyz subject to the conditions

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \quad x > 0, y > 0, z > 0 \qquad \ldots(1)$$

Let us consider a function F of three independent variables x, y, z, where

$$F = 8xyz + \lambda\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1\right)$$

$$\therefore \quad dF = \left(8yz+\frac{2x\lambda}{a^2}\right)dx + \left(8zx+\frac{2y\lambda}{b^2}\right)dy + \left(dxy+\frac{2z\lambda}{c^2}\right)dz$$

At stationary points,

$$8yz + \frac{2x\lambda}{a^2} = 0,\ 8zx + \frac{2y\lambda}{b^2} = 0,\ 8xy + \frac{2z\lambda}{c^2} = 0 \qquad \ldots(2)$$

Multiplying by x, y, z respectively and adding,

$$24xyz + 2\lambda = 0 \quad \text{or } \lambda = -12xyz. \quad \text{[using (1)]}$$

Hence, from (2), $x = a/\sqrt{3}$, $y = b/\sqrt{3}$, $z = c/\sqrt{3}$, and so $\lambda = -4abc/\sqrt{3}$

Again

$$d^2F = 2\lambda\left(\frac{dx^2}{a^2} + \frac{dy^2}{b^2} + \frac{dz^2}{c^2}\right) + 16z\, dx\, dy + 16x\, dy\, dz + 16\, dz\, dx$$

$$= -\frac{8abc}{\sqrt{3}}\Sigma\frac{1}{a^2}dx^2 + \frac{16}{\sqrt{3}}\Sigma c\, dx\, dy \qquad \ldots(3)$$

Now from (1) we have

$$x\frac{dx}{a^2} + y\frac{dy}{b^2} + z\frac{dz}{c^2} = 0 \quad \text{or} \quad \frac{dx}{a} + \frac{dy}{b} + \frac{dz}{c} = 0. \qquad \ldots(4)$$

Hence squaring,

$$\Sigma\frac{dx^2}{a^2} + 2\Sigma\frac{dx\, dy}{ab} = 0$$

or

$$abc\,\Sigma\frac{dx^2}{a^2} = -2\Sigma c\, dx\, dy$$

$$\therefore\ d^2F = -\frac{16}{\sqrt{3}}abc\,\Sigma\frac{dx^2}{a^2}$$

which is always negative.

Hence $\left(\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}, \frac{c}{\sqrt{3}}\right)$ is a point of maxima and the maximum value of 8 xyz is $\frac{8abc}{3\sqrt{3}}$.

MAL-512: M. Sc. Mathematics (Real Analysis)

Lesson No. VI **Written by Dr. Nawneet Hooda**

Lesson: The Riemann-Stieltjes integrals **Vetted by Dr. Pankaj Kumar**

Definition . Let f and α be bounded functions on [a, b] and α be monotonic increasing on [a, b], $b \geq a$.

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be any partition of [a, b] and let

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

Note that $\Delta\alpha_i \geq 0$. Let us define two sums,

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

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

where m_i, M_i, are infimum and supremum respectively of f in Δx_i. Here $U(P, f, \alpha)$ is called the Upper sum and $L(P, f, \alpha)$ is called the Lower sum of f corresponding to the partition P.

Let m, M be the lower and the upper bounds of f on [a, b], then we have

$$m \leq m_i \leq M_i \leq M$$

$$\Rightarrow \quad m\,\Delta\,\alpha_i \leq m_i\,\Delta\,\alpha_i \leq M_i\Delta\alpha_i, \leq M\Delta\alpha_i,\ \Delta\alpha_i \geq 0$$

Putting i = 1, 2,…, n and adding all inequalities, we get

$$m\{\alpha(b) - \alpha(a)\} \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq M\{\alpha(b) - \alpha(a)\} \qquad \ldots(1)$$

As in case of Riemann integration, we define two integrals,

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

$$\underline{\int_a^b} f\ d\alpha = \text{sup. } L(P, f, \alpha) \qquad \ldots(2)$$

where the infimum and supremum is taken over all partitions of [a, b]. These are respectively called the upper and the lower integrals of f with respect to α.

These two integrals may or may not be equal. In case these two integrals are equal, i.e.,

$$\overline{\int_a^b} f \quad d\alpha = \underline{\int_a^b} f\, d\alpha,$$

we say that f is integrable with respect to α in the Riemann sense and write f∈ R(α). Their common value is called the Riemann-Stieltjes integral of f with respect to α, over [a, b] and isdenoted by

$$\int_a^b f\ d\alpha \text{ or } \int_a^b f(x)\, d\alpha(x).$$

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

$$m\{\alpha(b) - \alpha(a)\} \le L(P, f, \alpha) \le \overline{\int_a^b} f \quad d\alpha \le \underline{\int_a^b} f \ f\, d\alpha$$

$$\le U(P, f, \alpha) \le M\{\alpha(b) - \alpha(a)\} \qquad \ldots(3)$$

Note 1. The upper and the lower integrals always exist for bounded functions but these may not be equal for all bounded functions.The Riemann-Stieltjes integral reduces to Riemann integral when α(x) = x.

Note 2. As in case of Riemann integration, we have

(1) If f ∈ R(α), then there exits a number λ between the bounds of f such that

$$\int_a^b f\ d\alpha = \lambda\{\alpha(b) - \alpha(a)\} \qquad \text{(by 3)}$$

(2) If f is continuous on [a, b], then there exits a number ξ ∈[a, b] such that

$$\int_a^b f\ d\alpha = f(\xi)\ \{\alpha(b) - \alpha(a)\}$$

(3) If f ∈ R(α), and k is a number such that

$|f(x)| \le k$, for all x ∈ [a, b]

then

$$\left|\int_a^b f\ d\alpha\right| \le k\{\alpha(b) - \alpha(a)\}$$

(4) If $f \in R(\alpha)$ over $[a, b]$ and $f(x) \geq 0$, for all $x \in [a, b]$, then

$$\int_a^b f \; d\alpha \begin{cases} \geq 0, & b \geq \\ \leq 0, & b \leq a \end{cases}$$

Since $f(x) \geq 0$, the lower bound $m \geq 0$ and therefore the result follows from (3).

(5) If $f \in R(\alpha)$, $g \in R(\alpha)$ over $[a, b]$ with $f(x) \geq g(x)$, then

$$\int_a^b f \; d\alpha \geq \int_a^b g \, d\alpha \, , \; b \geq a,$$

and

$$\int_a^b f \; d\alpha \leq \int_a^b g \; d\alpha, \; b \leq a.$$

Theorem . If P^* is a refinement of P, then

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

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

Proof. (a) Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be a partition of the given interval. Let P^* contain just one point more than P. Let this extra point ξ belongs to Δx_i, i.e., $x_{i-1} < \xi \; x_i$.

As f is bounded over the interval $[a, b]$, it is bounded on every sub-interval $\Delta x_i (i = 1, 2, \ldots, n)$. Let V_1, V_2, M_i be the upper bounds (supremum) of f in the intervals $[x_{i-1}, \xi]$, $[\xi, x_i]$, $[x_{i-1}, x_i]$, respectively.

Clearly

$$V_1 \leq M_i, \; V_2 \leq M_i.$$

$\therefore \; U(P^*, f, \alpha) - U(P, f, \alpha) = V_1\{\alpha(\xi) - \alpha(x_{i-1})\} + V_2\{\alpha(x_i) - \alpha(\xi)\} - M_i\{\alpha(x_i) - \alpha(x_{i-1}\}$

$$= (V_1 - M_1) \; \{\alpha(\xi) - \alpha(x_{i-1})\}$$
$$+ (V_2 - M_i) \; \{\alpha(x_i) - \alpha(\xi)\} \leq 0$$

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

If P^* contains m points more than P, we repeat the above arguments m times and get the result .

The proof of (b) runs on the same arguments.

Theorem . A function f is integrable with respect to α on [a, b] if and only if for every $\varepsilon > 0$ there exists a partition P of [a, b] such that

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

Proof. Let $f \in F(\alpha)$ over [a, b]

$$\therefore \quad \underline{\int_a^b} f\, d\alpha = \overline{\int_a^b} f \quad d\alpha = \int_a^b \quad f\, d\alpha$$

Let $\varepsilon > 0$ be any number.

Since the upper and the lower integrals are the infimum and the supremum of the upper and the lower sums, therefore ∃ partitions P_1 and P_2 such that

$$U(P_1, f, \alpha) < \overline{\int_a^b} f \quad f\, d\alpha + \frac{1}{2}\varepsilon = \int_a^b f\, d\alpha + \frac{1}{2}\varepsilon$$

$$L(P_2, f, \alpha) > \underline{\int_a^b} f \quad d\alpha - \frac{1}{2}\ \varepsilon = \int_a^b f\, d\alpha - \frac{1}{2}\varepsilon$$

Let $P = P_1 \cup P_2$ be the common refinement of P_1 and P_2.

$$\therefore \quad U(P, f, \alpha) \le U(P_1, f, \alpha)$$

$$< \int_a^b f\, d\alpha + \frac{1}{2}\varepsilon < L\,(P_2, f, \alpha) + \varepsilon$$

$$\le L(P, f, \alpha) + \in$$

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

Conversely . For $\in > 0$, let P be a partition for which

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

For any partition P, we have

$$L(P, f, \alpha) \le \underline{\int_a^b} f\, d\alpha \le \overline{\int_a^b} f \quad d\alpha \le U(P, f, \alpha)$$

$$\therefore \quad \overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha \le U(P, f, \alpha) - L(P, f, \alpha) < \in$$

$$\therefore \quad \overline{\int_a^b} f \quad = \underline{\int_a^b} f\, d\alpha$$

so that $f \in R(\alpha)$ over [a, b].

Theorem. If $f_1 \in R(\alpha)$ and $f_2 \in R(\alpha)$ over [a, b], then

$$f_1 + f_2 \in R(\alpha) \text{ and } \int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$$

Proof. Let $f = f_1 + f_2$.

Then f is bounded on [a, b].

