Institute of Distance and Open Learning (IDOL)
University of Mumbai.

M.Sc. (Mathematics), SEM- I

Paper - II

ANALYSIS – I

PSMT102

Note- There will be some addition to this study material. You should download it again after few weeks.

CONTENT

***

SYLLABUS

Unit I. Euclidean space $\mathbb{R}^{\boldsymbol{n}}$

Euclidean space $\mathbb{R}^{\boldsymbol{n}}$: inner product $< x, y > = \sum_{j=1}^{n} x_j y_j \; of \; x = (x_1, .., x_n), y = (y_1, .., y_n) \in \mathbb{R}^n$ and properties, norm

$||x|| = \sqrt{\sum_{j=1}^{n} x_j^2} \; of \; x = (x_1, .., x_n) \in \mathbb{R}^n$, Cauchy-Schwarz inequality, properties of the norm function $||x|| \; on \; \mathbb{R}^n$.

(Ref. W. Rudin or M. Spivak).

Standard topology on $\mathbb{R}^{\boldsymbol{n}}$: open subsets of $\mathbb{R}^n$, closed subsets of $\mathbb{R}^n$, interior A^o and boundary ∂A of a subset A of $\mathbb{R}^n$.

(ref. M. Spivak)

Operator norm $\left||T|\right|$ of a linear transformation $T: \mathbb{R} \to \mathbb{R}^m$

($||T|| = sup\{||T(v)||$: v ∈ $\mathbb{R}$ &$||v|| \leq 1\}$) and its properties such as:

For all linear maps $S, T: \mathbb{R} \to \mathbb{R}^m$ and $R : \mathbb{R}^m \to \mathbb{R}^k$

1. $||S + T|| \leq ||S|| + ||T||$,

2. $||RoS|| \leq ||R||||S||$,and

3. $||cT|| = |c|||T||(c \in \mathbb{R})$.

(Ref. C. C. Pugh or A. Browder)

Compactness: Open cover of a subset of $\mathbb{R}^n$, Compact subsets of $\mathbb{R}^n$ (A subset K of $\mathbb{R}^n$ is compact if every open cover of K contains a finite subover), Heine-Borel theorem (statement only), the Cartesian product of two compact subsets of $\mathbb{R}^n$ compact (statement only), every closed and bounded subset of $\mathbb{R}^n$ is compact.

Bolzano-Weierstrass theorem: Any bounded sequence in $\mathbb{R}^n$ has a converging subsequence.

Brief review of following three topics:

1. Functions and Continuity Notation: $A \subset \mathbb{R}^n$ arbitary non-empty set. A function f : A → $\mathbb{R}^m$ and its component functions, continuity of a function(∈, δ definition). A function f : A → $\mathbb{R}^m$ is continuous if and only if for every open subset V ⊂ $\mathbb{R}^m$ there is an open subset U of $\mathbb{R}^n$ such that $f^{-1}(V) = A \cap U$.

2. Continuity and compactness: Let K ⊂ $\mathbb{R}^n$ be a compact subset and f : K → $\mathbb{R}^m$ be any continuous function. Then f is uniformly continuous, and f(K) is a compact subset of $\mathbb{R}^m$.

3. Continuity and connectedness: Connected subsets of $\mathbb{R}$ are intervals. If $f : E \to \mathbb{R}$ is continuous where $E \subset \mathbb{R}^n$ and E is connected, then f(E) ⊂$\mathbb{R}$ is connected.

Unit II. Differentiable functions

Differentiable functions on $\mathbb{R}^n$, the total derivative $(Df)_p$ of a differentiable function $f : U \to \mathbb{R}^m$ at $p \in U$ where U is open in $\mathbb{R}^n$, uniqueness of total derivative, differentiability implies continuity.

(ref:[1] C.C.Pugh or[2] A.Browder)

Chain rule. Applications of chain rule such as:

1. Let γ be a differentiable curve in an open subset U of $\mathbb{R}^n$. Let $f : U \rightarrow \mathbb{R}$ be a differentiable function and let $g(t) = f(\gamma(t))$. Then

$g'(t) =< (\nabla f)(\gamma(t)), \gamma'(t) >$.

2. Computation of total derivatives of real valued functions such as

(a) the determinant function $det(X), X \in M_n(\mathbb{R})$,

(b) the Euclidean inner product function $< x, y >, (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$.

(ref. M. Spivak, W. Rudin)

Results on total derivative:

1. If $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a constant function,then $(Df)_p = 0 \ \forall p \in \mathbb{R}^n$.

2. If $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map, then $(Df)_p = f \ \forall p \in \mathbb{R}^n$.

3. A function $= (f_1, f_2, .. f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable at $p \in \mathbb{R}^n$ if and only if each f_j is differentiable at $p \in \mathbb{R}^n$, and $(Df)p = ((Df_1)p, (Df_2)\, p, .., (Df_m)p)$.

(ref. M. Spivak).

Partial derivatives, directional derivative $(D_u f)(p)$ of a function f at p in the direction of the unit vector, Jacobian matrix, Jacobian determinant. Results such as :

1. If the total derivative of a map $f = (f_1, .., f_m) : U \rightarrow \mathbb{R}^m$($U$ open subset of $\mathbb{R}^n$) exists at $p \in U$, then all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exists at p.

2. If all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ of a map $f = (f_1, .., f_m) : U \rightarrow \mathbb{R}^m$($U$ open subset of $\mathbb{R}^n$) exist and are continuous on U, then f is differentiable.

(ref. W. Rudin)

Derivatives of higher order, C^k-functions, C^∞-functions.(ref. T. Apostol)

Unit III. Inverse function theorem and Implicit function theorem

Theorem (Mean Value Inequality): Suppose $f : U \to \mathbb{R}^m$ is differentiable on an open subset U of $\mathbb{R}^n$ and there is a real number such that $||(Df)_x|| \leq M \; \forall x \in U$. If the segment $[p, q]$ is contained in U, then $||f(q) - f(p)|| \leq M||q - p||$.

(ref. C. C. Pugh or A. Browder).

Mean Value Theorem: Let $f : U \to \mathbb{R}^m$is a differentiable on an open subset U of $\mathbb{R}^n$. Let $p, q \in U$ such that the segment $[p, q]$ is contained in U. Then for every vector $v \in \mathbb{R}^n$ there is a point $x \in [p, q]$ such that $<v, f(q) - f(p)> = <v, (Df)_x(q - p)>$. (ref:T. Apostol)

If $f : U \to \mathbb{R}^m$ is differentiable on a connected open subset U of $\mathbb{R}^n$ and $(Df)_x = 0 \; \forall x \in U$, then f is a constant map.

Taylor expansion for a real valued C^m-function defined on an open subset of $\mathbb{R}^n$, stationary points(critical points), maxima, minima, saddle points, second derivative test for extrema at a stationary point of a real valued C^2 -function defined on an open subset of $\mathbb{R}^n$. Lagrange's method of undetermined multipliers. (ref. T. Apostol)

Contraction mapping theorem. Inverse function theorem, Implicit function theorem.(ref. A. Browder)

Unit IV. Riemann Integration(15 Lectures)

Riemann Integration over a rectangle in $\mathbb{R}^n$, Riemann Integrable functions, Continuous functions are Riemann integrable, Measure zero sets, Lebesgues Theorem(statement only), Fubini's Theorem and applications.

(Reference for Unit IV: M. Spivak, Calculus on Manifolds)

DIFFERENTIATION OF FUNCTIONS OF SEVERAL VARIABLES

Unit Structure

5.0 OBJECTIVES

After reading this unit you should be able to

- define a differentiable function of several variables
- define and calculate the partial and directional derivatives (if they exist) of a function of several variables
- establish the connection between the total, partial and directional derivatives of a differentiable function at a point

5.1 INTRODUCTION

You have seen how to extend the concepts of limit and continuity to functions between metric spaces. Another important concept is differentiation. If we try to apply this to functions between metric spaces, we encounter a problem. We realise that apart from the distance notion, the domain and codomain also need to have an algebraic structure. So, let us consider Euclidean spaces like $\boldsymbol{R}^n$, which have which have both metric and algebraic structures. Functions between two Euclidean spaces are what we call functions of several variables.

In this chapter we shall introduce the concept of differentiability of a function of several variables. The extension of this concept from one to several variables was not easy. Many different approaches were tried before this final one was accepted. The definition may seem a little difficult in the beginning, but as you will see, it allows us to extend all our knowledge of derivatives of functions one variable to the several variables case. You may have studied these concepts in T. Y. So, here we shall try to go a little deeper into these concepts, and deal with vector functions of several variables.

5.2 TOTAL DERIVATIVE

To arrive at a suitable definition of differentiability of functions of several variables, mathematicians had to closely examine the concept of derivative of a function of a single variable. To decide on the approach to extension of the concept, it was important to know what was the essence and role of a derivative. So, let us recall the definition of the derivative of a function f: $\boldsymbol{R} \to \boldsymbol{R}$.

We say that f is differentiable at $a \in \boldsymbol{R}$, if the limit, $\lim\limits_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$ exists.

In that case, we say that the derivative of f at a, $f^1(a) = \lim\limits_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$ **.......(**5.1)

So, we take the limit of the ratio of the increment in $f(x)$ to the increment in x. Now, when our function is defined on $\boldsymbol{R^n}$, the increment in the independent variable will be a vector. Since division by a vector is not defined, we cannot write a ratio similar to the one in (5.1). But (5.1) can be rewritten as

$$\lim_{h\to 0} \left[\frac{f(a+h)-f(a)}{h} - f^1(a) \right] = 0, \text{ or}$$

$$\lim_{h\to 0} \left[\frac{f(a+h)-f(a)-f^1(a).h}{h} \right] = 0, \text{ or}$$

$$\lim_{h\to 0} \frac{r(h)}{h} = 0, \text{ where } r(h) = f(a+h) - f(a) - f^1(a).h.$$

So, we can write $f(a+h) = f(a) + f^1(a).h + r(h)$, ...(5.2)

where the "remainder" r(h) is so small, that $\dfrac{r(h)}{h}$ tends to zero as h tends to zero.

For a fixed a, $f(a)$, and $f^1(a)$ are fixed real numbers. This means, except for the remainder, $r(h)$, (5.2) expresses $f(a + h)$ as a linear function of h. This also helps us in "linearizing" f. We say that for points close to a, the graph of the function f can be approximated by a line. Thus, $f^1(a)$ gives rise to a linear function L from $\boldsymbol{R}$ to $\boldsymbol{R}$.

L: $\boldsymbol{R} \to \boldsymbol{R}$, $h \to f^1(a).h$, which helps us in linearizing the given function f near the given

point a. (5.2) then transforms to

$f(a + h) = f(a) + \mathrm{L}(h) + r(h)$ . ...(5.3)

It is this idea of linearization that we are now going to extend to a function of several variables.

Definition 5.1 Suppose E is an open set in $\boldsymbol{R^n}$, $\boldsymbol{f}$: E $\to \boldsymbol{R^m}$, and $\boldsymbol{a} \in$ E. We say that **f** is differentiable at $\boldsymbol{a}$, if there exists a linear transformation T : $\boldsymbol{R^n} \to \boldsymbol{R^m}$, such that

$$\lim_{h\to 0} \frac{\|\boldsymbol{f(a+h)-f(a)-T(h)}\|}{\|\boldsymbol{h}\|} = \mathbf{0} \qquad ....(5.4)$$

and we write $\boldsymbol{f}^1(\boldsymbol{a}) = \mathbf{T}$.

If $\boldsymbol{f}$ is differentiable at every point in E, we say that $\boldsymbol{f}$ is differentiable in E.

Remark 5.1 i) Bold letters indicate vectors.

ii) Since E is open, $\exists\, r > 0$, such that B($\boldsymbol{a}$, r) $\subset$ E. We choose $\boldsymbol{h}$, such that $\|\boldsymbol{h}\| < r$, so that

$\boldsymbol{a} + \boldsymbol{h} \in$ E.

iii) The norm in the numerator of (5.4) is the norm in $\boldsymbol{R^m}$, whereas the one in the denominator is the norm in $\boldsymbol{R^n}$.

iv) The linear transformation **T** depends on the point $\boldsymbol{a}$. So, when we have to deal with more than one point, we use the notation, $\mathbf{T}_{\boldsymbol{a}}$, $\mathbf{T}_{\boldsymbol{b}}$, and so on.

We have seen that in the one variable case, the derivative defines a linear function,

$h \to f^1(a).h$ from **R** to $\boldsymbol{R}$. Similarly, here the derivative is a linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$. With every such transformation, we have an associated m × n matrix. The jth column of this matrix is $\mathbf{T}(\boldsymbol{e_j})$, where $\boldsymbol{e_j}$ is a basis vector in the standard basis of $\boldsymbol{R^n}$.

For a given point $\boldsymbol{a}$, the linear transformation $\mathbf{T}_{\boldsymbol{a}}$ is called the total derivative of $\boldsymbol{f}$ at $\boldsymbol{a}$, and is denoted by $\boldsymbol{f^1}(\boldsymbol{a})$ or D$\boldsymbol{f}(\boldsymbol{a})$. We can then write

$$\boldsymbol{f(a + h) = f(a) + T_a(h) + r(h)}, \textit{ where } \frac{\boldsymbol{r(h)}}{\|\boldsymbol{h}\|} \to \mathbf{0}, \textit{ as } \boldsymbol{h} \to \mathbf{0}\ . \qquad \text{.............................}(5.5)$$

We now give a few examples.

Example 5.1 : Consider $\boldsymbol{f}: \boldsymbol{R^n} \to \boldsymbol{R^n}$, $\boldsymbol{f(x) = a + x}$, where $\boldsymbol{a}$ is a fixed vector in $\boldsymbol{R^n}$. Find the total derivative of $\boldsymbol{f}$ at a point $\boldsymbol{p} \in \boldsymbol{R^n}$, if it exists.

Solution : Now, $\boldsymbol{f(p + h) - f(p) = h}$. So, if we take **T** to be the identity transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^n}$, then we get

$\boldsymbol{f(p + h) - f(p) - T(h)} = \mathbf{0}$, and hence

$$\lim_{h\to 0} \frac{\|\boldsymbol{f(p+h) - f(p) - T(h)}\|}{\|\boldsymbol{h}\|} = \mathbf{0}.$$

Comparing this with 5.5, we conclude that the identity transformation is the total derivative of $\boldsymbol{f}$ at the point $\boldsymbol{p}$.

Example 5.2 : Find the total derivative, if it exists, for $\boldsymbol{f} : \boldsymbol{R^2} \to \boldsymbol{R^2}$, $\boldsymbol{f}(x, y) = (x^2, y^2)$, at a point

$\boldsymbol{a} = (a_1, a_2)$.

Solution : If $\boldsymbol{f}$ is differentiable, we expect $\mathbf{T}_{\boldsymbol{a}}$ to be a 2×2 matrix. Let $\boldsymbol{h} = (h_1, h_2)$. Now,

$$\begin{aligned}\boldsymbol{f(a + h) - f(a)} &= ((a_1 + h_1)^2, (a_2 + h_2)^2) - (a_1^2, a_2^2)\\ &= (2a_1h_1 + h_1^2,\ 2a_2h_2 + h_2^2)\\ &= (2a_1h_1, 2a_2h_2) + (h_1^2, h_2^2)\end{aligned}$$

$$=\begin{pmatrix} 2a_1 & 0 \\ 0 & 2a_2 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + (h_1^2, h_2^2)$$

We take $\mathbf{T}_{\boldsymbol{a}} = \begin{pmatrix} 2a_1 & 0 \\ 0 & 2a_2 \end{pmatrix}$, and $\boldsymbol{r}(\boldsymbol{h}) = (h_1^2, h_2^2)$, and write

$\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{a}) + \boldsymbol{T}_{\boldsymbol{a}}(\boldsymbol{h}) + \boldsymbol{r}(\boldsymbol{h})$, where $\frac{\boldsymbol{r}(\boldsymbol{h})}{\|\boldsymbol{h}\|} = \frac{(h_1^2,\ h_2^2)}{\sqrt{(h_1^2+h_2^2)}} \to \mathbf{0}$, as $\boldsymbol{h} \to \mathbf{0}$.

Thus $\mathbf{T}_{\boldsymbol{a}}$ is the total derivative of $\boldsymbol{f}$ at $\boldsymbol{a}$.

Now that we have defined the total derivative, let us see how many of the results that we know about derivatives of functions of a single variable, hold for these total derivatives.

Theorem 5.1: If $\boldsymbol{f} : \boldsymbol{R}^n \to \boldsymbol{R}^m$ is differentiable at $\boldsymbol{a} \in \boldsymbol{R}^n$, then its total derivative is unique.

Proof : Suppose $\boldsymbol{f}$ has two derivatives, $\mathbf{T}_1$ and $\mathbf{T}_2$ at $\boldsymbol{a}$, and let $\mathbf{T} = \mathbf{T}_1 - \mathbf{T}_2$. Let $\boldsymbol{h} \in \boldsymbol{R}^n$,

$\boldsymbol{h} \neq \mathbf{0}$, and $t \in \boldsymbol{R}$, such that $t \to 0$.

Then $t\boldsymbol{h} \to \boldsymbol{0}$ as $t \to 0$.

Since $\mathbf{T}_1$ is a total derivative of $\boldsymbol{f}$ at $\boldsymbol{a}$,

$$\lim_{t\to 0} \frac{\|\boldsymbol{r}_1(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} = \lim_{t\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_1(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} = 0 \qquad \text{......................(5.6)}$$

Since $\mathbf{T}_2$ is also a total derivative of $\boldsymbol{f}$ at $\mathbf{a}$,

$$\lim_{t\to 0} \frac{\|\boldsymbol{r}_2(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} = \lim_{t\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_2(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} = 0 \qquad \text{......................(5.7)}$$

Thus, $\| \boldsymbol{T}(t\boldsymbol{h}) \| = \| (\boldsymbol{T}_1 - \boldsymbol{T}_2)(t\boldsymbol{h}) \| = \| \boldsymbol{T}_1(t\boldsymbol{h}) - \boldsymbol{T}_2(t\boldsymbol{h}) \|$

$$= \| \boldsymbol{f}(\boldsymbol{a} + t\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{T}_2(t\boldsymbol{h}) - [\boldsymbol{f}(\boldsymbol{a} + t\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{T}_1(t\boldsymbol{h})] \|$$

$$\leq \| \boldsymbol{f}(\boldsymbol{a} + t\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{T}_2(t\boldsymbol{h}) \| + \| \boldsymbol{f}(\boldsymbol{a} + t\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{T}_1(t\boldsymbol{h}) \|$$

Therefore, $\frac{\|\boldsymbol{T}(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} \leq \frac{\| \boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_2(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} + \frac{\| \boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_1(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|}$

Since T is a linear transformation, $T(t\boldsymbol{h}) = tT(\boldsymbol{h})$. Therefore,

$$\frac{|t|\|\boldsymbol{T}(\boldsymbol{h})\|}{|t|\|\boldsymbol{h}\|} \leq \frac{\| \boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_2(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} + \frac{\| \boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_1(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|}.$$

So, using (5.6) and (5.7) , we get

$$0 \leq \lim_{t\to 0} \frac{\|\boldsymbol{T}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} \leq \lim_{t\to 0} \frac{\| \boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_2(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} + \lim_{t\to 0} \frac{\| \boldsymbol{f}(\boldsymbol{a}+t\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T}_1(t\boldsymbol{h})\|}{\|t\boldsymbol{h}\|} = 0$$

Since $\frac{\|\boldsymbol{T}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|}$ is independent of t, this means $\frac{\|\boldsymbol{T}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = 0$, which means that $\| \boldsymbol{T}(\boldsymbol{h}) \| = 0$.