If $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be any partition of [a, b] and m_i' M_1'; m_i'', M_i''; m_i, M_i the bounds of f_1, f_2 and f, respectively on Δx_i, then

$$m_i' + m_i'' \le m_i \le M_i \le M_i' + M_i''$$

Multiplying by $\Delta\alpha_i$ and adding all these inequalities for i = 1, 2, 3,…, n, we get

$$L(P, f_1, \alpha) + L(P, f_2, \alpha) \le L(P, f, \alpha) \le U(P, f, \alpha)$$

$$\le U(P, f_1, \alpha) + U(P, f_2, \alpha) \qquad \ldots(1)$$

Let $\in > 0$ by any number.

Since $f_1 \in R(\alpha)$, $f_2 \in R(\alpha)$, therefore $\exists$ partitions P_1, P_2 such that

$$U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha) < \tfrac{1}{2} \in$$

$$U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha) < \tfrac{1}{2} \in$$

Let $P = P_1 \cup P_2$, a refinement of P_1 and P_2.

$$\therefore \quad U(P, f_1, \alpha) - L(P, f_1, \alpha) < \tfrac{1}{2} \in$$

$$U(P, f_2, \alpha) - L(P, f_2, \alpha) < \tfrac{1}{2} \in$$

Thus for partition P, we get from (1) and (2).

$$U(P, f, \alpha) - L(P, f, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha) - L(P, f_1, \alpha) - L(P, f_2, \alpha)$$

$$< \tfrac{1}{2} \in + \tfrac{1}{2} \in = \in$$

$$\Rightarrow \quad f \in R(\alpha) \text{ over } [a, b]$$

Since the upper integral is the infimum of the upper sums, therefore ∃ partitions P_1, P_2 such that

$$U(P_1, f_1, \alpha) < \int_a^b f_1 \, d\alpha + \tfrac{1}{2} \in$$

$$U(P_2, f_2, \alpha) < \int_a^b f_2 \, d\alpha + \tfrac{1}{2} \in$$

If $P = P_1 \cup P_2$, we have

$$\left.\begin{aligned} U(P, f_1, \alpha) &< \int_a^b f_1 \, d\alpha + \tfrac{1}{2}\varepsilon \\ U(P, f_2, \alpha) &< \int_a^b f_2 \, d\alpha + \tfrac{1}{2}\varepsilon \end{aligned}\right\} \qquad \ldots(3)$$

For such a partition P,

$$\int_a^b f \, d\alpha \le U(P, f, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha) \qquad \text{[from (1)]}$$

$$\le \int_a^b f_1 \, d\alpha + \int_a f_2 + \varepsilon \qquad \text{[by (3)]}$$

Since ε is arbitrary, we get

$$\int^b f \, d\alpha \le \int^b f_1 \, d\alpha + \int^b f_2 \, d\alpha \qquad \ldots(4)$$

Taking $(-f_1)$ and $(-f_2)$ in place of f_1 and f_2, we get

$$\int_a^b f \, d\alpha \ge \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha \qquad \ldots(5)$$

(4) and (5) give

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

Theorem. If $f \in R(\alpha)$, and c is a constant, then

$$cf \in R(\alpha) \text{ and } \int_a^b cf \; d\alpha = c\int_a^b f \; d\alpha$$

Proof. 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

$$\int_a^b g\ d\alpha \ge \int_a^b f\ d\alpha$$

Hence $\int_a^b (cf)d\alpha = c\int_a^b f\, d\alpha$

Theorem. If f ∈ R(α) on [a, b] , then f ∈R(α) on [a, c] and f∈R(α) 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 ∈ R(α), there exists a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \in, \quad \in > 0.$$

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

$$U(P, f, \alpha) - L(P^*, f, \alpha) \le U(P, f, \alpha) - L(P, f, \alpha) \qquad (1)$$

Let P_1 and P_2 denote the sets of point of P* between [a, b], [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

$$U(P^*, f, \alpha) = U(P_1, f, \alpha) + U(P_2, f, \alpha) \qquad (2)$$

and

$$L(P^*, f, \alpha) = L(P_1, f, \alpha) + L(P_2, f, \alpha) \qquad (3)$$

Then (1), (2) and (3) 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 (2) yields

$$\overline{\int} f\ d\alpha \geq \overline{\int_{a}^{c}} f\ d\alpha + \overline{\int_{c}^{b}} f\ d\alpha \qquad (4)$$

But since f integrable on [a, c] and [c, b], we have

$$\int_a^b f(x) d\alpha \geq \int_a^c f\ d\alpha + \int_c^b f\ d\alpha \qquad (5)$$

The relation (3) similarly yields

$$\int_a^b f\ d\alpha \leq \int_a^c f\ d\alpha + \int_c^b f\, d\alpha \qquad (6)$$

Hence (5) and (6) imply that

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

Theorem. If $f \in \Re(\alpha)$ and if $|f(x)| \leq K$ on [a, b], then

$$\left| \int_a^b f\ d\alpha \right| \leq K[\alpha(b) - \alpha(a)].$$

Proof. If M and m are bounds of $f \in \Re(\alpha)$ on [a, b], then it follows that

$$m[\alpha(b) - \alpha(a)] \leq \int_a^b f\, d\alpha \leq M[\alpha(b) - \alpha(a)] \text{ for } b \geq a. \qquad (1)$$

In fact, if a = b, then (1) is trivial. If b > a, then for any partition P, we have

$$m[\alpha(b) - \alpha(a)] \leq \sum_{i=1}^{n} m_i\ \Delta\alpha_i = L(P, f, \alpha)$$

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

$$\leq U(P, f, \alpha) = \sum M_i \ \Delta \alpha_i$$

$$\leq M(b - a)$$

which yields

$$m[\alpha(b) - \alpha(a)] \leq \int_a^b f\, d\alpha \leq M(b - a) \qquad (2)$$

Since $|f(x)| \leq k$ for all $x \in (a, b)$, we have

$$-k \leq f(x) \leq k$$

so if m and M are the bounds of f in (a, b),

$$-k \leq m \leq f(x) \leq M \leq k \text{ for all } x \in (a, b).$$

If $b \geq a$, then $\alpha(b) - \alpha(a) \geq 0$ and we have by (2)

$$-k[\alpha(b) - \alpha(a)] \leq m[\alpha(b) - \alpha(a)] \leq \int_a^b f\ d\alpha$$

$$\leq M[\alpha(b) - \alpha(a)] \leq k[\alpha(b) - \alpha(a)]$$

Hence

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

Theorem. 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 $h(x) = \phi[(x)] = f^2 \in \Re(\alpha)$.

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 $\phi(f) = |t|$, then $|f| \in \Re(\alpha)$. We choose c = ± 1 so that

$$C \int f \, d\alpha \geq 0$$

Then

$$\left|\int f \, d\alpha\right| = \int f \, d\alpha = \int c \, f \, d\alpha \leq \int |f| \, d\alpha$$

Because cf ≤ |f|.

(3) if $f_1 \in R(\alpha)$, $f_2 \in R(\alpha)$ and $f_1(x) \leq f_2(x)$ on [a, b] then

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

Theorem. If $f \in R(\alpha_1)$ and $f \in R(\alpha_2)$, then

$$f \in R(\alpha_1 + \alpha_2) \text{ and } \int_a^b f \, d(\alpha_1 + \alpha_2) = \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2$$

and if $f \in R(\alpha)$ and c a positive constant, then

$$f \in R(c\alpha) \text{ and } \int_a^b f \, d(c\alpha) = c\int_a^b f \, d\alpha .$$

Proof. Since $f \in R(\alpha_1)$ and $f \in R(\alpha_2)$, therefore for $\varepsilon > 0$, $\exists$ partitions P_1, P_2 of [a, b] such that

$$U(P_1, f, \alpha_1) - L(P_1, f, \alpha_1) < \tfrac{1}{2}\varepsilon$$

$$U(P_2, f, \alpha_2) - L(P_2, f, \alpha_2) < \tfrac{1}{2}\varepsilon$$

Let $P = P_1 \cup P_2$

$\therefore \quad U(P, f, \alpha_1) - L(P, f, \alpha_1) < \frac{1}{2}\varepsilon$

$U(P, f, \alpha_2) - L(P, f, \alpha_2) < \frac{1}{2}\varepsilon$ ...(1)

Let the partition P be $\{a = x_0, x_1, x_2, \ldots, x_n = b\}$, and m_i, M_i be bounds of f in Δx_i.

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

$\therefore \quad \alpha(x) = \alpha_1(x) + \alpha_2(x)$

$\Delta\alpha_{1i} = \alpha_1(x_i) - \alpha_1(x_{i-1})$

$\Delta\alpha_{2i} = \alpha_2(x_i) - \alpha_2(x_{i-1})$

$\therefore \quad \Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$

$= \alpha_1(x_i) + \alpha_2(x_i) - \alpha_1(x_{i-1}) - \alpha_2(x_{i-1})$

$= \Delta\alpha_{1i} + \Delta\alpha_{2i}$

$\therefore \quad U(P, f, \alpha) = \sum_i M_i \, \Delta\alpha_i$

$= \sum_i M_i \, (\Delta\alpha_{1i} + \Delta\alpha_{2i})$

$= U(P, f, \alpha_1) + U(P, f, \alpha_2)$ ...(2)

Similarly,

$L(P, f, \alpha) = L(P, f, \alpha_1) + L(P, f, \alpha_2)$ ...(3)

$\therefore \quad U(P, f, \alpha) - L(P, f, \alpha) = U(P, f, \alpha_1) - L(P, f, \alpha_1)$

$+ U(P, f, \alpha_2) - L(P, f, \alpha_2)$

$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$ [using (1)]

$\Rightarrow \quad f \in R(\alpha)$, where $\alpha = \alpha_1 + \alpha_2$

Now ,we notice that

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

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

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

$$= \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2 \qquad \ldots(4)$$

Similarly,

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

$$\leq \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2 \qquad \ldots(5)$$

From (4) and (5),

$$\int_a^b f \, d\alpha = \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2$$

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

Integral as a limit sum.