Now, $\boldsymbol{h}$ was any non-zero vector in $\boldsymbol{R^n}$. Further, $\mathbf{T}(\mathbf{0}) = \mathbf{0}$. Hence we conclude that $\mathbf{T}(\boldsymbol{h}) = \mathbf{0}$ for all $\boldsymbol{h} \in \boldsymbol{R^n}$. Thus $\mathbf{T} = \mathbf{T_1} - \mathbf{T_2}$ is the zero linear transformation. Thus, $\mathbf{T_1} = \mathbf{T_2}$. That is, the derivative is unique.

In the next example we find the derivatives of some standard functions.

Example 5.3 : i) Find the total derivative $\boldsymbol{f^1(a)}$, if $\boldsymbol{f : R^n \to R^m}$, $\boldsymbol{f(x) = c}$, where $\boldsymbol{c}$ is a fixed vector in $\boldsymbol{R^m}$ and $\boldsymbol{a \in R^n}$.

ii) If $\boldsymbol{f : R^n \to R^m}$ is a linear transformation, show that $D\boldsymbol{f(a)} = \boldsymbol{f}$ for every $\boldsymbol{a} \in \boldsymbol{R^n}$.

Solution : i) Since $\boldsymbol{f}$ is a constant function, we expect its derivative to be the zero transformation.

Here $\boldsymbol{f(a + h) - f(a) = c - c} = \mathbf{0}$.

If we take $\mathbf{T}$ to be the zero transformation,

$$\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = \lim_{h\to 0} \frac{\|f(a+th)-f(a)-T(h)\|}{\|h\|} = 0.$$

Hence $\boldsymbol{f^1}(\boldsymbol{a})$ exists and is equal to $\mathbf{0}$ for every $\boldsymbol{a} \in \boldsymbol{R^n}$.

ii) Since $\boldsymbol{f}$ is a linear transformation, $\boldsymbol{f(a + h) = f(a) + f(h)}$. If we take $\mathbf{T} = \boldsymbol{f}$,

$$\boldsymbol{r(h) = f(a + h) - f(a) - f(h)} = \mathbf{0} \Longrightarrow \frac{\|r(h)\|}{\|h\|} = 0 .$$

We have defined the total derivative of a function as a linear transformation. Now we prove a result about linear transformations which we may use later.

Proposition 5.1 : Every linear transformation $\mathbf{T}$ from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$ is continuous on $\boldsymbol{R^n}$.

Proof : If T is the zero linear transformation, it is clearly continuous. If T $\neq$ 0, let $\boldsymbol{p} \in \boldsymbol{R^n}$,

$\boldsymbol{p} = (p_1, p_2, ..., p_n)$, and $\varepsilon > 0$. Suppose $\{\boldsymbol{e_1}, \boldsymbol{e_2}, ..., \boldsymbol{e_n}\}$ is the standard basis for $\boldsymbol{R^n}$. Choose $\delta = \varepsilon/\mathrm{M}$, where M = $\| T(e_1) \| + \| T(e_2) \| +....... + \| T(e_n) \|$.

If $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ is such that $\| \boldsymbol{x} - \boldsymbol{p} \| < \delta$, then $|x_i - p_i| < \delta$ for i = 1, 2, ..., n.

Also, $\| \boldsymbol{x} - \boldsymbol{p} \| < \delta \Longrightarrow \| T(\boldsymbol{x}) - T(\mathbf{p}) \| = \| T(\mathbf{x} - \mathbf{p}) \|$

$$= \| T\big((x_1 - p_1)\boldsymbol{e_1} + (x_2 - p_2)\boldsymbol{e_2} + \cdots + (x_n - p_n)\boldsymbol{e_n}\big) \|$$

$$\leq |(x_1 - p_1)| \| T(\boldsymbol{e_1}) \| + |(x_2 - p_2)| \| T(\boldsymbol{e_2}) \| + ... + |(x_n - p_n)| \| T(\boldsymbol{e_n}) \|$$

$$< \delta(\| T(\boldsymbol{e_1}) \| + \| T(\boldsymbol{e_2}) \| +.......+\| T(\boldsymbol{e_n}) \|)$$

$$= \varepsilon$$

Thus, $\mathbf{T}$ is continuous at $\boldsymbol{p}$. Since $\boldsymbol{p}$ was an arbitrary point of $\boldsymbol{R^n}$, we conclude that $\mathbf{T}$ is continuous on $\boldsymbol{R^n}$.

In fact, since δ did not depend on $\boldsymbol{p}$, we can conclude that **T** is uniformly continuous on $\boldsymbol{R^n}$.

For functions of a single variable, we know that differentiability implies continuity. The next theorem shows that this holds for functions of several variables too.

Theorem 5.2 : If $\boldsymbol{f : R^n \to R^m}$ is differentiable at $\boldsymbol{p}$, then $\boldsymbol{f}$ is continuous at $\boldsymbol{p}$.

Proof : Since **f** is differentiable at $\boldsymbol{p}$, there exists a linear transformation $\mathbf{T_p}$ such that

$$\lim_{h\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{p}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{p})-\boldsymbol{T}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = \mathbf{0.}$$

Thus, $\forall\, \varepsilon > 0, \exists\, \delta_1 > 0,$ such that

$$\|\, h \,\| < \delta_1 \implies \frac{\|\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{a})-\boldsymbol{T_p}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} < \varepsilon/2$$

Choose $\delta_2 = \min(1, \delta_1)$. Then

$$\|\, h \,\| < \delta_2 \implies \|\, f(p+h) - f(p) - T_p(h) \,\| < (\varepsilon/2)\|\, h \,\| \le \varepsilon/2$$

By Proposition 5.1, $\mathbf{T_p}$ is continuous at **0**, and $\mathbf{T_p(0)} = \mathbf{0}$. So, there exists $\delta_3 > 0$, such that

$$\|\, \boldsymbol{h} \,\| < \delta_3 \implies \|\, \mathbf{T_p}(\boldsymbol{h}) \,\| < \varepsilon/2.$$

Now choose $\delta = \min(\delta_2, \delta_3)$. Then

$$\|\, \boldsymbol{h} \,\| < \delta \implies \|\, \boldsymbol{f}(\boldsymbol{p}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{p}) \,\| \le \|\, \boldsymbol{f}(\boldsymbol{p}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{p}) - \boldsymbol{T_p}(\boldsymbol{h}) \,\| + \|\, \boldsymbol{T_p}(\boldsymbol{h}) \,\|$$

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

Thus, $\lim_{\boldsymbol{h}\to\boldsymbol{0}} \boldsymbol{f}(\boldsymbol{p}+\boldsymbol{h}) = \boldsymbol{f}(\boldsymbol{p})$, and $\boldsymbol{f}$ is continuous at $\boldsymbol{p}$.

With your knowledge of functions of one variable, you would expect that the converse of Theorem 5.2 does not hold. That is, continuity does not imply differentiability. The following example shows that it is indeed so.

Example 5.4 : Consider the function $\boldsymbol{f : R \to R^2}$, $\boldsymbol{f}(x) = (|x|, |x|)$. We shall show that $\boldsymbol{f}$ is continuous at 0, but is not differentiable there.

Given $\varepsilon > 0$, choose $\delta = \varepsilon/\sqrt{2}$. Then

$$|x| < \delta \implies \|\, \boldsymbol{f}(x) \,\| = \|\, (|x|, |x|) \,\| < \sqrt{\delta^2 + \delta^2} = \sqrt{2}\,\delta = \varepsilon.$$

Hence, $\boldsymbol{f}$ is continuous at $x = 0$.

Now suppose $\boldsymbol{f}$ is differentiable at $x = 0$. Then there exists a linear transformation

$\mathbf{T} : \boldsymbol{R} \to \boldsymbol{R^2}$, such that

$$\lim_{h\to 0} \frac{\boldsymbol{f}(h)-\boldsymbol{f}(0)-\boldsymbol{T}(h)}{h} = \mathbf{0} \implies \lim_{h\to 0} \frac{(|h|,|h|)-\boldsymbol{T}(h)}{h} = \mathbf{0}$$

$$\implies \lim_{h\to 0}\left(\frac{|h|}{h}(1,1) - \boldsymbol{T}(1)\right) = \mathbf{0}$$

Now, (1, 1) and (−1, −1) are two distinct points in $\boldsymbol{R}^2$, and B((1, 1), 1) ∩ B((−1, −1), 1) = ∅.

For ε = 1, ∃ δ > 0, such that

$\| h \| < \delta \implies \| \frac{|h|}{h}(1,1) - \boldsymbol{T}(1)) \| < \varepsilon.$ ....(5.8)

Putting h = $\delta/2$ in (5.8), we get $\| \frac{|h|}{h}(1,1) - \boldsymbol{T}(1)) \| = \|(1,1) - T(1)\| < 1$. This means **T**(1) ∈ B((1, 1), 1).

Similarly, taking h = $-\delta/2$, we get that **T**(1) ∈ B((−1, −1), 1). But this contradicts the fact that B((1, 1), 1) and B((−1, − 1), 1) are disjoint.

Thus, ***f*** is not differentiable at x = 0.

If $\boldsymbol{f} : \boldsymbol{R}^n \to \boldsymbol{R}^m$, then, as you know, we can write $\boldsymbol{f} = (f_1, f_2, ..., f_m)$, where each $f_i : \boldsymbol{R}^n \to \boldsymbol{R}$,

i = 1, 2, ..., m. These f_is are called coordinate functions of ***f***. Similarly, a linear transformation

T : $\boldsymbol{R}^n \to \boldsymbol{R}^m$ can be written as **T** = $(T_1, T_2, ..., T_m)$, where each T_i is a linear transformation from $\boldsymbol{R}^n$ to $\boldsymbol{R}$.

Theorem 5.3 : Let $\boldsymbol{f} = (f_1, f_2, ..., f_m) : \boldsymbol{R}^n \to \boldsymbol{R}^m$, and $\boldsymbol{p} \in \boldsymbol{R}^n$. ***f*** is differentiable at ***p***, if and only if each f_i, $1 \le i \le$ m is differentiable at ***p***.

Proof : ***f*** is differentiable at ***p*** if and if there exists a linear transformation $\mathbf{T_p} : \boldsymbol{R}^n \to \boldsymbol{R}^m$, such

that $\lim_{h\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{p}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{p})-\boldsymbol{T}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = 0$, that is, if only if

$\lim_{h\to 0} \frac{\|\sum_1^m [f_i(\boldsymbol{p}+\boldsymbol{h})-f_i(\boldsymbol{p})-T_i(\boldsymbol{h})]\boldsymbol{e}_i\|}{\|\boldsymbol{h}\|} = 0$, where $\{\boldsymbol{e_1}, \boldsymbol{e_2}, ..., \boldsymbol{e_m}\}$ is the standard basis of $\boldsymbol{R}^m$,

if and only if, $\lim_{h\to 0} \frac{|f_i(\boldsymbol{p}+\boldsymbol{h})-f_i(\boldsymbol{p})-T_i(\boldsymbol{h})|}{\|\boldsymbol{h}\|} = 0, \forall\, i,\ 1 \le i \le$ m.

That is, if and only if each f_i is differentiable and $Df_i = T_i$, $\forall\, i$, $1 \le i \le$ m.

Thus, $D\boldsymbol{f}(\mathbf{p}) = \boldsymbol{T_p} = (Df_1(\mathbf{p}), Df_2(\mathbf{p}), ..., Df_m(\mathbf{p}))$.

Theorem 5.4 : Let $\boldsymbol{f} : \boldsymbol{R}^n \to \boldsymbol{R}^m$ and $\boldsymbol{g} : \boldsymbol{R}^n \to \boldsymbol{R}^m$ be two functions differentiable at $\boldsymbol{p} \in \boldsymbol{R}^n$. If $k \in \boldsymbol{R}$, then ***f*** + ***g*** and $k\boldsymbol{f}$ are also differentiable at ***p***. Moreover,

D(***f*** + ***g***)(**p**) = D***f***(**p**) + D***g***(**p**), and D(k***f***)(**p**) = kD***f***(**p**).

Proof : Let $Df(p) = \mathbf{T_1}$, and $Dg(p) = \mathbf{T_2}$. Then $\mathbf{T_1} + \mathbf{T_2}$ is also a linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$, and

$$0 \leq \lim_{h\to 0} \frac{\|(f+g)(p+h)-(f+g)(p)-(T_1+T_2)(h)\|}{\|h\|}$$

$$= \lim_{h\to 0} \frac{\|[f(p+h)-f(p)-T_1(h)]+ [g(p+h)-g(p)-T_2(h)]\|}{\|h\|}$$

$$\leq \lim_{h\to 0} \frac{\|f(p+h)-f(p)-T_1(h)\|}{\|h\|} + \lim_{h\to 0} \frac{\|g(p+h)-g(p)-T_2(h)\|}{\|h\|} = 0.$$

Therefore, $\boldsymbol{f} + \boldsymbol{g}$ is differentiable at $\boldsymbol{p}$, and $D(\boldsymbol{f} + \boldsymbol{g})(\boldsymbol{p}) = \mathbf{T_1} + \mathbf{T_2} = Df(\boldsymbol{p}) + Dg(\boldsymbol{p})$.

Now, $\displaystyle\lim_{h\to 0} \frac{\|kf(p+h)-kf(p)-kT_1(h)\|}{\|h\|} = |k| \lim_{h\to 0} \frac{\|f(p+h)-f(p)-T_1(h)\|}{\|h\|} = 0.$

Therefore, $k\mathbf{f}$ is also differentiable and $D(k\boldsymbol{f})(\boldsymbol{p}) = k\mathbf{T_1} = kD\boldsymbol{f}(\boldsymbol{p})$.

5.3 PARTIAL DERIVATIVES

We know that the derivative of a function of one variable denotes the rate at which the function value changes with change in the domain variable. In the case of functions of several variables, change in the domain vector variable means a change in any or all of its components. But if we consider change in only one component and study the rate at which the function value changes, we get what is known as the partial derivative of the function. Corresponding to each component of the variable, there will be a partial derivative. Here is the formal definition.

Definition 5.2 Let $\boldsymbol{f} : E \to \boldsymbol{R^m}$, where $E \subseteq \boldsymbol{R^n}$. Let $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ be an interior point of E. Then for every i, i = 1, 2, ..., n, the limit

$\displaystyle\lim_{h\to 0} \frac{f(x_1,x_2,...,x_i+h,x_{i+1},...,x_n)- f(x_1,x_2,...,x_n)}{h}$, if it exists, is called the ith partial derivative of $\boldsymbol{f}$ with respect to x_i at $\mathbf{x}$. It is denoted by $\frac{\partial f}{\partial x_i}$, $f_{x_i}, or\ D_i f$. We write $\frac{\partial f}{\partial x_i}(\boldsymbol{x})$ to indicate the point at which the partial derivative is calculated.

Remark 5.2 : i) If a function $\boldsymbol{f}$ has partial derivatives at every point of the set E, we say that $\boldsymbol{f}$ has partial derivatives on E.

ii) It is clear from the definition that a partial derivative can be defined at an interior point of E, and not on its boundary.

iii) If a function has a partial derivative at a point, its value depends on the values of the function in a neighbourhood of that point. So, if the function values outside this neighbourhood are changed, it does not affect the value of the partial derivative.

The following examples will make the concept clear.

Example 5.5 : Find the partial derivative of the function, $f(x, y, z) = xyz + x^2z$.

Solution : This is a real-valued function. You are already familiar with the partial differentiation of such a function.

$\frac{\partial f}{\partial x} = \lim\limits_{h\to 0} \frac{(x+h)yz+(x+h)^2z-xyz-x^2z}{h}$ = $yz + 2xz$. Similarly, you can check that $f_y = xz$, and $f_z = xy + x^2$.

Let us take a vector-valued function in the next example.

Example 5.6 : Find the partial derivatives of the function, $\boldsymbol{f} : \boldsymbol{R}^3 \to \boldsymbol{R}^2$, $\boldsymbol{f}(x, y, z) = (xy, z^2)$, if they exist.

Solution : $\lim\limits_{h\to 0} \frac{\left((x+h)y,z^2\right)-(xy,z^2)}{h} = \lim\limits_{h\to 0} \frac{((x+h)y-xy\ ,\ 0)}{h}$

$$= \left(\lim_{h\to 0} \frac{(x+h)y-xy}{h}, \lim_{h\to 0} \frac{0}{h}\right) = (y, 0).$$

Therefore, $\frac{\partial f}{\partial x}$ = (y, 0).

Proceeding similarly, we find that $\frac{\partial f}{\partial y}$ = (x, 0), and $\frac{\partial f}{\partial z}$ = (0, 2z).

You must have observed that the partial derivatives of a vector function are formed by taking the partial derivatives of its coordinate functions. In fact we have the following theorem, which establishes the connection between differentiability of a vector-valued function and the existence of partial derivatives of its coordinate functions

Theorem 5.5 : Let E be an open subset of $\boldsymbol{R}^n$, and $\boldsymbol{f}$: E $\to \boldsymbol{R}^m$. Suppose $\boldsymbol{f} = (f_1, f_2, ..., f_m)$ is differentiable at $\boldsymbol{p} \in$ E. Then the partial derivatives $\frac{\partial f_i}{\partial x_j}$ exist for i = 1, 2, ..., m, j = 1, 2, ..., n.

Proof : Since $\boldsymbol{f}$ is differentiable at $\boldsymbol{p}$, there exists a linear transformation **T**, such that $\lim\limits_{h\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{p}+\boldsymbol{h})-\boldsymbol{f}(\boldsymbol{p})-\boldsymbol{T}(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = 0.$ Let $\mathbf{h} = t\mathbf{e}_j$, where $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ is the standard basis of $\boldsymbol{R}^n$**.** Then, $\boldsymbol{h} \to \boldsymbol{0}$ if and only if t $\to$ 0. Thus,

$$\lim_{t\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{p}+t\boldsymbol{e}_j)-\boldsymbol{f}(\boldsymbol{p})-\boldsymbol{T}(t\boldsymbol{e}_j)\|}{|t|} = 0. \text{ Therefore, } \lim_{t\to 0} \frac{\boldsymbol{f}(\boldsymbol{p}+t\boldsymbol{e}_j)-\boldsymbol{f}(\boldsymbol{p})}{t} = \mathbf{T}(e_j).$$

That is,

$$\left(\lim_{t\to 0} \frac{f_1(\boldsymbol{p}+t\boldsymbol{e}_j)-f_1(\boldsymbol{p})}{t}, \lim_{t\to 0} \frac{f_2(\boldsymbol{p}+t\boldsymbol{e}_j)-f_2(\boldsymbol{p})}{t}, \dots, \lim_{t\to 0} \frac{f_m(\boldsymbol{p}+t\boldsymbol{e}_j)-f_m(\boldsymbol{p})}{t}\right)$$

= $\mathbf{T}(e_j)$.

Hence the limits exist, and $\frac{\partial f_i}{\partial x_j}(\boldsymbol{p})$ exists for all i = 1, 2, ..., m.

Since j was arbitrary, we conclude that $\frac{\partial f_i}{\partial x_j}(\boldsymbol{p})$ exists for all i = 1, 2, ..., m, j = 1, 2, ..., n.

If $\boldsymbol{f} : E \to \boldsymbol{R}^m$, where E is an open subset of $\boldsymbol{R}^n$, and if $\boldsymbol{f}$ is differentiable at $\boldsymbol{p} \in E$, then using Theorem 5.5, the matrix of the linear transformation **T** can be written as

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{p}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{p}) & \cdots & . & . & \frac{\partial f_1}{\partial x_n}(\boldsymbol{p}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{p}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{p}) & . & . & . & \frac{\partial f_2}{\partial x_n}(\boldsymbol{p}) \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{p}) & \frac{\partial f_m}{\partial x_2}(\boldsymbol{p}) & . & . & . & \frac{\partial f_m}{\partial x_n}(\boldsymbol{p}) \end{pmatrix}$$

This $m \times n$ matrix is called the Jacobian matrix of $\boldsymbol{f}$ at $\boldsymbol{p}$, and is denoted by $[\boldsymbol{f}'(\boldsymbol{p})]$ or $[D\boldsymbol{f}(\boldsymbol{p})]$.

If m = n, the determinant of the Jacobian matrix is called the Jacobian of $\boldsymbol{f}$ at $\boldsymbol{p}$, and is denoted by $\frac{\partial(f_1,f_2,\ldots,f_m)(p)}{\partial(x_1,x_2,\ldots,x_m)}$.

Thus, if $\boldsymbol{f}$ is differentiable at $\boldsymbol{p}$, then the total derivative of $\boldsymbol{f}$ at $\boldsymbol{p}$, $\mathbf{T} : \boldsymbol{R}^n \to \boldsymbol{R}^m$ is given by the Jacobian matrix. For $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \boldsymbol{R}^n$,

$$\mathbf{T}(\boldsymbol{x}) = [\boldsymbol{f}'(\boldsymbol{p})]\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ . \\ x_n \end{pmatrix}.$$

When m = 1, $\boldsymbol{f}$ is a real-valued function, and $\mathbf{T}(\boldsymbol{e}_j) = \frac{\partial f}{\partial x_j}(\boldsymbol{p})$. Hence, the Jacobian matrix of **T** is the row matrix, $[\frac{\partial f}{\partial x_1}(\boldsymbol{p}) \ \ \frac{\partial f}{\partial x_2}(\boldsymbol{p}) \ \ \ldots \ \ \frac{\partial f}{\partial x_n}(\boldsymbol{p})]$.

The vector form, $(\frac{\partial f}{\partial x_1}(\boldsymbol{p}), \frac{\partial f}{\partial x_2}(\boldsymbol{p}), \ldots, \frac{\partial f}{\partial x_n}(\boldsymbol{p}))$ is called the **gradient of *f* at *p*** , and is denoted by $\nabla f(\boldsymbol{p})$, or grad$f(\boldsymbol{p})$.

If $\boldsymbol{h} = (h_1, h_2, \ldots, h_n) \in \boldsymbol{R}^n$,

$$\mathbf{T}_{\mathbf{p}}(\boldsymbol{h}) = [\frac{\partial f}{\partial x_1}(\boldsymbol{p}) \ \ \frac{\partial f}{\partial x_2}(\boldsymbol{p}) \ \ \ldots \ \ \frac{\partial f}{\partial x_n}(\boldsymbol{p})]\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ . \\ h_n \end{pmatrix}.$$

Thus, $\mathbf{T}(\boldsymbol{h}) = \frac{\partial f}{\partial x_1}(\boldsymbol{p})h_1 + \frac{\partial f}{\partial x_2}(\boldsymbol{p})h_2 + \ \ldots \ + \frac{\partial f}{\partial x_n}(\boldsymbol{p})h_n$, or $\mathbf{T}_p(\boldsymbol{h}) = \nabla \boldsymbol{f}(\boldsymbol{p}) \bullet \boldsymbol{h}$.

So, we can say that the total derivative $\mathbf{T}_p$ of a real-valued function is given by

$\mathbf{T}_p(\boldsymbol{h}) = \nabla \boldsymbol{f}(\boldsymbol{p}) \bullet \boldsymbol{h}$.

Example 5.7 : Find the Jacobian matrix of i) $\boldsymbol{f}(x, y) = (x^2y, e^{xy})$

ii) $\boldsymbol{f}(x, y, z) = (x\sin z, -ye^z)$ at $(1, 2, -1)$.

Solution : i) $f_1(x, y) = x^2y$, and $f_2(x, y) = e^{xy}$. Therefore, $\frac{\partial f_1}{\partial x} = 2xy$, $\frac{\partial f_1}{\partial y} = x^2$,

$\frac{\partial f_2}{\partial x} = ye^{xy}$, and $\frac{\partial f_2}{\partial y} = xe^{xy}$.

Hence, $[f^{i}(x, y)] = \begin{bmatrix} 2xy & x^2 \\ ye^{xy} & xe^{xy} \end{bmatrix}$

$\frac{\partial f_1}{\partial x} = sinyz$, and $\frac{\partial f_1}{\partial x}$ $(1, 2, -1) = -\sin 2$

$\frac{\partial f_1}{\partial y}(1, 2, -1) = -\cos 2$, $\frac{\partial f_1}{\partial z}(1, 2, -1) = 2\cos 2$,

$\frac{\partial f_2}{\partial x}$ $(1, 2, -1) = 0$, $\frac{\partial f_2}{\partial y}(1, 2, -1) = -e^{-1}$, $\frac{\partial f_2}{\partial z}(1, 2, -1) = -2e^{-1}$.

Thus, $[f^{i}(1, 2, -1)] = \begin{pmatrix} -sin2 & -2\,cos2 & 2\,cos2 \\ 0 & -e^{-1} & -2e^{-1} \end{pmatrix}$

In the next section we shall consider yet another type of derivative.

5.4 DIRECTIONAL DERIVATIVES

Partial derivatives measure the rate of change of a function in the directions of the standard basis vectors. Directional derivatives measure the rate of change in any given direction.

Definition 5.3 : Let $f : E \to \boldsymbol{R}$, where E is an open subset of $\boldsymbol{R}^n$. Let $\boldsymbol{u}$ be a unit vector in $\boldsymbol{R}^n$, and $\boldsymbol{p} \in E$. If $\lim_{t\to 0} \frac{f(\boldsymbol{p}+t\boldsymbol{u})- f(\boldsymbol{p})}{t}$ exists, then it is called the directional derivative of f at $\boldsymbol{p}$ in the direction $\boldsymbol{u}$. It is denoted by $\frac{\partial f}{\partial u}(p)$ or $f_{\boldsymbol{u}}(\boldsymbol{p})$.

Example 5.8 : Find the directional derivatives of the following functions:

i) $f(x, y) = 2xy + 3y^2$ at $\boldsymbol{p} = (1, 1)$, in the direction of $\boldsymbol{v} = (1, 1)$.
ii) $f(x, y) = x^2y$ at $\boldsymbol{p} = (3, 4)$, in the direction of $\boldsymbol{v} = (1, 1)$.

Solution : i) The unit vector $\boldsymbol{u}$ in the given direction is $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Hence the required directional derivative is $\lim_{t\to 0} \frac{f\left((1,1)+ t(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})\right)- f(1,1)}{t}$.

$$= \lim_{t\to 0} \frac{f\left(\left(1+\frac{t}{\sqrt{2}},\ 1+\frac{t}{\sqrt{2}}\right)\right)- f(1,1)}{t}$$

$$= \lim_{t\to 0} \frac{2(1+\frac{t}{\sqrt{2}})^2+3(1+\frac{t}{\sqrt{2}})^2-5}{t} = \lim_{t\to 0} \frac{5\sqrt{2t}+5t^2/2}{t} = 5\sqrt{2}.$$

ii) We have the same unit vector $\boldsymbol{u}$ here. Therefore,

$$\mathrm{D}_{\boldsymbol{u}}f(\boldsymbol{p}) = \lim_{t\to 0} \frac{f\left(\left(3+\frac{t}{\sqrt{2}},\ 4+\frac{t}{\sqrt{2}}\right)\right) - f(3,4)}{t} = \lim_{t\to 0} \frac{\left(3+\frac{t}{\sqrt{2}}\right)^2 (4+t/\sqrt{2}) - 36}{t} = \frac{33\sqrt{2}}{2}.$$

Example 5.9 : Find the directional derivatives, if they exist, in the following cases:

i) $f(x, y) = \begin{Bmatrix} x+y, if\ xy = 0 \\ 1,\ otherwise \end{Bmatrix}$, at (0, 0), $\boldsymbol{u} = (u_1, u_2)$, $\|\boldsymbol{u}\| = 1$

ii) $f(x, y) = \begin{Bmatrix} \frac{xy^2}{x^2+y^4}, if\ (x,y) \neq (0,0) \\ 0, \quad\ if\ (x,y) = (0,0) \end{Bmatrix}$ at (0,0), $\boldsymbol{u} = (1/\sqrt{2}, 1/\sqrt{2})$.

Solution: i) if $u_1 \neq 0$, $u_2 \neq 0$, $\lim_{t\to 0} \frac{f((0+tu_1,\ 0+tu_2)) - f(0,0)}{t} = \lim_{t\to 0} \frac{1-0}{t}$, which does not exist. If either u_1 or u_2 is zero, we get the standard basis vectors, (1, 0) and (0, 1).

If $\boldsymbol{u}$ = (1, 0), $\lim_{t\to 0} \frac{f((0+t,\ 0)) - f(0,0)}{t} = \lim_{t\to 0} \frac{t-0}{t} = 1.$

Similarly, if $\boldsymbol{u}$ = (0, 1), $\lim_{t\to 0} \frac{f((0,\ 0+t)) - f(0,0)}{t} = 1.$

Thus, the directional derivatives in these two directions exist, and are equal to one. In any other direction, the derivative does not exist. Note that the directional derivative in the direction (1, 0) is f_x, and that in the direction (0, 1) is f_y. Thus, this function has both the partial derivatives at (0, 0).

ii) $\lim_{t\to 0} \frac{f\left(\left(0+\frac{t}{\sqrt{2}},\ 0+\frac{t}{\sqrt{2}}\right)\right) - f(0,0)}{t} == \lim_{t\to 0} \frac{\frac{t^3/2\sqrt{2}}{\frac{t^2}{2}+\frac{t^4}{4}} - 0}{t} == \lim_{t\to 0} \frac{\sqrt{2}}{(2+t^2)} = 1/\sqrt{2}.$

Thus, $\mathrm{D}_{\boldsymbol{u}}f(0, 0) = 1/\sqrt{2}$.
In fact, if we take $\boldsymbol{u} = (\cos\theta, \sin\theta)$, then we can show that f has directional derivative at (0, 0) in the direction of $\boldsymbol{u}$, whatever be θ. That is, the directional derivatives of f at (0, 0) exist in all directions. But you can easily show that this function is not continuous at (0, 0) by using the two-path test. Recall, that you need to show that the limits of f, at (0, 0) along two different paths are different. Then by Theorem 5.2 we can conclude that f is not differentiable at (0, 0).
This example shows that the existence of all directional derivatives at a point does not guarantee differentiability there. But we have the following theorem:

Theorem 5.7: Let f : E → $\boldsymbol{R}$, where E is an open subset of $\boldsymbol{R}^n$. If f is differentiable at $\boldsymbol{p} \in \boldsymbol{R}^n$, then the directional derivatives of f at $\boldsymbol{p}$ exist in all directions.
Proof : Since f is differentiable at $\boldsymbol{p}$, there exists a linear transformation, T: $\boldsymbol{R}^n \to \boldsymbol{R}$, such that

$$\lim_{h\to 0} \frac{|f(\boldsymbol{p}+\boldsymbol{h}) - f(\boldsymbol{p}) - T(\boldsymbol{h})|}{\|\boldsymbol{h}\|} = 0.$$

Let $\boldsymbol{u}$ be any unit vector in $\boldsymbol{R}^n$, and take $\boldsymbol{h} = t\boldsymbol{u}$. Then $\boldsymbol{h} \to \boldsymbol{0}$, as $t \to 0$. Therefore,
$\lim_{t\to 0} \frac{|f(\boldsymbol{p}+t\boldsymbol{u}) - f(\boldsymbol{p}) - T(t\boldsymbol{u})|}{|t|} = 0.$ This means,

$$\lim_{t\to 0} \left| \frac{f(\boldsymbol{p}+t\boldsymbol{u})-f(\boldsymbol{p})-tT(\boldsymbol{u})}{|t|} \right| = 0.$$ That is,

$$\lim_{t\to 0} \frac{f(\boldsymbol{p}+t\boldsymbol{u})-f(\boldsymbol{p})}{t} = \mathrm{T}(\boldsymbol{u}), \quad \text{or,} \quad \mathrm{D}_{\boldsymbol{u}}f(\boldsymbol{p}) = \mathrm{T}(\boldsymbol{u}). \qquad \text{......................(5.5)}$$

Since $\boldsymbol{u}$ was an arbitrary unit vector, we conclude that the directional derivatives of f at $\boldsymbol{p}$ exist in all directions.

Now, if $\boldsymbol{u} = (u_1, u_2, ..., u_n)$, $\mathrm{T}(\boldsymbol{u}) = \mathrm{T}(u_1\boldsymbol{e_1} + u_2\boldsymbol{e_2} + ... + u_n\boldsymbol{e_n})$, where $\{\boldsymbol{e_1}, \boldsymbol{e_2}, ..., \boldsymbol{e_n}\}$ is the standard basis of $\boldsymbol{R^n}$. Therefore, by (5.5),

$$\begin{aligned}
\mathrm{T}(\mathbf{u}) &= u_1\mathrm{T}(\boldsymbol{e_1}) + u_2\mathrm{T}(\boldsymbol{e_2}) + ... + u_n\mathrm{T}(\boldsymbol{e_n}) \\
&= u_1\, D_{e_1}f(\mathbf{p}) + u_2\, D_{e_2}f(\mathbf{p}) + ... + u_n\, D_{e_n}f(\mathbf{p}) \\
&= u_1 \frac{\partial f(p)}{\partial x_1} + u_2 \frac{\partial f(p)}{\partial x_2} + ... + u_n \frac{\partial f(p)}{\partial x_n} \\
&= \nabla \boldsymbol{f(p)} \bullet \boldsymbol{u}
\end{aligned}$$

Thus, $\mathrm{D}_{\boldsymbol{u}}f(\boldsymbol{p}) = \nabla \boldsymbol{f(p)} \bullet \boldsymbol{u}$ .... (5.6)

(5.6) gives an easy way to find a directional derivative of a differentiable function, if its partial derivatives are known. For example, if $f(x, y) = x^2 + y^2$, then f_x and f_y at (1, 2) are 2 and 4, respectively. So, the directional derivative of f at (1, 2) in the direction 2i – 3j is given by $(2i + 4j) \circ \left(\frac{2i-3j}{\sqrt{13}}\right) = \frac{-8}{\sqrt{13}}$.

This concept of directional derivatives can be extended to vector-valued functions. The directional derivative of a vector-valued function is a vector formed by the directional derivatives of its coordinate functions. Thus, to find the directional derivative of

$\boldsymbol{f}(x, y) = (x + y, x^2)$, at (1, 2) in the direction of (3, 4) , we first find the directional derivatives of $f_1(x, y) = x + y$, and $f_2(x, y) = x^2$. You can check that these are 7/5 and 6/5, respectively. Therefore, the required directional derivative of $\boldsymbol{f}$ is (7/5, 6/5).

We have seen in Theorems 5.6 and 5.7, that differentiability of f at a point guarantees the existence of partial and directional derivatives there. We have also noted that the converse statements are not true. Our next theorem gives us a sufficient condition which guarantees the differentiability of a function at a point.

Theorem 5.8 : Let E be an open subset of $\mathbf{R^n}$, and $\boldsymbol{f}: \mathrm{E} \to \boldsymbol{R^m}$, $\boldsymbol{f} = (f_1, f_2, ..., f_m)$. If all the partial derivatives, $\mathrm{D}_j f_i(\boldsymbol{x})$ of all the coordinate functions of $\boldsymbol{f}$ exist in an open set containing $\boldsymbol{a}$, and if each function $\mathrm{D}_j f_i$ is continuous at $\boldsymbol{a}$, then $\boldsymbol{f}$ is differentiable at $\boldsymbol{a}$.

Proof : In the light of Theorem 5.3, it is enough to prove this theorem for the case m = 1. So, we consider a scalar function f from $\boldsymbol{R^n}$ to $\boldsymbol{R}$, all whose partial derivatives $\mathrm{D}_j f$ are continuous at $\boldsymbol{a}$. Since E is open, for a given $\varepsilon > 0$, we can find r > 0, such that the open ball,

$\mathrm{B}(\boldsymbol{a}, \mathrm{r}) \subset E$, and $\|\, \boldsymbol{x} - \boldsymbol{a} \,\| < \mathrm{r} \Longrightarrow |\, \mathrm{D}_j f(\boldsymbol{x}) - \mathrm{D}_j f(\boldsymbol{a}) \,| < \varepsilon/\mathrm{n}$, for j = 1, 2, ... , n. ....(5.7)

Now, suppose $\boldsymbol{h} = (h_1, h_2, ... , h_n)$, $\|\boldsymbol{h}\| < \mathrm{r}$. Let $\boldsymbol{v_0} = \boldsymbol{0}$, $\boldsymbol{v_1} = h_1\boldsymbol{e_1}$, $\boldsymbol{v_2} = \boldsymbol{v_1} + h_2\boldsymbol{e_2}$, ... ,

$\boldsymbol{v_n} = \boldsymbol{v_{n-1}} + h_n\boldsymbol{e_n}$. Then $f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = \sum_{j=1}^{n}[f(\boldsymbol{a} + \boldsymbol{v_j}) - f(\boldsymbol{a} + \boldsymbol{v_{j-1}})]$. ....(5.8)

Since $||\boldsymbol{v_j}|| < r$, $\boldsymbol{v_j} \in B(\boldsymbol{a}, r)$, and since $B(\boldsymbol{a}, r)$ is convex, the line segment joining the points,

$\boldsymbol{a} + \boldsymbol{v_{j-1}}$ and $\boldsymbol{a} + \boldsymbol{v_j}$ lies in it, for all j = 1, 2, ... , n. Therefore, we can apply the Mean Value Theorem to the j^{th} term in the sum (5.8), and get

$f(\boldsymbol{a} + \boldsymbol{v_j}) - f(\boldsymbol{a} + \boldsymbol{v_{j-1}}) = h_j D_j f(\boldsymbol{a} + \boldsymbol{v_{j-1}} + \theta_j h_j \boldsymbol{e_j})$, for some $\theta_j \in (0, 1)$. Then, using (5.7), we can write

$|f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) - \sum_{j=1}^{n} h_j (D_j f)(\boldsymbol{a})| = |\sum_{j=1}^{n} h_j (D_j f)\ (\boldsymbol{a} + \boldsymbol{v_{j-1}} + \theta_j h_j \boldsymbol{e_j}) - \sum_{j=1}^{n} h_j (D_j f)(\boldsymbol{a})|$

$$\leq \frac{1}{n}\sum_{j=1}^{n} |h_j|\varepsilon \leq ||\boldsymbol{h}||\varepsilon, \text{ for all } \boldsymbol{h}, \text{ such that } ||\boldsymbol{h}|| < r.$$

This means that

$\lim\limits_{h\to 0} \frac{\|\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - \boldsymbol{f}'(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = \boldsymbol{0}$, where f' is the linear transformation, whose matrix $[f'(\boldsymbol{a})]$ consists of the row, $(D_1 f(\boldsymbol{a}), D_2 f(\boldsymbol{a}), ..., D_n f(\boldsymbol{a}))$.