For any partition P of [a, b] and $t_i \in \Delta x_i$, consider the sum

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

We say that $S(P, f, \alpha)$ converges to A as $\mu(P) \to 0$ if for every $\in > 0$ there exists $\delta > 0$ such that $|S(P, f, \alpha) - A| < \in$, for every partition $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$, of [a, b], with mesh $\mu(P) < \delta$ and every choice of t_i in Δx_i.

Theorem . If $S(P, f, \alpha)$ converges to A as $\mu(P) \to 0$, then

$$f \in R(\alpha), \text{ and } \lim_{\mu(P) \to 0} S(P, f, \alpha) = \int_a^b f \, d\alpha$$

Proof. Let us suppose that $\lim S(P, f, \alpha)$ exists as $\mu(P) \to 0$ and is equal to A. Therefore for $\in > 0$, $\exists \, \delta > 0$ such that for every partition P of [a, b]with mesh $\mu(P) < \delta$ and every choice of t_i in Δx_i, we have

$$|S(P, f, \alpha) - A| < \tfrac{1}{2} \in$$

or

$$A - \tfrac{1}{2}\in < S(P, f, \alpha) < A + \tfrac{1}{2}\in \qquad \ldots(1)$$

Let P be a partition. If we let the points t_i range over the interval Δx_i and take the infimum and the supremum of the sums $S(P, f, \alpha)$, (1) yields

$$A - \tfrac{1}{2}\in < L(P, f, \alpha) \le U(P, f, \alpha) < A + \tfrac{1}{2}\in \qquad \ldots(2)$$

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

$\Rightarrow \quad f \in R(\alpha)$ over $[a, b]$

Again, since $S(P, f, \alpha)$ and $\int_a^b f\, d\alpha$ lie between $U(P, f, \alpha)$ and $L(P, f, \alpha)$

$$\therefore \quad \left| S(P, f, \alpha) - \int_a^b f\, d\alpha \right| \le U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

$$\Rightarrow \quad \lim_{\mu(P)\to 0} S(P, f, \alpha) = \int_a^b f\, d\alpha$$

Theorem . If f is continuous on [a, b], then $f \in R(\alpha)$ over [a, b]. Also, to every $\varepsilon > 0$ there corresponds a $\delta > 0$ such that

$$\left| S(P, f, \alpha) - \int_a^b f\, d\alpha \right| < \varepsilon$$

for every partition $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ of [a, b] with $\mu(P) < \delta$, and for every choice of t_i in Δx_i, i.e.,

$$\lim_{\mu(P)\to 0} S(P, f, \alpha) = \int_a^b f\, d\alpha$$

Proof. Let $\varepsilon > 0$ be given, and let $\eta > 0$ such that

$$\eta\{\alpha(b) - \alpha(a)\} < \varepsilon \qquad \ldots(1)$$

Since continuity of f on the closed interval [a, b] implies its uniform continuity on [a, b], therefore for $\eta > 0$ there corresponds $\delta > 0$ such that

$$|f(t_i) - f(t_2)| < \eta, \quad \text{if } |t_1 - t_2| < \delta, \; t_1, t_2 \in [a, b] \qquad \ldots(2)$$

Let P be a partition of [a, b], with norm $\mu(P) < \delta$.

Then by (2),

$M_i - m_i \le \eta$, i = 1, 2,…, n

$\therefore$ $U(P, f, \alpha) - L(P, f, \alpha) = \sum_i (M_i - m_i)\Delta x_i$

$$\le \eta \sum_i \Delta x_i$$

$$= \eta(\alpha(b) - \alpha(a)\} < \varepsilon \qquad \ldots(3)$$

$\Rightarrow$ $f \in R(\alpha)$ over [a, b].

Again if $f \in R(\alpha)$, then for $\varepsilon > 0$, $\exists$ $\delta > 0$ such that for all partitions P with $\mu(P) < \delta$,

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

Since S(P, f, α) and $\int_a^b f\, d\alpha$ both lie between U(P, f, α) and L(P, f, α) for all partitions P with $\mu(P) < \delta$ and for all positions of t_i in Δx_i.

$$\therefore \quad \left| S(P, f, \alpha) - \int_a^b f\, d\alpha \right| < U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

$$\Rightarrow \quad \lim_{\mu(P)\to 0} S(P, f, \alpha) = \lim_{\mu(P)\to 0} \sum_{i=1}^{n} f(t_i)\ \Delta\alpha_i = \int_a^b f\, d\alpha$$

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

Proof. Let $\varepsilon > 0$ be a given positive number.

For any positive integer n, choose a partition $P = \{x_0, x_1,\ldots, x_n\}$ of [a, b] such that

$$\Delta\alpha_i = \frac{a(b) - \alpha(a)}{n}, \ i = 1, 2,\ldots,n$$

This is possible because α is continuous and monotonic increasing on the closed interval [a, b] and thus assumes every value between its bounds, α(a) and α(b).

Let f be monotonic increasing on [a, b], so that its lower and the upper bound, m_i, M_i, in Δx_i are given by

$$m_i = f(x_{i-1}),\ M_i = f(x_i),\ \ i = 1, 2,\ldots, n$$

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

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

$$= \frac{\alpha(b)-\alpha(a)}{n}\{f(b) - f(a)\}$$

$$< \varepsilon, \text{ for large } n$$

$\Rightarrow$ $f \in R(\alpha)$ over [a, b]

Example . Let a function α increase on [a, b] and is continuous at x' where $a \le x' \le b$ and a function f is such that

$$f(x') = 1, \text{ and } f(x) = 0, \text{ for } x \neq x'$$

then $f \in R(\alpha)$ over [a, b], and $\int_a^b f \, d\alpha = 0$

Solution. Let $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ be a partition of [a, b] and let $x' \in \Delta x_i$.

But since α is continuous at x' and increases on [a, b], therefore for $\varepsilon > 0$ we can choose $\delta > 0$ such that

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) < \varepsilon, \text{ for } \Delta x_i < \delta$$

Let P be a partition with $\mu(P) < \delta$. Now

$$U(P, f, \alpha) = \Delta\alpha_i$$

$$L(P, f, \alpha) = 0$$

$$\therefore \qquad \overline{\int_a^b} f \, d\alpha = \inf U(P, f, \alpha), \text{ over all partitions P with } \mu(P) < \delta$$

$$= 0 = \underline{\int_a^b} f \, d\alpha$$

$\Rightarrow$ $f \in R(\alpha)$, and $\int_a^b f \, d\alpha = 0$.

Theorem . If $f \in R[a, b]$ and α is monotone increasing on [a, b] such that $\alpha' \in R[a, b]$, then $f \in R(\alpha)$, and

$$\int_a^b f\, d\alpha = \int_a^b f\alpha'\, dx$$

Proof. Let $\varepsilon > 0$ be any given number.

Since f is bounded, there exists $M > 0$, such that

$$|f(x)| \le M, \quad \forall\, x \in [a, b]$$

Again since $f, \alpha' \in R[a, b]$, therefore $f\alpha' \in R[a, b]$ and consequently $\exists\, \delta_1 > 0$, $\delta_2 > 0$ such that

$$\left|\Sigma(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha'\, dx\right| < \varepsilon/2 \qquad \ldots(1)$$

for $\mu(P) < \delta_1$ and all $t_i \in \Delta x_i$, and

$$\left|\Sigma\alpha'(t_i)\Delta x_i - \int \alpha'\, dx\right| < \varepsilon/4M \qquad \ldots(2)$$

for $\mu(P) < \delta_2$ and all $t_i \in \Delta x_i$

Now for $\mu(P) < \delta_2$ and all $t_i \in \Delta x_i$, $s_i \in \Delta x_i$, (2) gives

$$\sum|\alpha'(t_i) - \alpha'(s_i)| < 2.\ \frac{\varepsilon}{4M} = \frac{\varepsilon}{2M} \qquad \ldots(3)$$

Let $\delta = \min(\delta_1, \delta_2)$, and P any partition with $\mu(P) < \delta$.

Then, for all $t_i \in \Delta x_i$, by Lagrange's Mean value Theorem, there are points $s_i \in \Delta x_i$ such that

$$\Delta\alpha_i = \alpha'(s_i)\, \Delta x_i \qquad \ldots(4)$$

Thus

$$\left|\Sigma f(t_i)\Delta\alpha_i - \int \alpha'\, dx\right| = \left|\Sigma f(t_i)\alpha'(s_i)\Delta x_i - \int f\alpha' dx\right|$$

$$= \left|\Sigma f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha' dx + \Sigma f(t_i)[\alpha'(s_i) - \alpha'(t_i)]\Delta x_i\right|$$

$$\le \left|\Sigma f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha'\, dx\right|$$

$$+ \sum |f(t_i)|\, |\alpha'(s_i) - \alpha'(t_i)|\, \Delta x_i$$

$$< \frac{\varepsilon}{2} + M\frac{\varepsilon}{2M} = \varepsilon$$

Hence for any $\varepsilon > 0$, $\exists\, \delta > 0$ such that for all partitions with $\mu(P) < \delta$, (5) holds

$\Rightarrow \quad \lim_{\mu(P)\to 0} \sum f(t_i)\, \Delta\alpha_i$ exists and equals $\int_a^b f\alpha' \, dx$

$\Rightarrow \quad f \in R(\alpha)$, and $\int_a^b f \;\; d\alpha \;=\; \int^b$

Theorem . If f is continuous on [a, b] and α a continuous derivative on [a, b], then

$$\int_a^b f \, d\alpha = \int_a^b f\alpha' \, dx$$

Proof. Let $P = \{a = x_0, \ldots., x_n = b\}$ be any partition of [a, b]. Thus, by Lagrange's Mean value Theorem it is possible to find $t_i \in \,]x_{i-1}, x_i[$, such that

$$\alpha(x_i) - \alpha(x_{i-1}) = \alpha'(t_i)\,(x_i - x_{i-1}),\; i = 1, 2, \ldots, n$$

or

$$\Delta\alpha_i = \alpha'(t_i)\, \Delta x_i$$

$$\therefore \quad S(P, f, \alpha) = \sum_{i=1}^{n} f(t_i)\, \Delta\alpha_i$$

$$= \sum_{i=1}^{n} f(t_i)\; \alpha'(t_i)\, \Delta x_i = S(P, f\alpha') \qquad \ldots(6)$$

Proceeding to limits as $\mu(P) \to 0$, since both the limits exist, we get

$$\int_a^b f \, d\alpha = \int_a^b f\alpha' \, dx$$

Examples.

(i) $\int_0^2 x^2 \, dx^2 = \int_0^2 x^2 \; 2x \, dx = \int_0^2 2x^3 \, dx = 8$

(ii) $\int_0^2 [x] dx^2 = \int_0^2 [x] 2x \, dx$

$$= \int_0^1 [x] 2x \, dx + \int_1^2 [x] 2x \;\, dx = 0 \quad 3 = 3$$

Ex. Evaluate the following integrals :

(i) $\int_1^4 (x-[x])dx^2$ (ii) $\int_0^3 \sqrt{x}\ dx^3$

(iii) $\int_0^3 [x]\ d(e^x)$ (iv) $\int_0^{\pi/2} x\ d(\sin x)$

Theorem . (First Mean Value Theorem). If a function f is continuous in [a, b] and α is monotonic increasing on [a, b], then there exists a number ξ in [a, b] such that

$$\int_a^b f\ d\alpha = f(\xi)\ [\alpha(b) - \alpha(a)\}$$

f is continuous and α is monotonic, therefore $f \in R(\alpha)$.

Proof. Let m, M be the infimum and supremum of f in [a, b]. Then

$$m\{(b) - \alpha(a)\} \le \int_a^b f\ d\alpha \le M\{\alpha(b) - \alpha(a)\}$$

Hence there exists a number μ, $m \le \mu \le M$ such that

$$\int_a^b f\ d\alpha = \mu\{\alpha(b) - \alpha(a)\}$$

Again, since f is continuous, there exists a number $\xi \in [a, b]$ such that $f(\xi) = \mu$

$$\therefore \qquad \int_a^b f\ d\alpha = f(\xi)\ \{\alpha(b) - \alpha(a)\}$$

Remark. It may not be possible always to choose ξ such that $a < \xi < b$.

Consider $\alpha(x) = \begin{cases} 0, & x = a \\ 1, & a < x \le b \end{cases}$

For a continuous function f, we have

$$\int_a^b f\ d\alpha\ = f(a) = f(a)\ \{\alpha(b) - \alpha(a)\}$$

Theorem . If f is continuous and α monotone on [a, b], then

$$\int_a^b f\ d\alpha = \left[f(x)\ \alpha(x)\right]_a^b - \int_a^b \alpha\ df$$

Proof. Let P = {a = x_n, x_1,…., x_n =b} be a partition of [a, b].

Let $t_1, t_2, \ldots, t_n$ such that $x_{i-1} \le t_i \le x_i$, and let $t_0 = a$, $t_{n+1} = b$, so that $t_{i-1} \le x_{i-1} \le t_i$.

Then Q = {a = $t_0, t_1, t_2, \ldots, t_n, t_{n+1}$} is also a partition of [a, b]

Now

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

$$f(t_1)\,[\alpha(x_1) - \alpha(x_0)] + f(t_2)[\alpha(x_2) - \alpha(x_1)] + \ldots$$

$$+ f(t_n)\,[\alpha(x_n) - \alpha(x_{n-1})]$$

$$= -\alpha(x_0)\,f(t_1) - \alpha(x_1)\,[f(t_2) - f(t_1)]$$

$$+ \alpha(x_2)\,[f(t_3) - f(t_2)] + \ldots$$

$$+ \alpha(x_{n-1})\,[f(t_n) - f(t_{n-1})] + \alpha\,(x_n)\,f(t_n)$$

Adding and subtracting $\alpha(x_0)\,f(t_0) + \alpha(x_n)\,f(t_{n+1})$, we get

$$S(P, f, \alpha) = \alpha(x_n)\,f(t_{n+1}) - \alpha(x_0)\,f(t_0) - \sum_{i=0}^{n} \alpha\,(x_i)\,\{f(t_{i+1}) - f(t_i)\}$$

$$= f(b)\,\alpha(b) - f(a)\,\alpha(a) - S(Q, \alpha, f) \qquad \ldots(1)$$

If $\mu(P) \to 0$, then $\mu(Q) \to 0$ and Theorem 5 shows that lim S(P, f, α) and lim S(Q, α, f) both exist and that

$$\lim S(P, f, \alpha) = \int_a^b f\;d\alpha$$

and

$$\lim S(Q, \alpha, f) = \int_a^b \alpha\;df$$

Hence proceeding to limits when $\mu(P) \to 0$, we get from (1),

$$\int_a^b f\;d\alpha = \big[f(x)\,\alpha(x)\big]_a^b - \int_a^b \alpha\;df \qquad \ldots(2)$$

where $\big[f(x)\,\alpha(x)\big]_a^b$ denotes the difference f(b) α(b) – f(a) α(a).

Corollary. The result of the theorem can be put in a slightly different form, by using Theorem 9, if, in addition to monotonicity α is continuous also

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

$$= f(b)\,\alpha(b) - f(a)\,\alpha(a) - \alpha(\xi)\,[f(b) - f(a)]$$

$$= f(a)\,[\alpha(\xi) - \alpha(a)] + f(b)\,[\alpha(b) - \alpha(\xi)]$$

where $\xi \in [a, b]$.

Stated in this form, it is called the **Second Mean Value Theorem**.

Integration and Differentiation.

Definition. If $f \in \pm$ 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. 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

$$|F(c + h) - F(c) = |\int_a^{c+h} f(x)dx - \int_a^c f(x)\,dx$$

$$= \left|\int_a^{c+h} f(x)dx\right|$$

$$\leq M\,|h|$$

$$< \in \text{ if } |h| < \frac{\in}{M}$$

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. If f is continuous on [a, b], then the integral function F is differentiable and $F'(x_0) = f(x_0)$, $x_0 \in [a, b]$.

Proof. Let f be continuous at x_0 in [a, b]. Then there exists $\delta > 0$ for every $\in > 0$ such that

(1) $\qquad |f(t) - f(x_0)| < \in$

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

$$\frac{F(t) - F(s)}{t - s} - f(x_0) = \left| \frac{1}{t - s} \int_s^t f(x)dx - f(x_0) \right|$$

$$= \left| \frac{1}{t - s} \int_s^t f(x)dx - \frac{1}{t - s} \int_s^t f(x_0)dx \right|$$

$$= \frac{1}{t - s} \left| \int_s^t [f(x) - f(x_0)]dx \right| \leq \frac{1}{t - s} \left| \int_s^t f(x) - f(x_0) \right| dx < \in,$$

(using (1)).

Hence $F'(x_0) = f(x_0)$. This completes the proof of the theorem.

Theorem (Fundamental Theorem of the Integral (Calculus). If $f \in \Re$ 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 $x_{i-1} \leq t_i \leq x_i$. 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

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

$$= \sum_{i=1}^{n} f(t_i)\ (x_i - x_{i-1})$$

$$= \sum_{i=1}^{n} f(t_i)\ \Delta x_i$$

and the last sum tends to $\int_a^b f(x)\ dx$ as $|P+ \to 0$, taking $\alpha(x) = x$. Hence

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

Integration of Vector – Valued Functions. Let f_1. $f_2, \ldots, f_k$ be real valued functions defined on [a, b] and let $\mathbf{f} = (f_1, f_2, \ldots, f_k)$ 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 $\mathbf{f}$ 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 its 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) $$\int_a^b (f + g) d\alpha = \int_a^b f\, d\alpha + \int_a^b g\ d\alpha$$

(ii) $$\int_a^b f\ d\alpha = \int_a^c f\ d\alpha + \int_a^b f\ d\alpha, a < c < b.$$

(iii) if $\mathbf{f} \in \Re(\alpha_1)$, $\mathbf{f} \in \Re(\alpha_2)$, then $\mathbf{f} \in \Re(\alpha_1 + \alpha_2)$

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

Theorem. If **f** and **F** map [a, b] into $\Re^k$, if $\mathbf{f} \in \Re(\alpha)$ if $\mathbf{F}' = \mathbf{f}$, then

$$\int_a^b f(t)\, dt = \mathbf{F}(b) - \mathbf{F}(a)$$

Theorem. 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 $|f| \in \mathbf{R}(\alpha)$ and

$$\left|\int_a^b f\, d\alpha\right| \le \int_a^b |\mathbf{f}|\, d\alpha.$$

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 $t_i^2 \in \mathbf{R}(\alpha)$ 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 of 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

$$\mathbf{y} = \int f\, d\alpha$$

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

$$(1) \qquad |y|^2 \le |y| \int |f|\, d\alpha$$

If $\mathbf{y} = \mathbf{0}$, then the result follows. If $\mathbf{y} \neq \mathbf{0}$, then divide (1) by $|y|$ and get

$$|y| \le \int |f|\, d\alpha$$

or

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

Rectifiable Curves.

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 $\mathbf{f}$ on $[a, b]$.

The function $\mathbf{f}$ 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., $V(\gamma, a, b)$.

Theorem. 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. We have to show that $\int |\gamma'| = V(\gamma, a, b)$. Let $\{x_0, \ldots, x_n\}$ be a partition of $[a, b]$.

By 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

$$V(\gamma, a, b) \leq \int |\gamma'| \qquad (1)$$

Conversely. 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 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

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)|\ dt - \in \Delta x_i \leq \gamma'(x_i)|\ \Delta\ x_1$$

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

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

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

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

$$\int_a^b |\gamma'(t)|\, dt \le V(\gamma, a, b) \qquad (2)$$

Combining ((1) and (2), we have

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

Therefore, length of $\gamma = \int_a^b |\gamma'(t)|\, dt$.

MAL-512: M. Sc. Mathematics (Real Analysis)

Lesson No. VII **Written by Dr. Nawneet Hooda**

Lesson: Measure Theory **Vetted by Dr. Pankaj Kumar**

Measurable Sets

Definition. The length of an interval I =[a,b] is defined to be the difference of the endpoints of the interval I and is written as l(I)= b - a.