Thus, f is differentiable at $\boldsymbol{a}$.

Definition 5.4 : A function $\boldsymbol{f} : E \to \boldsymbol{R}^m$, $\boldsymbol{f} = (f_1, f_2, ..., f_m)$, where E is an open subset of $\boldsymbol{R}^n$, is said to be **continuously differentiable**, or, a C^1 function, if $D_j f_i$ is continuous on E for all j, j = 1, 2, ..., n, and for all i, i = 1, 2, ..., m.

The continuity of partial derivatives assumed in Theorem 5.8, is only a sufficient condition, and not a necessary one. That is, there may be functions which are differentiable at a point, but do not have continuous partial derivatives there. We now give you an example, and ask you to work out the details (See Exercise 3.)

Example 5.10 : Consider the function $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ given by

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

This function is differentiable at (0, 0), but neither $f_x = \begin{cases} 2x sin\frac{1}{x} - cos\frac{1}{x}, & if\ x \neq 0 \\ 0, & if\ x = 0 \end{cases}$,

nor $f_y = \begin{cases} 2y sin\frac{1}{y} - cos\frac{1}{y}, & if\ y \neq 0 \\ 0, & if\ y = 0 \end{cases}$ is continuous at (0, 0).

Here are some exercises that you should try.

Exercises:

1) Show that the following function is differentiable at all $\boldsymbol{x}$ in $\boldsymbol{R}^n$.

$\boldsymbol{f} : \boldsymbol{R}^n \to \boldsymbol{R}$, $f(\boldsymbol{x}) = \boldsymbol{x} \cdot \mathbf{T}(\boldsymbol{x})$, where $\mathbf{T} : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is a linear transformation.

2) Let $\boldsymbol{f}(x, y) = (x^3 + x, x^2 - y^2, 2x + 3y^3)$, $\boldsymbol{p} = (2, 1)$, $\boldsymbol{v} = (4, 5)$. Compute the partial derivatives of $\boldsymbol{f}$, and the directional derivative of $\boldsymbol{f}$ in the direction $\boldsymbol{v}$, at $\boldsymbol{p}$.

3) Prove the assertions in Example 5.10. (Hint : To show that f is differentiable, check that $f(h, k) - f(0, 0) - h(h\sin\frac{1}{h}) + k(k\sin\frac{1}{k}) = 0$, and so, $Df = (h\sin\frac{1}{h}, k\sin\frac{1}{k})$).

5.5 SUMMARY

In this unit we have extended the concept of differentiation from functions of one variable to functions of several variables. Apart from the total derivatives, we have also defined partial derivatives, and directional derivatives. We have proved that differentiability implies the existence of all partial and directional derivatives at a point, but the converse is not true. As in the case of functions of one variable, we prove that differentiable functions are continuous, but not vice versa. We have also derived a sufficient condition for differentiability in terms of the partial derivatives.

6

DERIVATIVES OF HIGHER ORDER

Unit Structure

6.0 OBJECTIVES

After reading this chapter, you should be able to

- differentiate a composite of two vector-valued functions
- define and calculate derivatives of higher order
- derive the conditions for the equality of mixed partial derivatives
- state and prove the Mean Value Theorem

6.1 INTRODUCTION

In the last chapter you have seen how functions of several variables are differentiated. Now we shall start by discussing how a composite function of two differentiable functions can be differentiated. The Jacobian matrix introduced in the last chapter proves useful in this.

One of the important applications of derivatives is the location of extreme points of a function. In the next chapter we are going to see how this concept can be extended to scalar functions of several variables. But we shall do the necessary spade-work in this chapter. So, we shall introduce higher order derivatives. We shall also study the conditions under which mixed partial derivatives are equal. You may recall that the Mean Value Theorem was one of the most important theorems that you studied in Calculus in F. Y. B. Sc. We shall see whether this theorem can be applied to functions of several variables.

6.2 JACOBIAN MATRIX AND CHAIN RULE

We have seen in Theorem 5.5, that if $\boldsymbol{f}$: $\boldsymbol{R^n} \to \boldsymbol{R^m}$, is differentiable at $\boldsymbol{p}$, then all partial derivatives of all coordinate functions of $\boldsymbol{f}$ exist at $\boldsymbol{p}$. That is, if $\boldsymbol{f} = (f_1, f_2, \dots, f_m)$, then $D_j f_i(\boldsymbol{p})$ exists for all i = 1, 2, ..., m and all j = 1, 2, ..., n. We have also seen that if $\{\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_n}\}$ is the standard basis for $\boldsymbol{R^n}$, then

$\boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{e_j}) = (D_j f_1(\boldsymbol{p}), D_j f_2(\boldsymbol{p}), \dots, D_j f_m(\boldsymbol{p}))$.

If $\boldsymbol{h} = \sum_{j=1}^{n} a_j \boldsymbol{e_j}$ is a vector in $\boldsymbol{R^n}$, then

$\boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{h}) = \sum_{j=1}^{m} a_j \boldsymbol{f}'(\boldsymbol{p})(\mathbf{e}_j)$. $\boldsymbol{f}'(\boldsymbol{p})$, which is a linear transformation from $\boldsymbol{R^n}$ to $\boldsymbol{R^m}$, thus has the matrix,

,

$$\begin{pmatrix} D_1 f_1(\boldsymbol{p}) & D_2 f_1(\boldsymbol{p}) & .. & . & . & D_n f_1(\boldsymbol{p}) \\ D_1 f_2(\boldsymbol{p}) & D_2 f_2(\boldsymbol{p}) & . & . & . & D_n f_2(\boldsymbol{p}) \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ D_1 f_m(\boldsymbol{p}) & D_2 f_m(\boldsymbol{p}) & . & . & . & D_n f_m(\boldsymbol{p}) \end{pmatrix}$$

As we have already mentioned in Chapter 5, this m × n matrix, called the Jacobian matrix, is denoted by $[D\boldsymbol{f}(\boldsymbol{p})]$. The k^{th} row of this matrix is the gradient vector, $\nabla f_k(\boldsymbol{p})$, and the j^{th} column is the image of $\boldsymbol{e}_j$ under the linear transformation $D_j\boldsymbol{f}(\boldsymbol{p})$.

Thus, the Jacobian matrix of $\boldsymbol{f}$ is formed by all first order partial derivatives of $\boldsymbol{f}$. This means, we can write the Jacobian matrix of any function, all of whose partial derivatives exist. As we have noted earlier, the existence of partial derivatives does not guarantee differentiability. So, even when a function is not differentiable we would be able to write its Jacobian matrix, provided all its partial derivatives exist.

If $f : \boldsymbol{R^n} \to \boldsymbol{R}$, then its Jacobian matrix, if it exists, will be a 1 x n matrix, or a matrix vector.

If $\boldsymbol{f} : \boldsymbol{R^n} \to \boldsymbol{R^m}$ is differentiable at $\boldsymbol{p} \in \boldsymbol{R^n}$, and if $\boldsymbol{h}$ is any vector in $\boldsymbol{R^n}$, then

$\boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{h}) = [D\boldsymbol{f}(\boldsymbol{p})]\boldsymbol{h}$ is obtained by multiplying the m × n matrix $[D\boldsymbol{f}(\boldsymbol{p})]$ with the n × 1 column matrix $\boldsymbol{h}$. Thus,

$||\boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{h})|| = || \sum_{j=1}^{m} (\nabla f_j(\boldsymbol{p}) \bullet \boldsymbol{h})\, \boldsymbol{e_j} || \leq \sum_{j=1}^{m} ||(\nabla f_j(\boldsymbol{p}) \bullet \boldsymbol{h})\, \boldsymbol{e_j} || = \sum_{j=1}^{m} |(\nabla f_j(\boldsymbol{p}) \bullet \boldsymbol{h})|$, since

$|| \boldsymbol{e}_j || = 1, \ 1 \leq j \leq n$.

Cauchy-Schwartz inequality for inner products says that $| \boldsymbol{u} \bullet \boldsymbol{v} | \leq || \boldsymbol{u} ||\, || \boldsymbol{v} ||$. Using this we get $||\boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{h})|| \leq \sum_{j=1}^{m} ||\nabla f_j(\boldsymbol{p})||\, ||\boldsymbol{h}|| = || \boldsymbol{h} || \sum_{j=1}^{m} ||\nabla f_j(\boldsymbol{p})||$.

If we take $M = \sum_{j=1}^{m} ||\nabla f_j(\mathbf{p})||$, then

$||\boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{h})|| \leq M\, || \boldsymbol{h} ||$. ....(6.1)

We have seen in Theorem 5.4 how to get the derivative of the sum of two differentiable functions, and also that of a scalar multiple of a differentiable function. The next theorem, which is known as the chain rule, tells us how to get the total derivative of a composite of two functions.

Theorem 6.1 (Chain Rule) : Let $\boldsymbol{f}$ and $\boldsymbol{g}$ be two differentiable functions, such that the composite function $\boldsymbol{f} \bullet \boldsymbol{g}$ is defined in a neighbourhood of a point $\boldsymbol{a} \in \boldsymbol{R}^n$. Suppose $\boldsymbol{g}$ is differentiable at $\boldsymbol{a}$, $\boldsymbol{g}(\boldsymbol{a}) = \boldsymbol{p}$, and $\boldsymbol{f}$ is differentiable at $\boldsymbol{p}$. Then $\boldsymbol{f} \bullet \boldsymbol{g}$ is differentiable at $\boldsymbol{a}$, and

$(\boldsymbol{f} \circ \boldsymbol{g})'(\boldsymbol{a}) = \boldsymbol{f}'(\boldsymbol{p}) \circ \boldsymbol{g}'(\boldsymbol{a}) = [\mathrm{D}\boldsymbol{f}(\boldsymbol{p})]\,[\mathrm{D}\boldsymbol{g}(\boldsymbol{a})]$.

Proof : If $\boldsymbol{h}$ *is such that* $\|\boldsymbol{h}\|$ *is small, then* $\boldsymbol{a} + \boldsymbol{h}$ *will belong to the above neighbourhood of* $\boldsymbol{a}$, *in which* $\boldsymbol{f} \bullet \mathbf{g}$ is defined. Now, since $\boldsymbol{g}$ is differentiable at $\boldsymbol{a}$,

$\boldsymbol{k} = \boldsymbol{g}(\boldsymbol{a} + \boldsymbol{h}) - \boldsymbol{g}(\boldsymbol{a}) = \boldsymbol{g}'(\boldsymbol{a})(\boldsymbol{h}) + \|\boldsymbol{h}\|\, \boldsymbol{E}_a(\boldsymbol{h})$, ...(6.2)

where $\boldsymbol{E}_a(\boldsymbol{h}) \to \boldsymbol{0}$, as $\boldsymbol{h} \to \boldsymbol{0}$.

$\boldsymbol{f}$ is differentiable at $\boldsymbol{p} = \boldsymbol{g}(\boldsymbol{a})$, and therefore,

$\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a} + \boldsymbol{h})) - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a})) = \boldsymbol{f}(\boldsymbol{p} + \boldsymbol{k}) - \boldsymbol{f}(\boldsymbol{p}) = \boldsymbol{f}'(\boldsymbol{p})(\boldsymbol{k}) + \|\boldsymbol{k}\|\, \boldsymbol{E}_p(\boldsymbol{k})$, where $\boldsymbol{E}_p(\boldsymbol{k}) \to \boldsymbol{0}$, as $\boldsymbol{k} \to \boldsymbol{0}$.

$= \boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))[\,\boldsymbol{g}(\boldsymbol{a} + \boldsymbol{h}) - \boldsymbol{g}(\boldsymbol{a})] + \|\boldsymbol{k}\|\, \boldsymbol{E}_p(\boldsymbol{k})$

$= \boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))[\,\boldsymbol{g}'(\boldsymbol{a})(\boldsymbol{h}) + \|\boldsymbol{h}\|\, \boldsymbol{E}_a(\boldsymbol{h})] + \|\boldsymbol{k}\|\, \boldsymbol{E}_p(\boldsymbol{k})$, using (6.2).

$= \boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))\, \boldsymbol{g}'(\boldsymbol{a})(\boldsymbol{h}) + \boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))\, [\|\boldsymbol{h}\|\, \boldsymbol{E}_a(\boldsymbol{h})] + \|\boldsymbol{k}\|\, \boldsymbol{E}_p(\boldsymbol{k})$, since $\mathbf{f}'(\mathbf{g}(\mathbf{a}))$ is a linear transformation. Thus, we can write

$\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a} + \boldsymbol{h})) - \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a})) = \boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))\, \boldsymbol{g}'(\boldsymbol{a})(\boldsymbol{h}) + \|\boldsymbol{h}\|[\,\boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))\, \boldsymbol{E}_a(\boldsymbol{h}) + \frac{\|\boldsymbol{k}\|}{\|\boldsymbol{h}\|} \boldsymbol{E}_p(\boldsymbol{k})]$, if $\boldsymbol{h} \neq \boldsymbol{0}$. ..(6.3)

To complete the proof we need to show that the vector in the square brackets in (6.3) tends to zero, as $\boldsymbol{h}$ tends to zero.

We know that $\boldsymbol{E}_a(\boldsymbol{h}) \to \mathbf{0}$, as $\boldsymbol{h} \to \mathbf{0}$. ....(*)

$\|\boldsymbol{k}\| = \|\boldsymbol{g}(\boldsymbol{a} + \boldsymbol{h}) - \boldsymbol{g}(\boldsymbol{a})\| \leq \|\boldsymbol{g}'(\boldsymbol{a})(\boldsymbol{h})\| + \|\boldsymbol{h}\|\, \|\boldsymbol{E}_a(\boldsymbol{h})\|$, using (6.2).

If $\mathrm{M} = \sum_{j=1}^{m} ||\nabla g_j(\boldsymbol{a})||$, then using (6.1), we can write $\|\mathbf{g}'(\mathbf{a})(\mathbf{h})\| \leq \mathrm{M}\, \|\boldsymbol{h}\|$. Thus,

$\|\boldsymbol{k}\| \leq \mathrm{M}\, \|\boldsymbol{h}\| + \|\boldsymbol{h}\|\, \|\boldsymbol{E}_a(\boldsymbol{h})\| = \|\boldsymbol{h}\|\, (\mathrm{M} + \|\boldsymbol{E}_a(\boldsymbol{h})\|)$. Therefore,

$\frac{||\boldsymbol{k}||}{||\boldsymbol{h}||} \leq \mathrm{M} + \|\boldsymbol{E}_a(\boldsymbol{h})\|$. This means that $\frac{||\boldsymbol{k}||}{||\boldsymbol{h}||}$ is bounded. Thus,

$\frac{||\boldsymbol{k}||}{||\boldsymbol{h}||} \boldsymbol{E}_p(\boldsymbol{k}) \to \boldsymbol{0}$, as $\boldsymbol{h} \to \boldsymbol{0}$, since $\boldsymbol{h} \to \boldsymbol{0} \Rightarrow \boldsymbol{k} \to \mathbf{0} \Rightarrow \boldsymbol{E}_p(\boldsymbol{k}) \to \mathbf{0}$. ....(**)

Using (*) and (**), we can say that the term in the square brackets in (6.3) tends to zero as

$\boldsymbol{h} \to \mathbf{0}$. Therefore,

$$\frac{\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a}+\boldsymbol{h}))-\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a}))-\boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a}))\,\boldsymbol{g}'(\boldsymbol{a})(\boldsymbol{h})}{||\boldsymbol{h}||} \to \boldsymbol{0} \text{ as } \boldsymbol{h} \to \boldsymbol{0}.$$

This shows that $\boldsymbol{f} \bullet \boldsymbol{g}$ is differentiable at $\boldsymbol{a}$, and $(\boldsymbol{f} \bullet \boldsymbol{g})'(\boldsymbol{a}) = \boldsymbol{f}'(\boldsymbol{g}(\boldsymbol{a})) \circ \boldsymbol{g}'(\boldsymbol{a})$.

The Chain Rule can be written in terms of Jacobian matrices as follows:

$\mathrm{D}(\boldsymbol{f} \bullet \boldsymbol{g})(\boldsymbol{a}) = [\mathrm{D}(\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a})))]\,[\mathrm{D}(\boldsymbol{g}(\boldsymbol{a}))]$.

Here the product on the right hand side is matrix multiplication. If $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$, and $\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{y})$, comparing the entries in the matrices in (6.3), we get

$$\frac{\partial z_i}{\partial x_k} = \sum_{j=1}^{n} \frac{\partial z_i}{\partial y_j}\frac{\partial y_j}{\partial x_k}, \qquad \text{where } \frac{\partial z_i}{\partial x_k} = \mathrm{D}_k(\boldsymbol{f} \bullet \boldsymbol{g})_i,\ \frac{\partial z_i}{\partial y_j} = \mathrm{D}_j(\boldsymbol{f})_i,\text{ and } \frac{\partial y_j}{\partial x_k} = \mathrm{D}_k(\boldsymbol{g})_j.$$

Example 6.1 : Write the matrices for $\boldsymbol{f}'$, $\boldsymbol{g}'$ and $(\boldsymbol{f} \bullet \boldsymbol{g})'$ for the following functions, and evaluate them at the point (2, 5). $\boldsymbol{f}(x, y) = (x + y, x^2 + y^2, 2x + 3y)$, $\boldsymbol{g}(u, v) = (x, y) = (u^2, v^3)$.

Solution : Here $f_1(x, y) = x + y$, $f_2(x, y) = x^2 + y^2$, $f_3(x, y) = 2x + 3y$,

$g_1(u, v) = u^2$ and $g_2(u, v) = v^3$. This means, $\mathrm{D}(\boldsymbol{f}) = \begin{pmatrix} 1 & 1 \\ 2x & 2y \\ 2 & 3 \end{pmatrix}$, and $\mathrm{D}(\boldsymbol{g}) = \begin{pmatrix} 2u & 0 \\ 0 & 3v^2 \end{pmatrix}$.

$(\boldsymbol{f} \bullet \boldsymbol{g})(u,v) = (u^2 + v^3, u^4 + v^6, 2u^2 + 3v^3)$. Hence,

$$\mathrm{D}(\boldsymbol{f} \bullet \boldsymbol{g}) = \begin{pmatrix} 2u & 3v^2 \\ 4u^3 & 6v^5 \\ 4u & 9v^2 \end{pmatrix}.$$

At $(u, v) = (2, 5)$, $(x, y) = (4, 125)$. Therefore,

, $\mathrm{D}(\boldsymbol{f})(4, 125) = \begin{pmatrix} 1 & 1 \\ 8 & 250 \\ 2 & 3 \end{pmatrix}$, $\mathrm{D}(\boldsymbol{g})(2, 5) = \begin{pmatrix} 4 & 0 \\ 0 & 75 \end{pmatrix}$, and $\mathrm{D}(\boldsymbol{f} \bullet \boldsymbol{g})(2, 5) = \begin{pmatrix} 4 & 75 \\ 32 & 18750 \\ 8 & 225 \end{pmatrix}$.