The interval I may be closed, open, open-closed or closed-open, the length l(I) is always equals b–a, where a < b. In case a = b, the interval [a, b] becomes a point with length zero.

Definition. A function whose domain of definition is a class of sets is called a set function. In the case of length, the domain is the collection of all intervals.

In the above, we have said that in the case of length, the domain is the collection of all intervals.

Definition. **(** length of a set**).** Let A be an open set in R and let A be written as a countable union of mutually disjoint open intervals $\{I_i\}$ i.e.,

$$A = \bigcup_i I_i .$$

Then the length of the open set A is defined by

$$l(A) = \sum_i l(I_i) .$$

Also, if A_1 and A_2 are two open sets in R such that $A_1 \subset A_2$, then

$$l(A_1) \le l(A_2).$$

Hence, for any open set A in [a, b], we have

$$0 \le l(A) \le b - a.$$

Definition. **(**outer measure) The Lebesgue outer measure or simply the outer measure m*(A) of an arbitrary set A is given by

$$m^*(A) = \inf \sum_i l(I_i),$$

where the infimum is taken over all countable collections $\{I_i\}$ of open intervals such that $A \subset \bigcup_i I_i$.

Remark. The outer measure m* is a set function which is defined from the power set P(R) into the set of all non-negative extended real numbers.

Theorem (a) $m^*(A) \geq 0$, for all sets A.

(b) $m^*(\phi) = 0$.

(c) If A and B are two sets with $A \subset B$, then $m^*(A) \leq m^*(B)$.

(d) $m^*(A) = 0$, for every singleton set A.

(e) m* is translation invariant, i.e., $m^*(A + x) = m^*(A)$, for every set A and for every $x \in R$.

Proof. (a) clear by definition .

(b) Since $\phi \subset I_n$ for every open interval in R such that

$$I_n = \left]x - \frac{1}{n}, x + \frac{1}{n}\right[$$

So $0 \leq m^*(\phi) \leq l(I_n)$ .and $0 \leq m^*(\phi) \leq \frac{2}{n}$, for each $n \in N$

(c) Let $\{I_n\}$ be a countable collection of disjoint open intervals such that $B \subset \bigcup_n I_n$. Then $A \subset \bigcup_n I_n$ and therefore

$$m^*(A) \leq \sum_n l(I_n) .$$

Since ,this inequality is true for any coverings $\{I_n\}$ of B, the result follows.

(This property is known as monotonicity.)

(d) Since

$$\{x\} \subset I_n = \left]x - \frac{1}{n}, x + \frac{1}{n}\right[$$

is an open covering of $\{x\}$, so $0 \leq m^*(\{x\}) \leq l(I_n)$ and $l(I_n) = \frac{2}{n}$, for each $n \in N$, the result follows by using (a).

(e) Let I be any interval with end points a and b, the set I + x defined by

$$I + x = \{y + x : y \in I\}$$

is an interval with endpoints a + x and b + x. Also,

$$l(I + x) = l(I).$$

Now, let $\in > 0$ be given. Then there is a countable collection $\{I_n\}$ of open intervals such that $A \subset \bigcup_n I_n$ and satisfies

$$\sum_n l(I_n) \le m^*(A) + \in.$$

Also $A + x \subset \bigcup_n (I_n + x)$. Therefore

$$m^*(A + x) \le \sum_n l(I_n + x) = \sum_n l(I_n) \le m^*(A) + \in$$

Since $\in > 0$ is arbitrary, we have $m^*(A + x) \le m^*(A)$. If we take $A = (A + x) - x$ and use the above arguments the reverse inequality follows.

Theorem. The outer measure of an interval I is its length.

Proof. Case 1. First let I be a closed finite interval [a, b]. Since, for each $\in > 0$, the open interval $(a - \frac{\in}{2}, b + \frac{\in}{2})$ covers [a, b], we have

$$m^*(I) \le l(a - \frac{\in}{2}, b + \frac{\in}{2}) = b - a + \in.$$

Since this is true for each $\in > 0$, we must have

$$m^*(I) \le b - a = l(I).$$

Now we will prove that

$$m^*(I) \ge b - a \qquad \ldots(1)$$

Let $\in > 0$ be given. Then there exists a countable collection $\{I_n\}$ of open intervals covering [a, b] such that

$$m^*(I) > \sum_n l(I_n) - \in \qquad \ldots(2)$$

By the Heine-Borel Theorem, any collection of open intervals which cover [a, b] has a finite sub-cover which covers [a, b], it suffices to establish the inequality (2) for finite collections $\{I_n\}$ which cover [a, b].

Since $a \in [a, b]$, there must be one of the intervals I_n which contains a and let it be (a_1, b_1). Then $a_1 < a < b_1$. If $b_1 \leq b$, then $b_1 \in [a, b]$, and since $b_1 \notin (a_1, b_1)$, there must be an interval (a_2, b_2) in the finite collection $\{I_n\}$ such that $b_1 \in (a_2, b_2)$; that is $a_2 < b_1 < b_2$. Continuing in this manner, we get intervals (a_1, b_1), (a_2, b_2),… from the collection $\{I_n\}$ such that

$$a_i < b_{i-1} < b_i, \quad i = 1, 2, \ldots$$

where $b_0 = a$. Since $\{I_n\}$ is a finite collection, this process must terminate with some interval (a_k, b_k) in the collection . Thus

$$\begin{aligned} \sum_n l(I_n) &\geq \sum_{i=1}^{k} l(a_i, b_i) \\ &= (b_k - a_k) + (b_{k-1} - a_{k-1}) + \ldots + (b_1 - a_1) \\ &= b_k - (a_k - b_{k-1}) - \ldots - (a_2 - b_1) - a_1 \\ &> b_k - a_1 \\ &> b - a, \end{aligned}$$

since $a_i - b_i\ 1 < 0$, $b_k > b$ and $a_1 < a$. This, in view of (2), verifies that

$$m^*(I) > b - a - \in.$$

Hence $m^*(I) \geq b - a$.

Case 2. Now let I is any finite interval. Then given an $\in > 0$, there exists a closed finite interval $J \subset I$ such that

$$l(J) > l(I) - \in.$$

Therefore,

$$l(I) - \in < l(J) = m^*(J) \leq m^*(I) \leq m^*(I) = l(I) = l(I)$$

$$\Rightarrow \qquad l(I) - \in < m^*(I) \leq l(I).$$

This is true for each $\in > 0$. Hence $m^*(I) = l(I)$.

Case 3. Suppose I is an infinite interval. Then given any real number $r > 0$, there exists a closed finite interval $J \subset I$ such that $l(J) = r$. Thus $m^*(I) \geq m^*(J) = l(J) = r$, that is $m^*(I) \geq r$ for any arbitrary real number $r > 0$. Hence $m^*(I) = \infty = l(I)$.

Theorem. Let $\{A_n\}$ be a countable collection of sets of real numbers. Then

$$m^*(\bigcup_n A_n) \le \sum_n m^*(A_n)$$

Proof. If m*(A_n) = ∞ for some n∈N, the inequality holds . Let us assume that m*(A_n) < ∞, for each n∈N. Then, for each n, and for a given ∈ > 0, ∃ a countable collection $\{I_{n,i}\}_i$ of open intervals such that that $A_n \subset \bigcup_i I_{n,i}$ satisfying

$$\sum_i l(I_{n,i}) < m^*(A_n) + 2^{-n} \in$$

Then

$$\bigcup_n A_n \subset \bigcup_n \bigcup_i I_{n,i} .$$

However, the collection $\{I_{n-i}\}_{n,i}$ forms a countable collection of open intervals, as the countable union of countable sets is countable and covers $\bigcup_n A_n$. Therefore

$$m^*(\bigcup_n A_n) \le \sum_n \sum_i l(I_{n,i})$$

$$\le \sum_n (m^*(A_n) + 2^{-n} \in)$$

$$= \sum_n m^*(A_n) + \in.$$

But ∈ > 0 being arbitrary, the result follows.

(This theorem shows that m* is countable sub-additivite).

Note. Each of the sets N, I, Q has outer measure zero since each one is countable..

Corollary . If a set A is countable ,then m* (A)=0.

Proof. Since every countable set can be written as the union of singleton sets and m* of a singleton set is zero, corollary follows by above theorem and the definition of m* .

Corollary. The set [0, 1] is uncountable.

Proof**.** Let the set [0, 1] be countable. The m*([0, 1]) = 0 and so *l*([0, 1]) = 0. This is absurd as the length is equal to 1. Hence the set [0, 1] is uncountable.

Theorem. The Cantor set C is uncountable with outer measure zero.

Proof. Let E_n be the union of the intervals left at the nth stage while constructing the Cantor set C . E_n consists of 2^n closed intervals, each of length 3^{-n}. Therefore

$$m^*(E_n) \leq 2^n . 3^{-n}.$$

But each point of C must be in one of the intervals comprising the union E_n, for each $n \in N$, and as such $C \subset E_n$, for all $n \in N$. Hence

$$m^*(E) \leq \left(\frac{2}{3}\right)^n .$$

This being true for each $n \in N$, letting $n \to \infty$ gives $m^*(E) = 0$.

Theorem. If $m^*(A) = 0$, then $m^*(A \cup B) = m^*(B)$.

Proof. By countable subadditivity of m* and $m^*(A) = 0$ we have

$$m^*(A \cup B) \leq m^*(A) + m^*(B) = m^*(B), \qquad (1)$$

But $B \subset A \cup B$ gives

$$m^*(B) \leq m^*(A \cup B). \qquad (2)$$

Hence the result follows by (1) and (2).

LEBESGUE MEASURE

The outer measure does not satisfy the countable additivity . To have the property of countable additivity satisfied, we restrict the domain of definition for the function m* to some suitable subset, *M*, of the power set *P(R)*. The members of *M* are called measurable sets and we defined as :

Definition. A set E is said to be **Lebesgue measurable** or simply measurable if for each set A, we have

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c). \qquad \ldots(3)$$

Since $A = (A \cap E) \cup (A \cap E^c)$ and m* is subadditive, we always have

$$m^*(A) \leq m^*(A \cap E) + m^*(A \cap E^c).$$

Thus to prove that E is measurable, we have to show, for any set A, that

$$m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c). \qquad \ldots(4)$$

The set A in reference is called test set.

Definition. The restriction of the set functions m* to that to m of measurable sets, is called Lebesgue measure function for the sets in *M*.

So, for each $E \in \mathcal{M}$, m(E) = m*(E). The extended real number m(E) is called the Lebesgue measure or simply measure of the set E.

Theorem . If E is a measurable set, then so is E^c.

Proof. If E is measurable, then for for any set A,

$$m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c).$$
$$= m^*(A \cap E^c) + m^*(A \cap E)$$
$$= m^*(A \cap E^c) + m^*(A \cap E^{cc})$$

Hence E^c is measurable.

Remark. The sets ϕ and R are measurable sets.

Theorem. If m*(E) = 0, then E is a measurable set.

Proof. Let A be any set. Then

$$A \cap E \subset E \Rightarrow m^*(A \cap E) \leq m^*(E) = 0$$

and $$A \cap E^c \subset A \Rightarrow m^*(A \cap E^c) \leq m^*(A).$$

Therefore $$m^*(A) \geq m^*(A \cap E) + m^*(A \cap E^c),$$

Hence E is measurable.

Theorem. If E_1 and E_2 are measurable sets, then so is $E_1 \cup E_2$.

Proof. Since E_1 and E_2 are measurable sets, so for any set A, we have

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c)$$
$$= m^*(A \cap E_1) + m^*([A \cap E_1^c] \cap E_2) + m^*([A \cap E_1^c] \cap E_2^c)$$
$$= m^*(A \cap E_1) + m^*([A \cap E_2] \cap E_1^c) + m^*(A \cap E_1^c \cap E_2^c)$$
$$= m^*(A \cap E_1) + m^*(A \cap E_2 \cap E_1^c) + m^*(A \cap [E_1 \cap E_2]^c)$$

But $A \cap (E_1 \cup E_2) = [A \cap E_1] \cup [A \cap E_2 \cap E_1^c]$ and

$$m^*(A \cap (E_1 \cup E_2)) \leq m^*[A \cap E_1] + m^*[A \cap E_2 \cap E_1^c].$$

Therefore $$m^*(A) \geq m^*(A \cap [E_1 \cup E_2]) + m^*(A \cap [E_1 \cup E_2]^2),$$

Hence $E_1 \cup E_2$ is a measurable set.

Theorem . The intersection and difference of two measurable sets are measurable.

Proof. For two sets E_1 and E_2, we can write $(E_1 \cap E_2)^c = E_1^c \cup E_2^c$ and $E_1 - E_2 = E_1 \cap E_2^c$]. Now using the fact that union of two measurable sets is measurable and complement of a measurable sets is measurable, we get the result.

Theorem. The symmetric difference of two measurable sets is measurable.

Proof. The symmetric difference of two sets E_1 and E_2 is given by $E_1 \Delta E_2 = (E_1 - E_2) \cup (E_2 - E_1)$ and by arguments we get the result.

Definition. A nonempty collection *A* of subsets of a set S is called an algebra (or Boolean algebra) of sets in P(S) if $\phi \in A$ and

(a) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

(b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.

By DeMorgan's law it follows that if A is an algebra of sets in P(S), then

(c) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$,

while, on the other hand, if any collection A of subsets of S satisfies (b) and (c), then it also satisfies (a) and hence A is an algebra of sets in P(S).

Corollary. The family *M* of all measurable sets (subsets of R) is an algebra of sets in *P(R)*. In particular, if $\{E_1, E_2, \ldots, E_n\}$ is any finite collection of measurable sets, then so are $\bigcup_{i=1}^{n} E_1$ and $\bigcap_{i=1}^{n} E_i$.

Theorem. Let $E_1, E_2, \ldots, E_n$ be a finite sequence of disjoint measurable sets. Then, for any set A,

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

Proof. We use induction on n. For n = 1, the result is clearly true. Let it be true for (n−1) sets, then we have

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n-1} E_i\right]\right) = \sum_{i=1}^{n-1} m^*(A \cap E_i).$$

Adding $m^*(A \cap E_n)$ on both the sides and since the sets $E_i (i = 1, 2, \ldots, n)$ are disjoint we get

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n-1} E_i\right]\right) + m^*(A \cap E_n) = \sum_{i=1}^{n} m^*(A \cap E_i)$$

$$\Rightarrow \qquad m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n^c\right) + \left(A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n\right) = \sum_{i=1}^{n} m^*(A \cap E_i),$$

. But the measurability of the set E_n, by taking $A \cap \left[\bigcup_{i=1}^{n} E_i\right]$ as a test set, we get

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) = m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n\right) + m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right] \cap E_n^c\right)$$

Hence

$$m^*\left(A \cap \left[\bigcup_{i=1}^{n} E_i\right]\right) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

Corollary. If $E_1, E_2, \ldots, E_n$ is a finite sequence of disjoint measurable sets, then

$$m\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{1=i}^{n} m(E_i).$$

Theorem. If E_1 and E_2 are any measurable sets, then

$$m(E_1 \cup E_2) + m(E_1 \cap E_2) = m(E_1) + m(E_2).$$

Proof. Let A be any set. Since E_1 is a measurable set, we have

$$m^*(A) = m^*(A \cap E_1) + m^*(A \cap E_1^c)$$

Take $A = E_1 \cup E_2$, and adding $m(E_1 \cap E_2)$ on both sides, we get

$$m(E_1 \cup E_2) + m(E_1 \cap E_2) = m(E_1) + m(([E_1 \cup E_2] \cap E_1^c) + m(E_1 \cap E_2).$$

Since

$$[(E_1 \cup E_2) \cap E_1^c] \cup [E_1 \cap E_2] = E_2$$

is a union of disjoint measurable sets, we note that

$$m([E_1 \cup E_2] \cap E_1^c) + m(E_1 \cap E_2) = m(E_2).$$

Hence the result follows.

Theorem Let *A* be an algebra of subsets of a set S. If $\{A_i\}$ is a sequence of sets in *A*, then there exists a sequence $\{B_i\}$ of mutually disjoint sets in *A* such that

$$\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i .$$

Proof. If the sequence $\{A_i\}$ is finite, the result is clear. Now let $\{A_i\}$ be an infinite sequence. Set $B_1 = A_1$, and for each $n \geq 2$, define

$$B_n = A_n - \left[\bigcup_{i=1}^{n-1} A_i\right]$$

$$= A_n \cap A_1^c \cap A_2^c \cap \ldots \cap A_{n-1}^c$$

Note that

(i) $B_n \in A$, for each $n \in N$, since *A* is closed under the complementation and finite intersection of sets in *A*.

(ii) $B_n \subset A_n$, for each $n \in N$.

(iii) $B_m \cap B_n = \phi$ for $m \neq n$. i.e.the sets B_n are mutually disjoint.

Let B_n and B_m to be two sets and with $m < n$. Then, because $B_m \subset A_m$, we have

$$B_m \cap B_n \subset A_m \cap B_n$$

$$= A_m \cap [A_n \cap A_1^c \cap \ldots \cap A_m^c \cap \ldots \cap A_{n-1}^c]$$

$$= [A_m \cap A_m^c] \cap \ldots$$

$$= \phi \cap \ldots$$

$$= \phi.$$

(iv) $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i .$

Since $B_i \subset A_i$, for each $i \in N$, we have

$$\bigcup_{i=1}^{\infty} B_i \subset \bigcup_{i=1}^{\infty} A_i .$$

Now let $x \in \bigcup_{i=1}^{\infty} A_i$. Then , x must be in at least one of the sets A_i's. Let n be the least value of i such that $x \in A_i$. Then $x \in B_n$, and so $x \in \bigcup_{n=1}^{\infty} B_n$.

Hence

$$\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i=1}^{\infty} B_i .$$

Hence the theorem.

Theorem. A countable union of measurable sets is a measurable set.

Proof . Let $\{E_i\}$ be a sequence of measurable sets and let

$E = \bigcup_{i=1}^{\infty} E_i$. To prove E to be a measurable set, we may assume, without any loss of generality, that the sets E_i are mutually disjoint.

For each $n \in N$, define $F_n = \bigcup_{i=1}^{n} E_i$. Since *M* is an algebra of sets and $E_1, E_2, \ldots E_n$ are in *M*, the sets F_n are measurable. Therefore, for any set A, we have

$$m^*(A) = m^*(A \cap F_n) + m^*(A \cap F_n^c)$$

$$\geq m^*(A \cap F_n) + m^*(A \cap E^c),$$

since

$$F_n^c = \left[\bigcup_{i=n+1}^{\infty} E_i\right] \cup E^c \supset E^c.$$

But we observe that

$$m^*(A \cap F_n) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

Therefore,

$$m^*(A) \geq \sum_{i=1}^{n} m^*(A \cap E_i) + m^*(A \cap E^c).$$

This inequality holds for every $n \in N$ and since the left-hand side is independent of n, letting $n \to \infty$, we obtain

$$m^*(A) \geq \sum_{i=1}^{\infty} m^*(A \cap E_i) + m^*(A \cap E^c)$$

$$\geq m^*(A \cap E) + m^*(A \cap E^c),$$

in view of the countable subadditivity of m*. Hence E is a measurable set.

6.1 Theorem. Let E be a measurable set. Then any translate E + y is measurable, where y is a real number. Furthermore,

$$m(E + y) = m(E).$$

Proof. Let A be any set. Since E is measurable, we have

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$$

$$\Rightarrow \qquad m^*(A + y) = m^*([A \cap E] + y) + m^*([A \cap E^c] + y),$$

in view of m* is invariant under translation. It can be verified that

$$\begin{cases} [A \cap E] + y = (A + y) \cap (E + y) \\ [A \cap E^c] + y = (A + y) \cap (E^c + y). \end{cases}$$

Hence

$$m^*(A + y) = m^*([A + y] \cap [E + y]) + m^*([A + y] \cap [E^c + y]).$$

Since A is arbitrary, replacing A with A – y, we obtain

$$m^*(A) = m^*(A \cap E + y) + m^*(A \cap E^c + y).$$

Now since m* is translation invariant, the measurability of E + y follows by taking into account that $(E + y)^c = E^c + y$.