You can now easily verify that $\mathrm{D}(\boldsymbol{f} \bullet \boldsymbol{g})(2, 5) = [\mathrm{D}(\boldsymbol{f}(4, 125))]\,[\mathrm{D}(\boldsymbol{g}(2, 5))]$.

6.3 HIGHER ORDER PARTIAL DERIVATIVES

You are familiar with the concept of partial derivatives. In the last chapter we have calculated the partial derivatives of some functions of n variables. If you take a look at those examples, you will realise that the partial derivatives are themselves functions of n variables. So, we can talk about their partial derivatives. These, if they exist, will be the second order partial derivatives of the original function. If we differentiate these again, we will get the third order partial derivatives of the original function, and so on. We take a simple example to illustrate.

Example 6.2 : Find partial derivatives of all possible orders for the function,

$\mathbf{f}(x, y, z) = (x^2y^2, 3xy^3z, xz^3)$.

Solution : Since $\mathbf{f}$ is a polynomial function, we do not have to worry about the existence of partial derivatives. We get

$\mathbf{f}_x = (2xy^2, 3y^3z, z^3)$, $\mathbf{f}_y = (2x^2y, 9xy^2z, 0)$, $\mathbf{f}_z = (0, 3xy^3, 3xz^2)$.

Then, $\mathbf{f}_{xx} = \frac{\partial^2 f}{\partial x^2} = (2y^2, 0, 0)$, $\mathbf{f}_{xy} = \frac{\partial}{\partial y}(\frac{\partial \mathbf{f}}{\partial x}) = \frac{\partial^2 \mathbf{f}}{\partial y \partial x} = (4xy, 9y^2z, 0)$, $\mathbf{f}_{xz} = (0, 3y^3, 3z^2)$.

Differentiating $\mathbf{f}_y$, we get $\mathbf{f}_{yx} = (4xy, 9y^2, 0)$, $\mathbf{f}_{yy} = (2x^2, 18xyz, 0)$, and $\mathbf{f}_{yz} = (0, 9xy^2, 0)$.

Then differentiating $\mathbf{f}_z$ we get $\mathbf{f}_{zx} = (0, 3y^3, 3z^2)$, $\mathbf{f}_{zy} = (0, 9xy^2, 0)$, and $\mathbf{f}_{zz} = (0, 0, 6xz)$.

These are all possible second order derivatives of **f**. Proceeding in this way, we can also get

$\mathbf{f}_{xyz} = (0, 9y^2, 0)$, $\mathbf{f}_{yxz} = (0, 0, 0)$, $\mathbf{f}_{zzz} = (0, 0, 6x)$, and so on. There will be 27 third order partial derivatives of $\mathbf{f}$. See if you can get the remaining.

You know that $\mathbf{f}_{xy}$ and $\mathbf{f}_{yx}$ differ in the order in which $\mathbf{f}$ is differentiated with respect to the variables x and y. These two derivatives have come out to be equal in Example 6.2. But you may have seen examples of scalar functions of several variables, for which the two may not be the same. Here is an example, to jog your memory.

Example 6.3 : Consider this function f from $\mathbf{R}^2$ to $\mathbf{R}$, $f(x, y) = \frac{xy(x^2-y^2)}{x^2+y^2}$ for $(x, y) \neq (0, 0)$, and $f(0, 0) = 0$. You can easily check that

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

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

$$\text{Then, } f_{xy}(0, 0) = \lim_{k\to 0} \frac{f_x(0,k)-f_x(0,0)}{k} = \lim_{k\to 0} \frac{-k-0}{k} = -1, \text{ and similarly, } f_{yx}(0, 0) = 1.$$

Thus, the mixed partial derivatives of this function both exist, but are not equal.

Remark 6.1 : If f is a function from $\mathbf{R}^n$ to $\mathbf{R}$, the partial derivative of f with respect to the ith variable, x_i, is denoted by $D_i f$, and the partial derivative of $D_i f$ with respect to x_j, that is, $D_j(D_i f)$ is denoted by $D_{ji} f$.

The following theorem gives a sufficient condition for the two mixed partial derivatives of a function to be equal. Since the behaviour of a vector-valued function is decided by the behaviour of its coordinate functions, it is enough to derive this sufficient condition for a scalar function. Without loss of generality, we state the theorem for a function of two variables.

Theorem 6.2 : Let $f : \boldsymbol{R}^2 \rightarrow \boldsymbol{R}$, such that the partial derivatives, D_1f, D_2f, $D_{12}f$ and $D_{21}f$ exist on an open set S in $\boldsymbol{R}^2$. If $(a, b) \in$ S, and $D_{12}f$ and $D_{21}f$ are both continuous at (a, b), then $D_{12}f(a, b) = D_{21}f(a, b)$.

Proof : We choose positive real numbers, h and k, which are small enough so that the rectangle with vertices $(a, b), (a + h, b), (a, b + k), (a + h, b + k)$ lies within S.

Now we consider a function

$\Delta(h, k) = [f(a + h, b + k) - f(a + h, b)] - [f(a, b + k) - f(a, b)]$.

We also define a function G on $[a, a + h]$, $G(x) = f(x, b + k) - f(x, b)$.

Now we can write $\Delta(h, k) = G(a + h) - G(a)$. Since G is defined in terms of f, and since f has all the necessary properties, G is continuous on $[a, a + h]$, and is differentiable in $(a, a + h)$. So, we apply the Mean Value Theorem for functions of a single variable to G, and get

$G(a + h) - G(a) = hG'(c)$, for some $c \in (a, a + h)$. Now $G'(x) = D_1f(x, b + k) - D_1f(x, b)$. So, we write $\Delta(h, k) = G(a + h) - G(a) = h[D_1f(c, b + k) - D_1f(c, b)]$.

Now $D_1f (c, y)$ is a differentiable function of one variable with derivative equal to $D_{21}f$. So applying MVT to $D_1f(c, y)$ on the interval $[b, b + k]$, we get

$$\Delta(h, k) = h[D_1f(c, b + k) - D_1f(c, b)] = hkD_{21}f(c, d), \qquad \text{..........................(6.4)}$$

for some $d \in (b, b + k)$.

We now write $\Delta(h, k) = [f(a + h, b + k) - f(a, b + k)] - [f(a + h, b) - f(a, b)]$, and define

$H(y) = f(a + h, y) - f(a, y)$, so that $\Delta(h, k) = H(b + k) - H(b)$. Using the same arguments that we used for G, we apply MVT to H, and then to $D_2f(x, p)$, we get

$$\Delta(h, k) = k[D_2f(a + h, p) - D_2f(a, p)] = khD_{12}f(q, p), \qquad \text{............................(6.5)}$$

for some $p \in (b, b + k)$, and $q \in (a, a + h)$.

From (6.4) and (6.5) we get $D_{21}f(c, d) = D_{12}f(q, p)$. Since $D_{12}f$ and $D_{21}f$ are continuous, taking the limit as $(h, k) \rightarrow (0,0)$, we get $D_{12}f(a, b) = D_{21}f(a, b)$.

As we have mentioned earlier, the conditions of this theorem are sufficient, and not necessary. In fact, the continuity of just one of the mixed partial derivatives is also sufficient to guarantee equality. Functions whose partial derivatives are continuous play an important role in Calculus. We classify these functions as follows:

Definition 6.1 : A function $\boldsymbol{f}$ from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$ is said to be **continuously differentiable**, or belong to class C^1, if all its partial derivatives $D_i\boldsymbol{f}$ are continuous. It is said to belong to class C'', if all its second order partial derivatives are continuous, and so on. If all its partial derivatives of all orders are continuous, then it is said to belong to class C^∞.

We have proved that a function in class C^1 is differentiable in Theorem 5.8. In Theorem 6.2 we have seen that the mixed partial derivatives of a function belonging to class C'' are equal.

In the next chapter we shall see that a C^k function, that is a function, all whose partial derivatives of order up to k are continuous, can be approximated by means of a polynomial of order k. We shall also discuss the technique to find the maximum and minimum values of a function belonging to class C''.

6.4 MEAN VALUE THEOREM

The Mean Value Theorem (MVT) is an important theorem in Calculus. It is used as a tool to derive many other results. In the last section we have used it in the proof of Theorem 6.2. In this section we shall see if it also holds good for functions of several variables. But first, let us recall the one-variable case.

MVT (single variable): If $f : [a, b] \to \boldsymbol{R}$ is continuous on $[a, b]$, and differentiable on (a, b), then there exists $c \in (a, b)$, such that

$f(b) - f(a) = (b - a)\, f'(c).$

If we write $b = a + h$, then there exists $\theta,\ 0 < \theta < 1$, such that

$f(a + h) - f(a) = h\, f'(a + \theta h).$

Unfortunately, it is not possible to extend this theorem to a function $\boldsymbol{f} : \boldsymbol{R}^n \to \boldsymbol{R}^m$, when

$m > 1$. This will be quite clear from the following example.

Example 6.4 : Consider $f : [0, 2\pi] \to \boldsymbol{R}^2$, $f(t) = (\cos t, \sin t)$. This function is continuous on

$[0, 2\pi]$ and differentiable on $(0, 2\pi)$. Now, $f(2\pi) - f(0) = (1, 0) - (1, 0) = (0, 0)$.

$f'(t) = (-\sin t, \cos t)$. For the extension of MVT to hold, we must have

$f(2\pi) - f(0) = 2\pi f'(c)$ for some c in $(0, 2\pi)$. So, we should have $(0, 0) = 2\pi(-\sin c, \cos c)$. But this is impossible, since $\sin c$ and $\cos c$ both cannot be zero.

So, the extension of MVT in its stated form does not hold. But there is a way around this difficulty. A slightly modified version of MVT does hold true for all functions of several variables. We now state and prove this modified theorem for functions from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$. As a special case of this theorem you will realize that MVT holds for real-valued functions of several variables.

Theorem 6.3 : (Mean Value Theorem) Let $\boldsymbol{f} : S \to \boldsymbol{R}^m$, where S is an open subset of $\boldsymbol{R}^n$. Suppose $\boldsymbol{f}$ is differentiable on S. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two points in S, such that the line segment joining $\boldsymbol{x}$ and $\boldsymbol{y}$, $L(\boldsymbol{x}, \boldsymbol{y}) = \{t\boldsymbol{x} + (1 - t)\boldsymbol{y} \mid 0 \le t \le 1\}$, also lies in S. Then for every $\boldsymbol{a} \in \boldsymbol{R}^m$, there is a point $\boldsymbol{z} \in S$, such that

$\boldsymbol{a} \bullet \{f(y) - f(x)\} = \boldsymbol{a} \bullet \{f'(z)(y - x)\}$ ....(6.6)

Before we start the proof, let us understand the geometry involved. Let $\boldsymbol{u} = \boldsymbol{y} - \boldsymbol{x}$. Then $\boldsymbol{x} + t\boldsymbol{u}$ gives us a point on the line segment L($\boldsymbol{x}, \boldsymbol{y}$), if $0 \leq t \leq 1$. Since S is open, we can find a

$\delta > 0$, such that $\overline{\mathrm{B}(\boldsymbol{x}, \delta)} \subset$ S, and $\overline{\mathrm{B}(\boldsymbol{y}, \delta)} \subset$ S. See Fig. 6.1, in which we show the situation

when n = 2. The point $\boldsymbol{p}$ is on the extension of L($\boldsymbol{x}, \boldsymbol{y}$) and is equal to $\boldsymbol{x} + (1 + \beta)\boldsymbol{u}$. Similarly the point $\boldsymbol{q}$ is also on the extension of L($\boldsymbol{x}, \boldsymbol{y}$), and is equal to $\boldsymbol{x} - \beta\boldsymbol{u}$ for some $\beta > 0$.

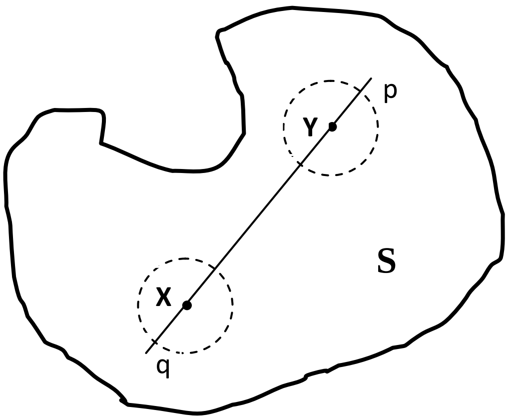


Figure 6.1

Thus we get a $\beta > 0$, such that $\boldsymbol{x} + t\boldsymbol{u} \in$ S for every $t \in (-\beta, 1 + \beta)$. Now we start the formal proof.

Proof : Let $\boldsymbol{a} \in \boldsymbol{R}^n$. We define a function $F : (-\beta, 1 + \beta) \rightarrow \boldsymbol{R}$, $F(t) = \boldsymbol{a} \bullet \boldsymbol{f}(\boldsymbol{x} + t\mathbf{u})$. This F is a differentiable function on $(-\beta, 1 + \beta)$, and

$F'(t) = \boldsymbol{a} \bullet \{\boldsymbol{f}'(\boldsymbol{x} + t\boldsymbol{u})(\boldsymbol{u})\}$, using chain rule.

(Recall, that $\boldsymbol{f}'(\boldsymbol{x} + t\boldsymbol{u})$ is a linear transformation.)

Thus, we can apply MVT for functions of a single variable, and get

$F(1) - F(0) = F'(\theta)$, for some $\theta, 0 < \theta < 1$. ....(6.7)

Now, $F(1) = \boldsymbol{a} \bullet \boldsymbol{f}(\boldsymbol{x} + \boldsymbol{u}) = \boldsymbol{a} \bullet \boldsymbol{f}(y)$, $F(0) = \boldsymbol{a} \bullet \boldsymbol{f}(\boldsymbol{x})$, and

$F'(\theta) = \boldsymbol{a} \bullet \{\boldsymbol{f}'(\boldsymbol{x} + \theta\boldsymbol{u})(\boldsymbol{u})\} = \boldsymbol{a} \bullet \{\boldsymbol{f}'(\boldsymbol{z})(\boldsymbol{y} - \boldsymbol{x})\}$, where $\boldsymbol{z} = \boldsymbol{x} + \theta\boldsymbol{u} \in L(x, y)$.

Therefore, from (6.7) we get $\boldsymbol{a} \bullet \{f(y) - f(x)\} = \boldsymbol{a} \bullet \{\boldsymbol{f}'(\boldsymbol{z})(\boldsymbol{y} - \boldsymbol{x})\}$ for some $\boldsymbol{z} \in$ S.

Remark 6.2 : i) (6.6) is true for all $\boldsymbol{x}, \boldsymbol{y}$ in S, such that the line segment joining $\boldsymbol{x}$ and $\boldsymbol{y}$ is also in S. This means, if S is a convex open set in $\boldsymbol{R}^n$, then (6.6) will be true for all $\boldsymbol{x}, \boldsymbol{y}$ in S.

ii) If f is a real-valued function, then m = 1, and $a \in \boldsymbol{R}$. Then for $a = 1$ we have

$1 . \{f(y) - f(x)\} = 1 . \{f'(\boldsymbol{z})(\boldsymbol{y} - \boldsymbol{x})\} = \nabla f(z) \bullet (y - x)$, for some $\boldsymbol{z} \in$ S.

So, the MVT for functions of a single variable extends directly to real-valued functions of several variables. We can also directly prove MVT for scalar functions. The proof runs exactly similar to that of Theorem 6.3, if we put $a = 1$.

The MVT has a well-known consequence, which we now state:

Theorem 6.4 : Let $\boldsymbol{f} : S \to \boldsymbol{R^m}$, where S is an open connected subset of $\boldsymbol{R^n}$. Suppose $\boldsymbol{f}$ is differentiable on S, and $\boldsymbol{f}'(\boldsymbol{p}) = \boldsymbol{0}$ for every $\boldsymbol{p} \in S$. Then $\boldsymbol{f}$ is a constant function on S.

Proof : The set S is polygonally connected, since it is open and connected. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two points in S. Then $\boldsymbol{x}$ and $\boldsymbol{y}$ are joined by line segments $L_1, L_2, L_3, \dots, L_r$, lying entirely in S. Suppose L_i is a line segment joining $\boldsymbol{p_i}$ and $\boldsymbol{p_{i+1}}$, $1 \le i \le r$, $\boldsymbol{p_1} = \boldsymbol{x}$, and $\boldsymbol{p_{r+1}} = \boldsymbol{y}$.

Let $\boldsymbol{a} \in \boldsymbol{R^m}$. Then using Theorem 6.3, we have

$$\boldsymbol{a} \bullet \{\boldsymbol{f}(\boldsymbol{p_{i+1}}) - \boldsymbol{f}(\boldsymbol{p_i})\} = \boldsymbol{a} \bullet \{\boldsymbol{f}'(\boldsymbol{z_i})(\boldsymbol{p_{i+1}} - \boldsymbol{p_i})\}, \ \boldsymbol{z_i} \in L_i$$

$$= 0, \text{ since } \boldsymbol{f}'(\boldsymbol{z_i}) = \boldsymbol{0}.$$

This means,

$$\boldsymbol{a} \bullet \{\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})\} = \boldsymbol{a} \bullet \{\boldsymbol{f}(\boldsymbol{p_{r+1}}) - \boldsymbol{f}(\boldsymbol{p_1})\} = \sum_{i=1}^{r} \boldsymbol{a} \bullet \{\boldsymbol{f}(\boldsymbol{p_{i+1}}) - \boldsymbol{f}(\boldsymbol{p_i})\} = 0. \quad \dots\dots\dots\dots(6.8)$$

(6.8) is true for every $\boldsymbol{a}$ in $\boldsymbol{R^m}$. So, in particular, it is true for $\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})$. Thus,

$$\{\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})\} \bullet \{\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})\} = ||\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})||^2 = 0.$$

So, $\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$, or $\boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x})$.

Since $\boldsymbol{x}$ and $\boldsymbol{y}$ were any arbitrary points in S, we have thus proved that $\boldsymbol{f}$ is a constant function on S.

Try a few exercises now.