Theorem. Let $\{E_i\}$ be an infinite decreasing sequence of measurable sets; that is, a sequence with $E_{i+1} \subset E_i$ for each $i \in N$. Let $m(E_1) < \infty$. Then

$$m\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n \to \infty} m(E_n).$$

Proof. Let $m(E_1) < \infty$.

Set $E = \bigcap_{i=1}^{\infty} E_i$ and $F_i = E_i - E_{i+1}$. Then the sets F_i are measurable and pairwise disjoint, and

$$E_1 - E = \bigcup_{i=1}^{\infty} F_i .$$

Therefore,

$$m(E_1 - E) = \sum_{i=1}^{\infty} m(F_i) = \sum_{i=1}^{\infty} m(E_i - E_{i+1}) .$$

But $m(E_1) = m(E) + m(E_1 - E)$ and

$m(E_i) = m(E_{i+1}) + m(E_i - E_{i+1})$, for all $i \geq 1$, since $E \subset E_1$ and $E_{i+1} \subset E_i$.

Further, using the fact that $m(E_i) < \infty$, for all $i \geq 1$, it follows that

$$m(E_1 - E) = m(E_1) - m(E)$$

and $$m(E_1 - E_{i+1}) = m(E_i) - m(E_{i+1}),\ \forall\ i \geq 1.$$

Hence,

$$m(E_1) - m(E) = \sum_{i=1}^{\infty} (m(E_i) - m(E_{i+1}))$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} (m(E_i) - m(E_{i+1})$$

$$= \lim_{n\to\infty} \{m(E_1) - m(E_n)\}$$

$$= (E_1) - \lim_{n\to\infty} m(E_n).$$

Since $m(E_1) < \infty$, it gives

$$m(E) = \lim_{n\to\infty} m(E_n).$$

Remark. The condition $m(E_1) < \infty$, in above Theorem cannot be relaxed.

Example. Consider the sets E_n given by $E_n =]n, \infty[$, $n \in N$. Then $\{E_n\}$ is a decreasing sequence of measurable sets such that $m(E_n) = \infty$ for each $n \in N$ and $\bigcap_{n=1}^{\infty} E_n = \phi$. Therefore,

$$\lim_{n\to\infty} m(E_n) = \infty, \quad \text{while} \quad m(\phi) = 0.$$

Theorem. Let $\{E_i\}$ be an infinite increasing sequence of measurable sets i.e. $E_{i+1} \subset E_i$ for each $i \in N$. Then

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} m(E_n).$$

Proof. If $m(E_i) = \infty$ for some $n \in N$, then the result is trivial, since

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) \geq m(E_i) = \infty,$$

and $m(E_n) = \infty$, for each $n \geq i$. Let $m(E_i) < \infty$, for each $i \in N$. Set

$$E = \bigcup_{i=1}^{\infty} E_i, \quad F_i = E_{i+1} - E_i.$$

Then the sets F_i are measurable and pairwise disjoint, and

$$E - E_1 = \bigcup_{i=1}^{\infty} F_i$$

$$\Rightarrow \qquad m(E-E_1) = m\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} m(F_i) = \sum_{i=1}^{\infty} m(E_{i+1} - E_i)$$

$$\Rightarrow \qquad m(E) - m(E_1) = \lim_{n\to\infty} \sum_{i=1}^{n} \{m(E_{i+1}) - m(E_i)\}$$

$$= \lim_{n\to\infty} \{m(E_{n+1}) - m(E_1)\}$$

$$\Rightarrow \qquad m(E) = \lim_{n\to\infty} m(E_n).$$

Definition. A set which is a countable (finite or infinite) union of closed sets is called an F_σ**-set**.

Example of F_σ-set are:A closed set, A countable set, A countable union of F_σ-sets, an open interval (a, b) since

$$(a, b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b - \frac{1}{n}\right],$$

and hence an open set.

Definition. A set which is a countable intersection of open sets is a G_δ-**set**.

Examples of G_δ-set are :An open set and, in particular, an open interval, A closed set, A countable intersection of G_δ-sets.

A closed interval [a, b] since

$$[a, b] = \bigcap_{n=1}^{\infty} \left] a - \frac{1}{n}, b + \frac{1}{n} \right[.$$

Remark. Each of the classes F_σ and G_δ of sets is wider than the classes of open and closed sets. The complement of an F_σ-set is a G_δ-set, and conversely.

Theorem. Let A be any set. Then :

(a) Given $\in > 0$, $\exists$ an open set $O \supset A$ such that

$$m^*(O) \le m^*(O) + \in$$

while the inequality is strict in case $m^*(A) < \infty$; and hence $m^*(A) = \inf_{A \subset O} m^*(O)$,

(b) $\exists$ a G_δ-set $G \supset A$ such that

$$m^*(A) = m^*(G).$$

Proof. (a) Assume first that $m^*(A) < \infty$. Then there exists a countable collection $\{I_n\}$ of open intervals such that $A \subset \bigcup_n I_n$ and

$$\sum_n l(I_n) < m^*(A) + \in$$

Set $O = \bigcup_{n=1}^{\infty} I_n$. Clearly O is an open set and

$$m^*(O) = m^*\left(\bigcup_n I_n\right)$$

$$\le \sum_n m^*(I_n) = \sum_n l(I_n) < m^*(A) + \in.$$

(b) Choose $\in = \frac{1}{n}$, $n \in N$ in (a). Then, for each $n \in N$, $\exists$ an open set $O_n \supset E$ such that

$$m^*(O_n) \le m^*(A) + \frac{1}{n}.$$

Define G = $\bigcap_{n=1}^{\infty} O_n$. Clearly, G is a G_δ-set and G ⊃ A. Moreover, we observe that

$$m^*(A) \leq m^*(G) \leq m^*(O_n) \leq m^*(A) + \frac{1}{n}, \qquad n \in N.$$

Letting $n \to \infty$ we have $m^*(G) = m^*(A)$.

Theorem. Let E be a given set. Then the following statements are equivalent:

(a) E is measurable.

(b) Given $\in > 0$, there is an open set $O \supset E$ such that $m^*(O - E) < \in$,

(c) There is a G_δ-set $G \supset E$ such that $m^*(G - E) = 0$.

(d) Given $\in > 0$, there is a closed set $F \subset E$ such that $m^*(E - F) < \in$,

(e) There is a F_σ-set $F \subset E$ such that $m^*(E - F) = 0$.

Proof. (a) ⇒ (b) : Suppose first that $m(E) < \infty$, then there is an open set $O \supset E$ such that

$$m^*(O) < m^*(E) + \in.$$

Since both the sets O and E are measurable, we have

$$m^*(O - E) = m^*(O) - m^*(E) < \in.$$

Now let $m(E) = \infty$. Write

$R = \bigcup_{n=1}^{\infty} I_n$ where R is setof real numbers and I_n are disjoint finite intervals Then, if $E_n = E \cap I_n$, $m(E_n) < \infty$. We can 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

$$m^*(O - E) \leq \sum_{n=1}^{\infty} m^*(O_n - E_n) < \in.$$

(b) $\Rightarrow$ (c) : Given $\in = 1/n$, there is an open set $O_n \supset E$ with $m^*(O_n - E) < 1/n$. Define $G = \bigcap_{n=1}^{\infty} O_n$. Then G is a G_δ-set such that $G \supset E$ and

$$m^*(G - E) \le m^*(O_n - E) < \frac{1}{n}, \quad \forall\, n \in N.$$

This on letting $n \to \infty$ proves (c).

(c) $\Rightarrow$ (a) : Write $E = G - (G - E)$. But the sets G and G–E are measurable since G is a Borel set and G–E is of outer measure zero. Hence E is measurable.

(a) $\Rightarrow$ (d) : E^c is measurable and so, in view of (b), there is an open set $O \supset E^c$ such that $m^*(O - E^c) < \in$. But $O - E^c = E - O^c$. Taking $F = O^c$, the assertion (d) follows.

(d) $\Rightarrow$ (e) : Given $\in = 1/n$, there is a closed set $F_n \subset E$ with $m^*(E - F_n) < 1/n$. Define $F = \bigcup_{n=1}^{\infty} F_n$. Then F is a F_σ-set such that $F \subset E$ and

$$m^*(E - F) \le m^*(E - F_n) < \frac{1}{n}, \quad \forall\, n \in N.$$

Hence the result in (e) follows on letting $n \to \infty$.

(e) $\Rightarrow$ (a) : The proof is similar to (c) $\Rightarrow$ (a).

Theorem. Let E be a set with $m^*(E) < \infty$. Then E is measurable if and only if, given $\in > 0$, there $\in > 0$, there is a finite union B of open intervals such that

$$m^*(E \,\Delta\, B) < \in.$$

Proof. Let E be measurable, and let $\in > 0$ be given. Then there exists an open set $O \supset E$ with $m^*(O - E) < \in/2$. As $m^*(E)$ is finite, so is $m^*(O)$. Further, since the open set O can be expressed as the union of disjoint countable open intervals $\{I_i\}$, there exists an $n \in N$ such that

$$\sum_{i=n+1}^{\infty} l(I_i) < \frac{\in}{2},$$

since $m^*(O) < \infty$.

Write $B = \bigcup_{i=1}^{n} I_i$. Then

$$E\Delta B = (E - B) \cup (B - E) \subset (O - B) \cup (O - E).$$

Hence

$$m^*(E \,\Delta\, B) \le m^*\left(\bigcup_{i=n+1}^{\infty} I_i\right) + m^*(O - E) < \in.$$

Conversely, assume that for a given $\in > 0$, there is a finite union, $B = \bigcup_{i=1}^{n} I_i$, of open intervals with $m^*(E\Delta B) < \in$. Then there is an open set $O \supset E$ such that

$$m^*(O) < m^*(E) + \in. \qquad \ldots(1)$$

If we can show that $m^*(O - E)$ is arbitrarily small, it follows that E is a measurable set.

Write $S = \bigcup_{i=1}^{n} (I_i \cap O)$. Then $S \subset B$ and so

$$S\Delta E = (E - S) \cup (S - E) \subset (E - S) \cup (B - E).$$

However,

$$E - S = (E \cap O^c) \cup (E \cap B^c) = E - B$$

Since $E \subset O$. 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

$$m^*(E) < m^*(S) + \in. \qquad \ldots(2)$$

Also $O - E \subset (O - S) \cup (S\Delta E)$ gives

$$m^*(O - E) < m^*(O) - m^*(S) + \in.$$

Hence, in view of (1) and (2), we get

$$m^*(O - E) < 3\in.$$

Definition. An algebra A of sets is called a σ-algebra (or σ-Boolean algebra or Borel field) if it is closed under countable union of sets; that is, $\bigcup_{i=1}^{\infty} A_i$, is in A whenever the countable collection $\{A_i\}$ of sets, is in A.

Note. It follows, from DeMorgan's law, that a σ-algebra is also closed under countable intersection of sets. The family M of all measurable sets (subsets of R) is a σ-algebra of sets in P(R).

Theorem. Let $\{E_i\}$ be an infinite sequence of disjoint measurable sets. Then

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m(E_i).$$

Proof. For each $n \in N$, we have

$$m\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} m(E_i).$$

But

$$\bigcup_{i=1}^{\infty} E_i \supset \bigcup_{i=1}^{n} E_i, \quad \forall\, n \in N.$$

Therefore, we obtain

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) \geq \sum_{i=1}^{n} m(E_i).$$

Since the left-hand side is independent of n, letting $n \to \infty$, we get

$$m\left(\bigcup_{i=1}^{\infty} E_i\right) \geq \sum_{i=1}^{\infty} m(E_i).$$

The reverse inequality is countable sub-additivity property of m*.

Definition. The σ-algebra generated by the family of all open sets in R, denoted *B*, is called the class of **Borel sets** in R. The sets in *B* are called Borel sets in R.

Examples: Each of the open sets, closed sets, G_δ-sets, F_σ-sets, $G_{\delta\sigma}$-sets, $F_{\sigma\delta}$-sets,… is a simple type of Borel set.

Theorem. Every Borel set in **R** is measurable; that is, $B \subset M$.

Proof. We prove the theorem in several steps by using the fact that M is a σ-algebra.

Step 1 : The interval (a, ∞) is measurable.

It is enough to show, for any set A, that

$$m^*(A) \geq m^*(A_1) + m^*(A_2),$$

where $A_1 = A \cap (a, \infty)$ and $A_2 = A \cap (-\infty, a)$.

If $m^*(A) = \infty$, our assertion is trivially true. Let $m^*(A) < \infty$. Then, for each $\in > 0$, $\exists$ a countable collection $\{I_n\}$ of open intervals that covers A and satisfies

$$\sum_n l(I_n) < m^*(A) + \in$$

Write $I'_n = I_n \cap (a, \infty)$ and $I''_n = I_n \cap (-\infty, a)$. Then,

$$\begin{aligned} I'_n \cup I''_n &= \{I_n \cap (a, \infty)\} \cup \{I_n \cap (-\infty, a)\} \\ &= I_n \cap (-\infty, \infty) \\ &= I_n, \end{aligned}$$

and $I'_n \cap I''_n = \phi$. Therefore,

$$\begin{aligned} l(I_n) &= l(I'_n) + l(I''_n) \\ &= m^*(I'_n) + m^*(I''_n) \end{aligned}$$

But

$$A_1 \subset [\cup I_n] \cap (a, \infty) = \cup (I_n \cap (a, \infty)) = \cup I'_n,$$

so that $m^*(A_1) \leq m^*\left(\bigcup_n I'_n\right) \leq \sum_n m^*(I'_n)$. Similarly $A_2 \subset \bigcup_n I''_n$ and so $m^*(A_2) \leq \sum_n m^*(I''_n)$. Hence,

$$\begin{aligned} m^*(A_1) + m^*(A_2) &\leq \sum_n \{m^*(I'_n) + m^*(I''_n)\} \\ &= \sum_n l(I_n) < m^*(A) + \in. \end{aligned}$$

Since $\in > 0$ is arbitrary, this verifies the result.

Step 2 : The interval $]-\infty, a]$ is measurable since

$$]-\infty, a] = [a, \infty[^c$$

Step 3: The interval $]-\infty, b[$ is measurable since it can be expressed as a countable union of the intervals of the form as in Step 2; that is,

$$]-\infty, b[\; = \bigcup_{n=1}^{\infty}]-\infty, b-\frac{1}{n}].$$

Step 4 : Since any open interval]a, b[can be expressed as

$$]a, b[\; = \;]-\infty, b[\cap]a, \infty[,$$

it is measurable.

Step 5 : Every open set is measurable. It is so because it can be expressed as a countable union of open intervals (disjoint).

Hence, in view of Step 5, the σ-algebra M contains all the open sets in R. Since B is the smallest σ-algebra containing all the open sets, we conclude that B⊂M. This completes the proof of the theorem.

5.4 Corollary. Each of the sets in R : an open set, a closed set, an F_σ-set and a G_δ-set is measurable.

Problem 13. Prove that every interval is a measurable set and its measure is its length.

Solution. It follows in view of the fact that an interval is a Borel set and the outer measure of an interval is its length.

8. NONMEASURABLE SETS

8.1 Definition. If x and y are real numbers in [0, 1[, then the sum modulo 1, $\overset{o}{+}$, of x and y is defined by

$$x \overset{o}{+} y = \begin{cases} x+y, & x+y<1 \\ x+y-1, & x+y\geq 1 \end{cases}$$

8.2 Definition. If E is a subset of [0, 1[, then the translate modulo 1 of E by y is defined to be the set given by

$$E \overset{o}{+} y = \{z : \; = x \overset{o}{+} y, x\in E\}.$$

Note that

(i) $x, y \in [0, 1[\; \Rightarrow x \overset{o}{+} y \in [0, 1[.$

(ii) The operation + is commutative and associative.

(iii)

Now weprove that the measure (Lebesgue) is invariant under translate modulo 1.

8.3 Theorem. Let $E \subset [0, l[$ be a measurable set and $y \in [0, 1[$ be given.

Then the set E + y is measurable and m(E + y) = m(E).

Proof. Define

$$\begin{cases} E_1 = E \cap [0,1-y[\\ E_2 = E \cap [1-y,1[. \end{cases}$$

Clearly E_1 and E_2 are two disjoint measurable sets such that$E_1 \cup E_2 = E$. Therefore

$$m(E) = m(E_1) + m(E_2).$$

Now, $E_1 + y = E_1 + y$ and $E_2 + y = E_2 + y - 1$ and so $E_1 + y$ and $E_2 + y$ are disjoint measurable sets with

$$\begin{cases} m(E_1 \overset{o}{+} y) = m(E_1 + y) = m(E_1) \\ m(E_2 + y) = m(E_2 + y - 1) = m(E_2), \end{cases}$$

since m is translation invariant (cf. Theorem 6.1). Also

$$\begin{aligned} E \overset{o}{+} y &= (E_1 \cup E_2) + y \\ &= (E_1 \overset{o}{+} y) \cup (E_2 \overset{o}{+} y) \end{aligned}$$

Hence $E \overset{o}{+} y$ is a measurable set with

$$\begin{aligned} m(E + y) &= m(E_1 + y) + m(E_2 + y) \\ &= m(E_1) + m(E_2) \\ &= m(E). \end{aligned}$$

8.4 Theorem. There exists a non-measurable set in the interval [0, 1[.

Proof. We define an equivalence relation '~' in the set I = [0, 1[by saying that x and y in I are equivalent, to be written x~y, if x – y is rational. Clearly, the relation ~ partitions the set I into mutually disjoint equivalence classes, that is, any two elements of the same class differ by a rational number while those of the different classes differ by an irrational number.

Construct a set P by choosing exactly one element from each equivalence class – this is possible by the axiom of choice. Clearly $P \subset [0, 1[$. We shall now show that P is a nonmeasurable set.

Let $\{r_i\}$ be an enumeration of rational numbers in [0, 1[with $r_0 = 0$. Define

$$P_i = P + r_i.$$

Then $P_0 = P$. We further observe that:

(a) $P_m \cap P_n = \phi$, $m \neq n$.

(b) $\bigcup_n P_n = [0, 1[$.

Proof of (a). Let if possible, $y \in P_m \cap P_n$. Then there exist p_m and p_n in P such that

$$y = p_m + r_m = p_n + r_n$$

$\Rightarrow$ $p_m - p_n$ is a rational number

$\Rightarrow$ $p_m - p_n$, by the definition of the set P

$\Rightarrow$ $m = n$.

This is a contradiction.

Proof of (b). Let $x \in [0, 1[$. Then x lies in one of the equivalence classes and as such x is equivalent to an element y (say) of P. Suppose r_i is the rational number by which x differs from y. Then $x \in P_i$ and hence $[0, 1[\subset \bigcup_n P_n$ while the reverse inclusion is obviously true.

Now, we turn towards the proof of the nonmeasurability of P. Assume that P is measurable. Then each P_i is measurable, and $m(P_i) = m(P)$, cf. Theorem 8.3. Therefore

$$m(\bigcup_i P_i) = \sum_{i=0}^{\infty} m(P_i)$$

$$= \sum_{i=0}^{\infty} m(P)$$

$$= \begin{cases} 0 & \text{if } m(P) = 0 \\ \infty & \text{if } m(P) > 0. \end{cases}$$

On the other hand

$$M(\bigcup_i P_i) = m([0, 1[) = 1.$$

These lead to contradictory statements. Hence P is a nonmeasurable set.