Exercises :

1) Find the partial derivatives, D_1f, D_2f, $D_{12}f$ and $D_{21}f$ at (0, 0) , if they exist, for the following function f from $\boldsymbol{R^2}$ to $\boldsymbol{R}$.
   $f(x, y) = y\frac{x^2-y^2}{x^2+y^2}$, if $(x, y) \neq (0, 0)$, and $f(0, 0) = 0$.
2) If $u(x, y) = x + y^2$, $x(t) = 3t^2 + 4$, and $y(t) = \sin 2t$, find $u'(t)$ and $u''(t)$.
3) If $u(x, y) = x - 2y + 3$, $x = r + s + t$, $y = rs + t^2$, find u_r, u_s and u_t at (1, 2, 4).
4) Let $\boldsymbol{f} : \boldsymbol{R^2} \to \boldsymbol{R^2}$, and $\boldsymbol{g} : \boldsymbol{R^3} \to \boldsymbol{R^2}$ be two vector functions, defined as:
   $\boldsymbol{f}(x, y) = (\sin(2x + y), \cos(x + 2y))$,
   $\boldsymbol{g}(r, s, t) = (2r - s - 3t, r^2 - 3st)$.
   i) Write the Jacobian matrices for $\boldsymbol{f}$ and $\boldsymbol{g}$. If $\boldsymbol{h}$ is the composite function, $\boldsymbol{f} \bullet \boldsymbol{g}$, compute the Jacobian matrix of $\boldsymbol{h}$ at the point (1, 0, - 2).
5) If f is a function from $\boldsymbol{R^2}$ to $\boldsymbol{R}$, and $D_1f = 0$ at all points, show that f is independent of the first variable. If $D_1f = D_2f = 0$ at all points, show that f is a constant function.

6.5 SUMMARY

In this chapter we have derived the chain rule for differentiation of composite of two functions. We have also seen that the Jacobian matrix for the composite function is the product of the Jacobian matrices of the two given functions. We have defined higher order partial derivatives of functions of several variables. We have seen functions, whose second order mixed partial derivatives depend on the order of the variables with respect to which the function is differentiated. On the other hand, we have derived sufficient conditions for such mixed partial derivatives to be equal. Finally, through an example we have seen that the Mean Value Theorem cannot be extended to all vector functions. We have proved a restricted form of the MVT for vector functions. Of course, MVT does extend to scalar-valued functions of several variables. As a result of MVT we have proved that a function defined on an open connected set is constant, if its derivative is uniformly zero over its domain.

7

APPLICATIONS OF DERIVATIVES

Unit Structure

7.0 OBJECTIVES

After reading this chapter, you should be able to

- state Taylor's theorem for real-valued functions of several variables
- obtain Taylor's expansions for some simple functions
- define, locate and classify extreme points of a function of several variables
- obtain the extreme values of a function of n variables, subject to some constraints

7.1 INTRODUCTION

In the two previous chapters we have discussed differentiation of scalar and vector functions of several variables. Now we shall tell you about some applications of derivatives. In your study of functions of one variable you have seen that a major application of the concept of derivatives is the location of maxima and minima of a function. This knowledge is very crucial for curve tracing. Here we shall see how the derivatives help us in locating the extreme values of a real-valued function of several variables. But before we do that, we are going to discuss Taylor's theorem and Taylor's expansions, which help us approximate a function with the help of polynomials. This knowledge will help us derive some tests for locating and classifying the extreme points of a function.

7.2 TAYLOR'S THEOREM

It will be useful to recall Taylor's theorem for functions of one variable, which you have studied in F. Y. B. Sc. Here we shall also give you the proof of this theorem. Our method of proof involves the use of Rolle's theorem. You have studied this theorem too in F. Y. We now state Rolle's theorem, and then move on to Taylor's theorem.

Theorem 7.1 (Rolle's Theorem): If f: $[a, b] \rightarrow \boldsymbol{R}$ is continuous on $[a, b]$, differentiable on

(a, b), and $f(a) = f(b)$, then there exists $c \in (a, b)$, such that $f'(c) = 0$.

Theorem 7.2 (Taylor's theorem for real functions of one variable): Let f be a real-valued function defined on the open interval (p, q). Suppose f has derivatives of all orders up to and including n +1 in (p, q). Let a be any point in (p, q). Then for any x ∈ (p, q),

$$f(x) = f(a) + \frac{x-a}{1!} f'(a) + \frac{(x-a)^2}{2!} f''(a) + \ldots + \frac{(x-a)^n}{n!} f^{(n)}(a) + \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(c), \ldots (7.1)$$

where $c \in (a, b)$.

Proof: We now define a new function g on $[a, x]$, or $[x, a]$, according as $a < x$, or $x < a$, by

$$g(y) = f(y) + \frac{(x-y)}{1!} f'(y) + \frac{(x-y)^2}{2!} f''(y) + \ldots + \frac{(x-y)^n}{n!} f^{(n)}(y) + (x-y)^{n+1} A, \quad \ldots (7.2)$$

where A is a constant, chosen so as to satisfy $g(x) = g(a)$. We can easily write the expression for A by using this condition. We leave this to you as an exercise. See Exercise 1).

Using the properties of f, we can see that g satisfies all the conditions of Rolle's theorem on its domain. Thus, we can conclude that there exists a point $c \in (a, x)$, (or (x, a)) such that $g'(c) = 0$. Now, differentiating (7.2), we see that

$$g'(y) = f'(y) - f'(y) + (x-y) f''(y) - (x-y) f''(y) + \frac{(x-y)^2}{2!} f'''(y) - \ldots - \frac{(x-y)^{(n-1)}}{(n-1)!}$$

$$f^{(n)}(y) + \frac{(x-y)^n}{n!} f^{(n+1)}(y) - (n+1)(x-y)^n A.$$

$$= (x-y)^n \left[\frac{f^{(n+1)}(y)}{n!} - (n+1)A \right].$$

Hence, $g'(c) = (x-c)^n \left[\frac{f^{(n+1)}(c)}{n!} - (n+1)A \right] = 0.$

This means that $A = \frac{f^{(n+1)}(c)}{n!}$

Substituting this value of A in (7.2), we get

$f(x) = g(x) =$

$$g(a) = f(a) + \frac{x-a}{1!} f'(a) + \frac{(x-a)^2}{2!} f''(a) + \ldots + \frac{(x-a)^n}{n!} f^{(n)}(a) + \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(c),$$

thus proving the theorem.

Remark 7.1 : If the function in Theorem 7.2 has derivatives of all orders in (p, q), then we can write a Taylor expansion as in (7.1) for any $n \in \boldsymbol{N}$. Further, if all the derivatives of all orders are bounded by a positive number M, that is, if $\left|\frac{d^n f}{dx^n}\right| < M$ for all n, and at all points in (p, q), then $\left|\frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(c)\right| \leq \left|\frac{(x-a)^{n+1}}{(n+1)!}\right| \to 0$ as $n \to \infty$ for every x in some interval

$\{x: |x - a| < R\}$. Therefore, in this case we can write

$$f(x) = f(a) + \frac{x-a}{1!} f'(a) + \frac{(x-a)^2}{2!} f''(a) + \ldots + \frac{(x-a)^n}{n!} f^{(n)}(a) + \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(c), \ldots (7.3)$$

The infinite series in (7.3) is convergent under the given conditions, and is called the Taylor series of f about a.

Now, (7.1) can be written as $f(x) = P_n(x) + R_n(x)$, where

$P_n(x) = f(a) + \frac{x-a}{1!} f'(\mathrm{a}) + \frac{(x-a)^2}{2!} f''(\mathrm{a}) + \ldots + \frac{(x-a)^n}{n!} f^{(\mathrm{n})}(\mathrm{a})$ is called the **n^th^ Taylor polynomial of f about a**, and $R_n(x) = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(c)$, is called the **remainder.**

We now state Taylor's theorem for functions of two variables, and then find Taylor expansions of some functions.

Theorem 7.3 (Taylor's theorem for $f: \boldsymbol{R}^2 \to \boldsymbol{R}$)**:** Let f be a real-valued C^{n+1} function on an open convex set $\mathrm{E} \subseteq \boldsymbol{R}^2$. Let $(a, b) \in \mathrm{E}$. Then for any $(x, y) \in \mathrm{E}$,

$$f(x, y) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a, b) + \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \ldots + \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a, b)$$
$$+ \frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n+1} f(c, d), \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7.4)$$

where $h = x - a$, $k = y - b$, and (c, d) is some point on the line segment joining (a, b) and (x, y).

We are not going to prove this theorem. But, note the following points:

1. Recall that f is C^{n+1} means f has continuous partial derivatives of all orders $\leq n + 1$. This ensures that all the relevant mixed partial derivatives are equal.
2. E is convex. This guarantees that the line segment joining any two points of E, lies in E, the domain of f.

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

where $h = x - a$, and $k = y - b$, is called the **n^th^ Taylor polynomial**, and

$R_n(x, y) = \frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n+1} f(c, d)$ is called the **remainder of order n.**

Let us use this theorem to get the expansions of some functions.

Example 7.1: Find the Taylor expansions of the following functions about the given points up to the third order.

i) $f(x, y) = x^3 + 2xy^2 - 3xy + 4x + 5$, $(a, b) = (1, 2)$

ii) $f(x, y) = \sin(2x + 3y)$ $(a, b) = (0, 0)$.

Solution: i) Since $f(x,y) = x^3 + 2xy^2 - 3xy + 4x + 5$ is a polynomial, it has partial derivatives of all orders. Further, its partial derivatives of order > 3 are all zero. In fact,

$f_x = 3x^2 + 2y^2 - 3y + 4$, $f_y = 4xy - 3x$, $f_{xx} = 6x$, $f_{xy} = 4y - 3$, $f_{yy} = 4x$, $f_{xxx} = 6$, $f_{xxy} = 0$,

$f_{xyy} = 4$, $f_{yyy} = 0$, and all higher partial derivatives are zero. Calculating all these partial derivatives at (1, 2), we write

$$f(1 + h, 2 + k) = 12 + 9h + 5k + \frac{1}{2!}(6h^2 + 10hk + 4k^2) + \frac{1}{3!}(6h^3 + 12hk^2) + R_3.$$

Now, R_3 involves all fourth order derivatives, and therefore is zero. Hence,

$$f(1 + h, 2 + k) = 12 + 9h + 5k + \frac{1}{2!}(6h^2 + 10hk + 4k^2) + \frac{1}{3!}(6h^3 + 12hk^2).$$

ii) $f(x, y) = \sin(2x + 3y)$ also has derivatives of all orders.
$f_x = 2\cos(2x + 3y) = 2$ at (0, 0), $f_y = 3\cos(2x + 3y) = 3$ at (0, 0),
$f_{xx} = -4\sin(2x + 3y)$, $f_{xy} = -6\sin(2x + 3y)$, $f_{yy} = -9\sin(2x + 3y)$. These second order derivatives are all zero at (0, 0).

$f_{xxx} = -8\cos(2x + 3y)$, $f_{xxy} = -12\cos(2x + 3y)$, $f_{xyy} = -18\cos(2x + 3y)$,

$f_{yyy} = -27\cos(2x + 3y)$.

These are, respectively, – 8, – 12, – 18, and – 27 at (0, 0). Thus,

$f(h, k) = 0 + (2h + 3k) + \frac{1}{2!}.0 + \frac{1}{3!}(-8h^3 - 3.12h^2k - 3.18hk^2 - 27h^3) + R_3$, where

$R_3 = \frac{1}{4!}(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^4\sin(2c + 3d)$, where (c, d) is some point on the line segment joining (0, 0) and (h, k).

We are now going to state Taylor's theorem for real-valued functions of n variables. For this, let us first take a close look at the Taylor expansion of a function of two variables.

If we write (x, y) as $(a + h, b + k)$, we get

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

If we take the variables to x_1, x_2, instead of x and y, take (a, b) to be (a_1, a_2), and (h, k) to be

$f(a_1 + h, a_2 + h_2) = f(a_1, a_2) + (h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}) f(a_1, a_2) + \frac{1}{2!}\left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}\right)^2 f(a_1, a_2) + ...$

$+ \frac{1}{n!}\left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}\right)^n f(a_1, a_2) + \frac{1}{(n+1)!}\left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}\right)^2 f(c, d),$

$= \sum_{k=0}^{n} \frac{1}{k!}\left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}\right)^k f(a_1, a_2) + R_n(c, d),$

$= \sum_{k=0}^{n} \frac{1}{k!} \sum D_{i_1 i_2 \ldots i_k} f(a_1, a_2) h_{i_1} h_{i_2} \ldots h_{i_k} + R_n(c, d),$

where $D_{i_1 i_2 \ldots i_k} = \frac{\partial^k}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}$, and $i_1, i_2, \ldots, i_k$ = 1 or 2, and the sum is taken over all ordered k-tuples $(i_1, i_2, \ldots, i_k)$. For example,

$\sum D_{i_1 i_2} f(a_1, a_2) h_{i_1} h_{i_2} = D_{11} f(a_1, a_2) h_1^2 + D_{12} f(a_1, a_2) h_1 h_2 + D_{21} f(a_1, a_2) h_2 h_1 + D_{22} f(a_1, a_2) h_2^2$

$$= (h_1^2 \frac{\partial^2}{\partial x_1^2} + 2 h_1 h_2 \frac{\partial^2}{\partial x_1 \partial x_2} + h_2^2 \frac{\partial^2}{\partial x_2^2}) f(a_1, a_2) .$$

Similarly,

$\sum D_{i_1 i_2 i_3} f(a_1, a_2) h_{i_1} h_{i_2} h_{i_3} = D_{111} f(a_1, a_2) h_1^3 + D_{112} f(a_1, a_2) h_1^2 h_2 + D_{121} f(a_1, a_2) h_1 h_2 h_1 + D_{211} f(a_1, a_2) h_2 h_1^2 + D_{122} f(a_1, a_2) h_1 h_2^2 + D_{212} f(a_1, a_2) h_2 h_1 h_2 + D_{221} f(a_1, a_2) h_2^2 h_1$

$+ D_{222} f(a_1, a_2) h_2^3$

$$= (h_1^3 \frac{\partial^3}{\partial x_1^3} + 3 h_1^2 h_2 \frac{\partial^3}{\partial x_1^2 \partial x_2} + 3 h_1 h_2^2 \frac{\partial^3}{\partial x_1 \partial x_2^2} + h_2^3 \frac{\partial^3}{\partial x_2^3}) f(a_1, a_2) .$$

You must have noticed that we have added the mixed partial derivative terms, for example, $D_{12}f$ and $D_{21}f$, or $D_{112}f$, $D_{121}f$, and $D_{211}f$. We could do this, since $f \in C^\infty$ ensures that that these partial derivatives are equal. Now we state Taylor's theorem for real-valued functions of several variables.

Theorem 7.4 : Let $f : E \to \boldsymbol{R}$, where E is a convex open subset of $\boldsymbol{R}^n$. Further, let

$\boldsymbol{a} = (a_1, a_2, ..., a_n) \in E$, $\boldsymbol{h} = (h_1, h_2, ..., h_n) \in \boldsymbol{R}^n$, such that $\boldsymbol{a} + \boldsymbol{h} \in D$. If $f \in C^m$, then

$f(\boldsymbol{a} + \boldsymbol{h}) = \sum_{k=0}^{m-1} \frac{1}{k!} \sum D_{i_1 i_2 \ldots i_k} f(\boldsymbol{a}) h_{i_1} h_{i_2} \ldots h_{i_k} + R_{m-1}(\boldsymbol{c}),$ ....(7.5)

where $i_1, i_2, \ldots, i_k$ take values from the set {1, 2, ..., n}, and the inner summation in (7.5) is taken over all possible such k-tuples.

Further, the remainder $R_{m-1}(\boldsymbol{c}) = \frac{1}{m!} \sum D_{i_1 i_2 \ldots i_m} f(\boldsymbol{c}) h_{i_1} h_{i_2} \ldots h_{i_m}$. This sum is taken over all possible m-tuples $(i_1, i_2, ..., i_m)$, where $i_1, i_2, ..., i_m$ take values from {1, 2, ..., n},and $\boldsymbol{c}$ is some point on the line segment joining $\boldsymbol{a}$ and $\boldsymbol{a} + \boldsymbol{h}$.

This theorem is used to approximate a given function by a polynomial. In the next section we shall use it to derive conditions for locating and classifying extreme points of a function.

Exercises: 1) Write the expression for *A* appearing in Theorem 7.2.

7.3 MAXIMA AND MINIMA

One of the most interesting and well-known applications of Calculus is the location and classification of extreme points of a function. You have solved many such problems involving functions of one or two variables. We shall now extend the definitions of maxima and minima to functions of n variables, and derive suitable tests for their location.

Definition 7.1 : Let $f : \mathbf{R}^n \to \mathbf{R}$. A point $\mathbf{a} \in \mathbf{R}^n$ is said to be a local maximum (or relative maximum) if there exists a neighbourhood N of $\mathbf{a}$, such that $f(\mathbf{x}) \le f(\mathbf{a})$ for every $\mathbf{x} \in \mathbf{N}$.

$f(\mathbf{a})$ is then called the local or relative maximum value.

A local minimum (or relative minimum) is defined in a similar manner. You will agree that

the function $f : \mathbf{R}^5 \to \mathbf{R}$, $f(x_1, x_2, x_3, x_4, x_5) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2$, clearly has a local minimum at (0, 0, 0, 0, 0). Can you find an example of a function with a local maximum?

Definition 7.2 : A point $\mathbf{a} \in \mathbf{R}^n$ is called a **saddle point** of a function $f : \mathbf{R}^n \to \mathbf{R}$, if every ball $B(\mathbf{a}, r)$, $r > 0$, contains points $\mathbf{x}$, such that $f(\mathbf{x}) \ge f(\mathbf{a})$, and also other points $\mathbf{y}$, such that $f(\mathbf{y}) \le f(\mathbf{a})$.

In general, it is not easy to spot the local maximum or local minimum merely by observation. For differentiable functions we can derive tests to locate these values. You know that in the case of a differentiable function of a single variable, the derivative vanishes at an extreme point. We have a very similar test for the location of extreme points of a function of n variables, as you can see in the next theorem.

Theorem 7.5 : If $f : \mathbf{R}^n \to \mathbf{R}$ has a local maximum at $\mathbf{a} \in \mathbf{R}^n$, then $\forall\ i = 1, 2, ..., n$,

$\frac{\partial f}{\partial x_i}(\mathbf{a})$, if it exists, is equal to zero.

Proof : Since f has a local maximum at $\mathbf{a}$, $\exists\ r > 0$, such that $\mathbf{x} \in B(\mathbf{a}, r) \Longrightarrow f(\mathbf{x}) \le f(\mathbf{a})$.

For $i = 1, 2, ..., n$, consider a function $g_i : (a_i - r, a_i + r) \to \mathbf{R}$, such that

$g_i(x) = f(a_1, a_2, ..., a_{i-1}, x, a_{i+1}, ..., a_n)$. Since $f(\mathbf{a})$ is the local maximum value of f, $g_i(a_i)$ is the maximum value of g_i. If $\frac{\partial f}{\partial x_i}(\mathbf{a})$ exists, then $g_i'(a_i)$ also exists, and the two are equal. By applying the first derivative test for functions of one variable to g_i, we get

$\frac{\partial f}{\partial x_i}(\mathbf{a}) = g_i'(a_i) = 0.$

An exactly similar proof will help us conclude that $\frac{\partial f}{\partial x_i}(\boldsymbol{a})$, if it exists, is equal to zero, even when $\boldsymbol{a}$ is a local minimum of f.

Thus, if f has a local extremum at $\boldsymbol{a}$, and all the partial derivatives exist at $\boldsymbol{a}$, then $\nabla f(\boldsymbol{a}) = \mathbf{0}$.

As in the case of functions of one variable, the condition in theorem 7.5 is a necessary one, and is not sufficient. That is, if all the partial derivatives of a function at a point $\boldsymbol{a}$ are zero, we cannot say that $\boldsymbol{a}$ is a local maximum or local minimum point. It may be neither.

An example is the function $f : \boldsymbol{R}^2 \to \boldsymbol{R}$, $f(x, y) = 1 - x^2 + y^2$. Here $f_x = -2x$, and $f_y = 2y$. So, $f_x(0, 0) = 0$ and $f_y(0, 0) = 0$. But you can see clearly, that f has a maximum in the direction of the x-axis, and a minimum in the direction of the y-axis at (0, 0). This means, f has neither a minimum, nor a maximum at (0, 0). In fact (0, 0) is a saddle point for this function.

Definition 7.3 : Let $f : \boldsymbol{R}^n \to \boldsymbol{R}$ be differentiable, and $\boldsymbol{a} \in \boldsymbol{R}^n$. If $\frac{\partial f}{\partial x_i}(\boldsymbol{a})$ is equal to zero

for i = 1, 2, ..., n, then $\boldsymbol{a}$ is called a **critical point,** or a **stationary point** of f.

Theorem 7.5, tells us to look for extreme points among the critical points of a function. We shall now see how to classify these points as local maxima, local minima, or saddle points. This involves second order partial derivatives. This is to be expected, since in one variable functions too, we have a second derivative test to classify stationary points. The proof of the test for several variables involves quadratic forms. You have studied them in T. Y. B. A. /B. Sc. We start with a definition and recall the relevant results.

Definition 7.4 : If $A = (a_{ij})$ is a real symmetric n x n matrix, and $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in \boldsymbol{R}^n$, then $Q(\boldsymbol{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$ is called a **quadratic form associated with *A*.**

We can write $Q(\boldsymbol{x}) = \boldsymbol{x}A\boldsymbol{x}^t$. If A is a diagonal matrix, then $Q(\boldsymbol{x}) = \sum_{i=1}^{n} a_{ii}x_i^2$ is called a diagonal form. Since A is real symmetric, its eigen values are all real. If all the eigen values of A are positive, then $Q(\boldsymbol{x}) \geq 0$ for every $\boldsymbol{x}$, and $Q(\boldsymbol{x}) = 0 \Longrightarrow \boldsymbol{x} = \boldsymbol{0}$. Such a quadratic form is said to be positive definite. If all the eigen values of A are negative, then $Q(\boldsymbol{x}) \leq 0$ for every $\boldsymbol{x}$, and $Q(\boldsymbol{x}) = 0 \Longrightarrow \boldsymbol{x} = \boldsymbol{0}$. Such a quadratic form is called negative definite.

It may not be very easy to get the eigen values. But we have an easier way to decide.

A principal minor of a square matrix, A, is the determinant of the matrix obtained by taking the first k rows, and the first k columns of A, $1 \leq k \leq n$.

If all the principal minors are positive, then the associated quadratic form is positive definite.

If the principal minors are alternately positive and negative, starting with a negative minor for $k = 1$, then the associated quadratic form is negative definite.

If a principal minor of order k is negative, when k is an even number, then $Q(\boldsymbol{x})$ takes both positive and negative values.

We now use these facts about quadratic forms to derive the second derivative test. A definition first.

Definition 7.5 : If f is a C^2 function from $\boldsymbol{R^n}$ to $\boldsymbol{R}$, then the symmetric matrix

$A = H(\boldsymbol{x}) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)$ is called the **Hessian matrix** of f at $\boldsymbol{x}$. Thus,

$$A = H(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

If $\boldsymbol{a} \in \boldsymbol{R^n}$, the first order Taylor formula for f about $\boldsymbol{a}$ gives us the value of $f(\boldsymbol{a} + \boldsymbol{h})$ for small values of $||\boldsymbol{h}||$ as

$$f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a}) + \boldsymbol{\nabla} f(\boldsymbol{a}) \bullet \boldsymbol{h} + R_1(\boldsymbol{c}).$$

If $\boldsymbol{a}$ is a critical point, then $\boldsymbol{\nabla} f(\boldsymbol{a}) = \boldsymbol{0},$ and therefore we get

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = R_1(\boldsymbol{c}).$$

Now, $R_1(\boldsymbol{c}) = \frac{1}{2!}\sum\sum D_{ij} f(\boldsymbol{a} + \theta\boldsymbol{h})\, h_i h_j$, where $0 < \theta < 1$.

$$= \frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a}+\theta\boldsymbol{h})\boldsymbol{h}^t .$$ We write,

$$\frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a}+\theta\boldsymbol{h})\boldsymbol{h}^t - \frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a})\boldsymbol{h}^t = \frac{1}{2!}\boldsymbol{h}[H(\boldsymbol{a}+\theta\boldsymbol{h}) - H(\boldsymbol{a})]\boldsymbol{h}^t = ||\boldsymbol{h}||^2 E(\boldsymbol{a}, \theta)$$. Thus,

$$||\boldsymbol{h}||^2\, |E(\boldsymbol{a}, \theta)| = \frac{1}{2!}\left|\sum\sum\{D_{ij} f(\boldsymbol{a} + \theta\boldsymbol{h}) - D_{ij} f(\boldsymbol{a})\}h_i h_j\right|$$

$$\leq \frac{1}{2!}\sum\sum |D_{ij} f(\boldsymbol{a} + \theta\boldsymbol{h}) - D_{ij} f(\boldsymbol{a})|\, ||\boldsymbol{h}||^2$$

Therefore, $|E(\boldsymbol{a}, \theta)| \leq \frac{1}{2!}\sum\sum |D_{ij} f(\boldsymbol{a} + \theta\boldsymbol{h}) - D_{ij} f(\boldsymbol{a})|$, when $\boldsymbol{h} \neq \boldsymbol{0}$. ....(7.6)
Each term in the finite sum on the right hand side tends to zero as $\boldsymbol{h} \to \boldsymbol{0}$, since $f \in C^2$, and hence the second order derivatives are continuous. Therefore, $E(\boldsymbol{a}, \theta) \to \boldsymbol{0}$, as $\boldsymbol{h} \to \boldsymbol{0}$. We write $\frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a}+\theta\boldsymbol{h})\boldsymbol{h}^t = \frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a})\boldsymbol{h}^t + ||\boldsymbol{h}||^2 E(\boldsymbol{a}, \theta)$, where $E(\boldsymbol{a}, \theta) \to \boldsymbol{0}$, *as* $\boldsymbol{h} \to \boldsymbol{0}$.

Hence, $f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = \frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a})\boldsymbol{h}^t + ||\boldsymbol{h}||^2 E(\boldsymbol{a}, \theta)$. ....(7.7)

Theorem 7.6 : If f is a function from $\boldsymbol{R^n}$ to $\boldsymbol{R}$, and has continuous second order partial derivatives in a ball $B(\boldsymbol{a}; r)$ around a stationary point $\boldsymbol{a}$ of f, then

i) f has a relative minimum at $\boldsymbol{a}$, if $H(\boldsymbol{a})$ is positive definite
ii) f has a relative maximum at $\boldsymbol{a}$, if $h(\boldsymbol{a})$ is negative definite
iii) f has a saddle point at $\boldsymbol{a}$, if $H(\boldsymbol{a})$ has both positive and negative eigen values.

Proof : Using the notations that we have used in the discussion just before this theorem, we can write $f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = \frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a})\boldsymbol{h}^t + ||\boldsymbol{h}||^2 E(\boldsymbol{a}, \theta)$. Since $E(\boldsymbol{a}, \theta) \to \boldsymbol{0}$, as $\boldsymbol{h} \to \boldsymbol{0}$, we can conclude that the sign of $f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})$ will depend on that of $\frac{1}{2!}\boldsymbol{h}H(\boldsymbol{a})\boldsymbol{h}^t$.

i) This value will be positive for all $\boldsymbol{h}$, if $H(\boldsymbol{a})$ is positive definite. Hence,

$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) > 0$ for all $\boldsymbol{h}$, such that $0 < ||\boldsymbol{h}|| < r$. This tells us that $f(\boldsymbol{a} + \boldsymbol{h}) \geq f(\boldsymbol{a})$ for every $\boldsymbol{h} \in B(\boldsymbol{a}; r)$, that is , $\boldsymbol{a}$ is a relative minimum point of f.

The argument for proving ii) and iii) are exactly similar, and we are sure you can write those.

.**Remark 7.2** : i) If an even principal minor, that is a principal minor of even order is negative, then the point is a saddle point.

ii) If $detH(\boldsymbol{a}) = 0$, the test is inconclusive, and $\boldsymbol{a}$ is called a **degenerate stationary point of f.**

Go through the following examples carefully, they illustrate our discussion here.

Example 7.2: Locate and classify the stationary points of the functions given by

i) $x^2 + xy + 2x + 2y + 1$, ii) $x^3 + y^3 - 3xy$, iii) $(x - 1)e^{xy}$.

Solution : i) Let $f(x, y) = x^2 + xy + 2x + 2y + 1$. Then $f_x = 2x + y + 2$, $f_y = x + 2$.

$f_x = f_y = 0 \implies x + 2 = 0$, and $2x + y + 2 = 0 \implies x = -2$ and $y = 2$. Therefore, f has only one stationary point, $(-2, 2)$. Now, $f_{xx} = 2$, $f_{yy} = 1$, and $f_{xy} = 0$.

Thus, $H((-2, 2)) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$, and $det(H((-2, 2))) = -1$.

Therefore, f has a saddle point at $(-2, 2)$.

ii) Let $f(x, y) = x^3 + y^3 - 3xy$. Then, $f_x = 3x^2 - 3y$, $f_y = 3y^2 - 3x$.

$f_x = f_y = 0 \implies y = x^2$, and $x = y^2 \implies x = y = 0$, or $x = y = 1$. Therefore, the stationary points are (0, 0) and (1, 1). Now, $f_{xx} = 6x$, $f_{yy} = 6y$, and $f_{xy} = -3$. Hence,

$H((0, 0)) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$. $det(H(0, 0)) = -9 < 0$, and (0, 0) is a saddle point.

$H((1, 1)) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$. The principal minors are 6, and 27. Both are positive, and hence f has a local minimum at (1, 1).

iii) Let $f(x, y) = (x - 1)e^{xy}$. Then $f_x = e^{xy}(xy - y + 1)$, $f_y = x(x - 1)e^{xy}$

$f_x = 0 \implies xy - y + 1 = 0$, and $f_y = 0 \implies x(x - 1) = 0 \implies x = 0$, or $x = 1$.

$x = 0 \implies y = 1$, and $x = 1$ contradicts $f_x = 0$. So, (0, 1) is the only stationary point.

$f_{xx} = e^{xy}(y + xy^2 - y^2 + y)$, $f_{xy} = e^{xy}(x - 1 + x^2y - xy + x)$, $f_{yy} = x^2(x - 1)e^{xy}$.

Therefore, $H((0, 1)) = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$. $det(H(0, 1)) = -1 < 0$. Hence, (0, 1) is a saddle point.

Example 7.3 : Locate and classify the stationary points of $f(x, y, z) =$ *i)* $xyze^{-x^2-y^2-z^2}$,

ii) $x^2y + y^2z + z^2 - 8\sqrt{2}x$, *iii)* $x^2 - xy + yz^3 - 6z$.

Solution : i) $f_x = yze^{-x^2-y^2-z^2} - 2x^2yze^{-x^2-y^2-z^2} = e^{-x^2-y^2-z^2}yz(1 - 2x^2)$

$f_y = e^{-x^2-y^2-z^2}xz(1 - 2y^2)$, $f_z = e^{-x^2-y^2-z^2}xy(1 - 2z^2)$. Equating to zero these partial derivatives, and solving the resultant equations, we get $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$, $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$, where a, b, c are real numbers, as the stationary points.

$f_{xx} = -4xyze^{-x^2-y^2-z^2} - 2xyz(1 - 2x^2)\,e^{-x^2-y^2-z^2}$

$f_{xy} = z(1 - 2x^2)\,e^{-x^2-y^2-z^2} - 2y^2z(1 - 2x^2)\,e^{-x^2-y^2-z^2}$,

$f_{yz} = e^{-x^2-y^2-z^2}x(1 - 2y^2) - 2xz^2e^{-x^2-y^2-z^2}(1 - 2y^2)$.

We have indicated the procedure. We are sure now you will be able to get f_{xz}, f_{yy}, and f_{zz}. Evaluating these second order partial derivatives at the stationary points, we find,

$H((a, 0, 0)) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & ae^{-a^2} \\ 0 & ae^{-a^2} & 0 \end{pmatrix}$ $detH((a, 0, 0)) = 0$. Therefore, $(a, 0, 0)$ is a degenerate point of f. Similarly, $(0, b, 0)$ and $(0, 0, c)$ are also degenerate points.

$H((\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})) = \begin{pmatrix} -\sqrt{2}e^{-3/2} & 0 & 0 \\ 0 & -\sqrt{2}e^{-3/2} & 0 \\ 0 & 0 & -\sqrt{2}e^{-3/2} \end{pmatrix}$. The minors of this matrix are $-\sqrt{2}e^{-3/2}$, $2e^{-3}$, $-2\sqrt{2}e^{-3/2}$. Therefore, $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ is a local maximum. Check the remaining 7 points. You should get local maxima at $(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$, $(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$, $(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, and local minima at $(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$, $(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, $(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$.

ii) $f_x = 2xy - 8\sqrt{2}$, $f_y = x^2 + 2yz$, $f_z = y^2 + 2z$. Equating these to zero, we get $xy = 4\sqrt{2}$,

$x^2 = -2yz$, $y^2 = -2z$. If x, y, and z are non-zero, we get $x = 2\sqrt{2}$, $y = 2$, and $z = -2$. So, the stationary points are (0, 0, 0) and $(2\sqrt{2}, 2, -2)$.

You will find that (0, 0, 0) is a degenerate stationary point, and $(2\sqrt{2}, 2, -2)$ is a saddle point.

iii) $f_x = 2x - y$, $f_y = -x + z^3$, $f_z = 3yz^2 - 6$. Equating these to zero, we get (1, 2, 1) as the stationary point. Check that $H((1, 2, 1)) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 0 & 3 \\ 0 & 3 & 12 \end{pmatrix}$, and the principal minors are 2, -1, - 6. Hence, (1, 2, 1) is a saddle point.

See if you can solve these exercises now.

Exercises:

1) Find the stationary points of $f(x, y) =$ i) $\dfrac{x}{x^2 + y^2 - 4}$ ii) $(x + y)e^{xy}$. $\dfrac{x}{x^2 + y^2 - 4}$

2) Find the extreme values of $f(x, y) = x^2 + y^3 + 3xy^2 - 2x$.

3) Is (0, 0) an extreme point of $2\cos(x + y) + e^{xy}$?

4) Locate and classify the stationary points of

i) $f(x, y) = (2 - x)(4 - y)(x + y - 3)$, ii) $f(x, y, z) = 4xyz - x^4 - y^4 - z^4$,

iii) $f(x, y, z) = 64x^2y^2 - z^2 + 16x + 32y + z$, iv) $f(x, y, z) = xyz(x + y + z - 1)$.

7.4 LAGRANGE'S MULTIPLIERS

Look at these situations: i) A rectangular cardboard sheet is given. We have to make a closed box out of it. What is the maximum volume that is possible?

ii) Temperature varies on a metal surface according to some formula. Where do the maximum and minimum temperature occur on the surface?

In both these problems we have to maximize or minimize a certain function: volume in the first case, and temperature in the second. So these are max-min. Problems. But there is a difference between these and the problems considered in the last section. Here, an additional constraint or condition is imposed. The given cardboard sheet has a fixed area. The maximum/minimum temperature points are to be on the given surface.

In this section we shall see how such problems are solved. A very useful method was developed by Joseph Louis Lagrange. This method gives a necessary condition for the extreme points of a function. We now state the theorem and then illustrate its use through some examples.

Theorem 7.7 : Let $f : \boldsymbol{R}^n \to \boldsymbol{R}$, and $f \in C^1$. Suppose $g_1, g_2, \ldots, g_m$ $(m < n)$ are functions belonging to C^1, which vanish on an open set E in $\boldsymbol{R}^n$. If $\boldsymbol{a} \in$ E is an extreme point of f, and if $\nabla g_1(\boldsymbol{a}), \nabla g_2(\boldsymbol{a}), \ldots, \nabla g_m(\boldsymbol{a})$ are independent vectors in $\boldsymbol{R}^n$, then there exist real numbers, λ_1, $\lambda_2, \ldots, \lambda_m$, such that

$D_if(\boldsymbol{a}) + \lambda_1 D_ig_1(\boldsymbol{a}) + \lambda_2 D_ig_2(\boldsymbol{a}) + \ldots + \lambda_m D_ig_m(\boldsymbol{a}) = 0, \quad i = 1, 2, \ldots, n.$

We can also write the vector equation $\nabla f(\boldsymbol{a}) + \sum_1^n \lambda_i \nabla g_i(\boldsymbol{a}) = \boldsymbol{0}$.

When we want to find the extreme values of a function $f : \boldsymbol{R^n} \to \boldsymbol{R}, f \in C^1$, subject to some constraints, $g_1(x_1, x_2, \ldots, x_n) = 0, g_2(x_1, x_2, \ldots, x_n) = 0, \ldots, g_m(x_1, x_2, \ldots, x_n) = 0,$ where m < n, we set up the n equations

$D_if(\boldsymbol{a}) + \lambda_1 D_ig_1(\boldsymbol{a}) + \lambda_2 D_ig_2(\boldsymbol{a}) + \ldots + \lambda_m D_ig_m(\boldsymbol{a}) = 0, \quad i = 1, 2, \ldots, n.$

These n equations, along with the m equations, $g_1(x_1, x_2, \ldots, x_n) = 0, g_2(x_1, x_2, \ldots, x_n) = 0, \ldots, g_m(x_1, x_2, \ldots, x_n) = 0,$ are then solved to get the values of the n + m unknowns, $x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m$. The solutions $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ are the stationary points, and contain the extreme points of f.

$\lambda_1, \lambda_2, \ldots, \lambda_m$ are called **Lagrange's Multipliers**. We use one multiplier for each constraint.

To analytically classify these stationary points into local maximum, minimum, or saddle, is a very complicated process. It is usually easier to look at the physical or geometrical aspect of the problem to arrive at any conclusion. We now solve a few problems, so that the entire process is clear to you.

Example 7.4 : Find the dimensions of the box with maximum volume that can be made with a cardboard sheet of size 12 cm^2.

Solution : If the dimensions of the box are x, y, z cms, then its volume $V = xyz$ c. cms. And surface area is $2(xy + yz + xz)$ sq. cms. Here we have to maximize V, subject to a constraint $2(xy + yz + xz) = 12$, or $(xy + yz + xz) = 6$. So, $f(x, y, z) = xyz$, and

$g(x, y, z) = xy + yz + xz - 6$. Hence,

$\boldsymbol{\nabla} f(x, y, z) + \lambda \boldsymbol{\nabla} g(x, y, z) = \boldsymbol{0} \Longrightarrow$

$f_x + \lambda g_x = 0 \Longrightarrow yz + \lambda(y + z) = 0,\; f_y + \lambda g_y = 0 \Longrightarrow xz + \lambda(x + z) = 0,\; f_z + \lambda g_z = 0 \Longrightarrow xy + \lambda(x + y) = 0.$

$xyz = -\lambda(xy + xz) = -\lambda(xy + yz) = -\lambda(xz + yz)$. If $\lambda = 0$, then $V = 0$, which is the minimum volume. If $\lambda \neq 0$, then $xy + xz = xy + yz = xz + yz$. That is, $x = y = z$ (unless, of course, $x = y = z = 0$).

Therefore, $xy + yz + xz = 6 \Longrightarrow 3x^2 = 6 \Longrightarrow x = \sqrt{2}$ cms. Thus, $V = 2\sqrt{2}$ c. cms. is the maximum volume.

Example 7.5 : Find the extreme values of the function given by $f(x, y, z) = 2x + y + 3z$, subject to $x^2 + y^2 = 2, x + z = 5$.

Solution : Let $g_1(x, y, z) = x^2 + y^2 - 2 = 0$, and $g_2(x, y, z) = x + z - 5 = 0$. Then

$\boldsymbol{\nabla} f + \lambda_1 \boldsymbol{\nabla} g_1 + \lambda_2 \boldsymbol{\nabla} g_2 = \boldsymbol{0} \Longrightarrow$

$f_x + \lambda_1 g_{1x} + \lambda_2 g_{2x} = 0 \Longrightarrow 2 + 2\lambda_1 x + \lambda_2 = 0$

$f_y + \lambda_1 g_{1y} + \lambda_2 g_{2y} = 0 \Longrightarrow 1 + 2\lambda_1 y = 0$

$f_z + \lambda_1 g_{1z} + \lambda_2 g_{2z} = 0 \Longrightarrow 3 + \lambda_2 = 0$. Therefore, $\lambda_2 = -3$, $2\lambda_1 x = 1$, and $2\lambda_1 y = -1$.

$\lambda_1 = 0 \Longrightarrow \lambda_2 = -2$. But $\lambda_2 = -3$. Therefore λ_1 cannot be zero. Hence, $x = \frac{1}{2\lambda_1}$, $y = \frac{-1}{2\lambda_1}$.
Substituting these values in $x^2 + y^2 = 2$, we get $\lambda_1 = \pm\frac{1}{2}$. This gives, $x = \pm 1$, $y = \mp 1$. Hence, the stationary points are (1, - 1, 4) and (- 1, 1, 6), and the extreme values are 13 and 17.

Example 7.6 : Find the minimum distance of a point on the intersection of the planes,

$x + y - z = 0$, and $x + 3y + z = 2$ from the origin.

Solution : The distance of $P(x, y, z)$ from the origin is $\sqrt{x^2 + y^2 + z^2}$. So, we need to minimize $f(x, y, z) = x^2 + y^2 + z^2$, subject *to* $g_1(x, y, z) = x + y - z = 0$, and

$g_2(x, y, z) = x + 3y + z - 2 = 0$.

$\boldsymbol{\nabla} f + \lambda_1 \boldsymbol{\nabla} g_1 + \lambda_2 \boldsymbol{\nabla} g_2 = \boldsymbol{0} \Longrightarrow$

$f_x + \lambda_1 g_{1x} + \lambda_2 g_{2x} = 0 \Longrightarrow 2x + \lambda_1 + \lambda_2 = 0$

$f_y + \lambda_1 g_{1y} + \lambda_2 g_{2y} = 0 \Longrightarrow 2y + \lambda_1 + 3\lambda_2 = 0$

$f_z + \lambda_1 g_{1z} + \lambda_2 g_{2z} = 0 \Longrightarrow 2x - \lambda_1 + \lambda_2 = 0$. Therefore, $x = \frac{-(\lambda_1 + \lambda_2)}{2}$, $y = \frac{-(\lambda_1 + 3\lambda_2)}{2}$,

$z = \frac{(\lambda_1 - \lambda_2)}{2}$. Putting these values in $x + y - z = 0$, we get $\lambda_1 + \lambda_2 = 0$. Therefore, $x = 0$ and

$y = z$. Using this in $x + 3y + z - 2 = 0$, we get $y = z = ½$. Thus, the stationary point is

(0, 1/2, 1/2). The distance of this point from the origin is $\frac{1}{\sqrt{2}}$.

Geometrically, the constraints are equations of two planes. There is no maximum to the distance of a point on their line of intersection from the origin. So, the stationary point is a minimum point.

Here are some problems you can try.

1) Find the extreme values of the function $f(x, y) = xy$ on the surface $\frac{x^2}{8} + \frac{y^2}{2} = 1$.

2) Find the extreme values of $z = \frac{x}{2} + \frac{y}{3}$ on the unit circle in the *xy*-plane.

3) Find the distance of the point (10, 1, − 6) from the intersection of the planes,

$x + y + 2z = 5$ and $2x - 3y + z = 12$.

7.5 SUMMARY

In this chapter we have introduced Taylor's theorem for functions of several variables. We have also seen how to get Taylor polynomials of a given order for a given function. Of course, to be able to do this, the function must have continuous partial derivatives of higher orders.

We have then discussed the location of maxima and minima of a real-valued function of several variables. This has tremendous applications in diverse fields of study. In particular, we have proved that the extreme points of a function are located among the points at which the gradient vector of the function is zero. That is, the points at which all the first order partial derivatives are zero. The classification of these points into maxima, minima, or saddle points depends on the signs of the principal minors of the Hessian matrix.

We pointed out that there are some situations, where we need to find the extreme values subject to certain constraints. Such problems, and the method of tackling them is also discussed, and illustrated through some examples.

8

INVERSE AND IMPLICIT FUNCTION THEOREMS

Unit Structure

8.0 OBJECTIVES

After reading this chapter, you should be able to

- state and prove Inverse Function Theorem for functions of several variables
- check if some simple functions are locally invertible
- state and prove Implicit Function Theorem for functions of several variables

8.1 INTRODUCTION

In this chapter we introduce two very important theorems. You have not come across these theorems even for functions of a single variable. In each case, we shall first discuss the single variable case, and then extend the concept to functions of several variables. A word of caution : these theorems are not easy. To help you understand them better, we are going to prove some smaller results, and then use them in the proof of the theorems. Do study this chapter carefully and we are sure you would have no difficulty in digesting the concepts.

8.2 INVERSE FUNCTION THEOREM

The inverse function theorem is a very important theorem in Calculus. You may be familiar with its one dimensional version. Before we introduce the theorem for functions from $\boldsymbol{R}^n$ to $\boldsymbol{R}^n$, we shall recall some results about functions of one variable:

1) If $f : [a, b] \rightarrow \boldsymbol{R}$ is continuous, and $f(c) > 0$ for some $c \in (a, b)$, then $\exists\ \varepsilon > 0$, such that

$(c-\varepsilon,\ c+\varepsilon) \subseteq (a, b)$, and $f(x) > 0\ \ \forall\ x \in (c-\varepsilon,\ c+\varepsilon)$. In other words, we can always find a neighbourhood of the point c, in which $f(x)$ has the same sign as $f(c)$.

2) If $f : [a, b] \rightarrow \boldsymbol{R}$ is a continuously differentiable function, and $f'(c) \neq 0$ for some

$c \in (a, b)$, then using 1) we can prove that $\exists\ \varepsilon > 0$, such that f is an injective function on

$(c-\varepsilon,\ c+\varepsilon) \subseteq (a, b)$. Further, $f^{-1}: f(c-\varepsilon,\ c+\varepsilon) \to (c-\varepsilon,\ c+\varepsilon)$ is differentiable at $f(c)$,

The statement in 2) is the inverse function theorem. Note that we do not know whether the inverse of f exists on $[a, b]$. But what this theorem tells us, is that if $f'(c) \neq 0$, then f is "locally invertible" at c. For example, we know that the function $f : [0, 2\pi] \to \boldsymbol{R}$, $f(x) = sinx$ does not have an inverse. But $f'(x) = cosx$ is a continuous function, and $f'\left(\frac{\pi}{3}\right) = \frac{1}{2} \neq 0$. So, the theorem says that f is locally invertible at $\pi/3$. That is, we can find a neighbourhood N of $\pi/3$, such that f restricted to N has an inverse. Check that f is injective when restricted to $N = (\frac{\pi}{4}, \frac{5\pi}{12})$, and hence has an inverse on N.

We shall now see if this theorem extends to functions of several variables. Let us start with a definition.

Definition 8.1 : Let $\boldsymbol{f} : \mathrm{E} \to \boldsymbol{R}^{\boldsymbol{n}}$, where $\mathrm{E} \subseteq \boldsymbol{R}^{\boldsymbol{n}}$. If $\boldsymbol{f} \in C^1$, $\boldsymbol{f}$ is said to be **locally invertible** at $\boldsymbol{a} \in \mathrm{E}$, if there exists a neighbourhood N_1 of $\boldsymbol{a}$, $N_1 \subseteq \mathrm{E}$, and a neighbourhood N_2 of $\boldsymbol{f(a)}$, such that $\boldsymbol{f}(N_1) = N_2$, $\boldsymbol{f}$ is injective on N_1, and $\boldsymbol{f}^{-1} : N_2 \to N_1$ is a C^1 function.

We shall soon state and prove the inverse function theorem. In the proof, we are going to use some minor results. You have already studied some in the earlier chapters of this course. Next we state and prove one other result, which will be useful to us.

Theorem 8.1 : Let $\boldsymbol{f} = (f_1\, f_2, \ldots, f_n) : \mathrm{E} \to \boldsymbol{R}^{\boldsymbol{n}}$, where E is an open set in $\boldsymbol{R}^{\boldsymbol{n}}$. Suppose $\boldsymbol{f} \in C^1$. If the Jacobian of $\boldsymbol{f}$, $J(\boldsymbol{a}) \neq 0$ for some $\boldsymbol{a} \in \mathrm{E}$, then $\boldsymbol{f}$ is injective on a neighbourhood of $\boldsymbol{a}$ in E.

Proof : If $\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n} \in \mathrm{E}$, we consider a point $\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}) \in \boldsymbol{R}^{\boldsymbol{n}^2}$, whose first n coordinates are the coordinates of $\boldsymbol{X_1}$, the next n are the coordinates of $\boldsymbol{X_2}$, and so on. We define a function, j, such that

$$j(\boldsymbol{X}) = det[D_jf_i(\boldsymbol{X_i})] = det\begin{pmatrix} D_1f_1(X_1) & D_2f_1(X_1) & .. & . & . & D_nf_1(X_1) \\ D_1f_2(X_2) & D_2f_2(X_2) & . & . & . & D_nf_2(X_2) \\ . & & & & & \\ . & & & & & \\ . & & & & . & . \\ D_1f_n(X_n) & D_nf_n(X_n) & . & . & . & D_nf_n(X_n) \end{pmatrix}.$$

Now, the function j, being an n×n determinant, is a polynomial of its n^2 entries, and each entry, $D_jf_i(X_i)$ is a continuous function, since $f \in C^1$. Thus, j is a continuous function on its domain. We write $\boldsymbol{A} = (\boldsymbol{a}, \boldsymbol{a}, \ldots, \boldsymbol{a})$. Then $j(\boldsymbol{A}) = det[D_jf_i(\boldsymbol{a})] = J(\boldsymbol{a}) \neq 0$. Now, since $\boldsymbol{f} \in C^1$, all the entries of $j(\boldsymbol{A})$ are continuous, and hence, $j(\boldsymbol{A})$ is also continuous. The continuity of $j(\boldsymbol{A})$ ensures that there exists a neighbourhood N of $\boldsymbol{A}$, such that $j(\boldsymbol{X}) \neq 0$, if $\boldsymbol{X} \in N$.

In other words, there exists a convex neighbourhood N_a of $\boldsymbol{a}$, such that $j(\boldsymbol{X}) \neq 0$, if

$\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n})$ is a point, for which $\boldsymbol{X_i} \in N_a$ for every i = 1, 2, . . . , n. ....(8.1)

This N_a is the required neighbourhood. We have to show that $\boldsymbol{f}$ is injective on N_a. For this, suppose $\boldsymbol{x}, \boldsymbol{y} \in N_a$, such that $\mathbf{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{y})$. Then $f_i(\boldsymbol{x}) = f_i(\boldsymbol{y})$ for every i = 1, 2, . . . , n.

Then, using the Mean Value Theorem for scalar fields (See Remark 6.2 ii).), we get

$f_i(\boldsymbol{x}) - f_i(\boldsymbol{y}) = \nabla f_i(\boldsymbol{c_i}) \circ (\boldsymbol{x} - \boldsymbol{y}) \Longrightarrow \nabla f_i(\boldsymbol{c_i}) \circ (\boldsymbol{x} - \boldsymbol{y}) = 0$ for some $\boldsymbol{c_i}$ on the line segment joining $\boldsymbol{x}$ and $\boldsymbol{y}$. So, if $\boldsymbol{x} - \boldsymbol{y} \neq \mathbf{0}$, then $\nabla f_i(\boldsymbol{c_i}) = \mathbf{0}$ for some $\boldsymbol{c_i}$ on the line segment joining $\boldsymbol{x}$ and $\boldsymbol{y}$, that is, in the neighbourhood N_a, since N_a is convex. This means, $D_j f_i(\boldsymbol{c_i}) = 0$ for every j, $1 \leq j \leq n$, $1 \leq i \leq n$. Thus, if $\boldsymbol{C} = (\boldsymbol{c_1}, \boldsymbol{c_2}, . . . , \boldsymbol{c_n})$, then $j(\boldsymbol{C}) = det[D_j f_i(\boldsymbol{c_i})] = 0$. But this contradicts (8.1). So, we conclude that $\boldsymbol{x} - \boldsymbol{y} = \mathbf{0}$, which proves that $\boldsymbol{f}$ is injective on N_a.

Remark 8.1 : i) A function may not be injective on its entire domain. But if its Jacobian is non-zero at a point, then it is injective on a neighbourhood of that point. In other words, it is locally injective.

ii) If the Jacobian is non-zero, then the linear transformation D$\boldsymbol{f}$, which represents the derivative of $\boldsymbol{f}$, is non-singular, and hence, is a linear isomorphism.

Example 8.1 : a) Consider the function $\boldsymbol{f}(x, y) = (e^x cosy, e^x siny)$. This function is not injective, since $\boldsymbol{f}(x, 0) = \boldsymbol{f}(x, 2\pi)$. But,

$J(x, y) = \begin{vmatrix} e^x cosy & -e^x siny \\ e^x siny & e^x cosy \end{vmatrix} = e^{2x} \neq 0$. Thus, $\boldsymbol{f}$ is locally injective at each point in $\boldsymbol{R^2}$.

Here we have a function, which is locally injective at every point of its domain, but is not injective on the domain.

b) Consider the function $\boldsymbol{f}(x, y) = (x^3, y^3)$, defined on $\boldsymbol{R^2}$. The Jacobian of this function is zero at (0, 0). But the function is locally invertible at (0, 0). In fact, it is an invertible function.

Theorem 8.2 (The Inverse Function Theorem): Let $\boldsymbol{f} = (f_1, f_2, . . . , f_n) \in C^1$, f: E $\rightarrow \boldsymbol{R^n}$, where E is an open set in $\boldsymbol{R^n}$. Let T = $\boldsymbol{f}$(E). Suppose $J(\boldsymbol{a}) \neq 0$ for some $\boldsymbol{a} \in$ E. Then there exists a unique function $\boldsymbol{f^{-1}}$ from Y to X, where X is open in E, Y is open in T, such that

i) $\boldsymbol{a} \in$ X, $\boldsymbol{f}(\boldsymbol{a}) \in$ Y, ii) Y = $\boldsymbol{f}$(X), iii) f is injective on X, iv) $\boldsymbol{f^{-1}}$: Y$\rightarrow$ X, $\boldsymbol{f^{-1}}$(Y) = X, v) $\boldsymbol{f^{-1}} \in C^1$ on Y.

Proof : Using Theorem 8.1, we can conclude that $\boldsymbol{f}$ is injective on a neighbourhood N of $\boldsymbol{a}$ in E. So, $\boldsymbol{f} : N \rightarrow \boldsymbol{f}(N)$ is bijective, and hence has an inverse, $\boldsymbol{f^{-1}} : \boldsymbol{f}(N) \rightarrow N$. Let r > 0 be such that $\overline{B(\boldsymbol{a}, r)} \subseteq N$. Since $\overline{B(\boldsymbol{a}, r)}$ is compact in $\boldsymbol{R^n}$, we use Theorem 3.4.1 to conclude that $\boldsymbol{f}(\overline{B(\boldsymbol{a}, r)})$ is also compact in $\boldsymbol{R^n}$. Now $\boldsymbol{f}$ is continuous and injective on the compact set $\overline{B(\boldsymbol{a}, r)}$. Hence, using Theorem 3.4.2, we can say that $\boldsymbol{f^{-1}}$ is continuous on $\boldsymbol{f}(\overline{B(\boldsymbol{a}, r)})$.

Now, B($\boldsymbol{a}$, r) is an open set in $\overline{B(\boldsymbol{a}, r)}$, and therefore,

$(\boldsymbol{f}^{-1})^{-1}(\mathrm{B}(\boldsymbol{a}, \mathrm{r}))$ is open in $\boldsymbol{f}(\overline{\mathrm{B}(\boldsymbol{a}, \mathrm{r})})$. That is, $\boldsymbol{f}(\mathrm{B}(\boldsymbol{a}, \mathrm{r}))$ is open in $\boldsymbol{f}(\overline{\mathrm{B}(\boldsymbol{a}, \mathrm{r})})$.

Also, $\boldsymbol{f}(\boldsymbol{a}) \in \boldsymbol{f}(\mathrm{B}(\boldsymbol{a}, \mathrm{r}))$. Therefore, there exists a $\delta > 0$, such that $\mathrm{B}(\boldsymbol{f}(\boldsymbol{a}), \delta) \subseteq \boldsymbol{f}(\mathrm{B}(\boldsymbol{a}, \mathrm{r}))$.

Take $\mathrm{X} = \boldsymbol{f}^{-1}(\mathrm{B}(\boldsymbol{f}(\boldsymbol{a}), \delta))$, and $\mathrm{Y} = \mathrm{B}(\boldsymbol{f}(\boldsymbol{a}), \delta)$. Then X and Y satisfy i), ii), iii) and iv) in the statement of the theorem.

To prove the last assertion v) in the statement, we have to show that all the partial derivatives of all the component functions of $\boldsymbol{f}^{-1}$ are continuous on Y. For this we first define the function $j(\boldsymbol{X}) = det[D_j f_i(\boldsymbol{x_i})]$, as in Theorem 8.1. Here $\boldsymbol{X} = (\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n})$. Then, as before, there is a neighbourhood N_a of $\boldsymbol{a}$, such that $j(\boldsymbol{X}) \neq 0$, whenever each $\boldsymbol{X_i} \in N_a$. We can assume that the neighbourhood $N \subseteq N_a$. This ensures that $j(\boldsymbol{X}) \neq 0$, whenever each $\boldsymbol{X_i} \in \overline{\mathrm{B}(\boldsymbol{a}, \mathrm{r})}$.

Now we first prove that $\mathrm{D}_i\boldsymbol{f}^{-1}$ exists on Y. Let $\boldsymbol{y} \in \mathrm{Y}$, and consider $\dfrac{f^{-1}(y+te_i)-f^{-1}(y)}{t}$,

where $\boldsymbol{e_i}$ is the i[th] coordinate vector, and t is a scalar. Let $\boldsymbol{x} = \boldsymbol{f}^{-1}(\boldsymbol{y})$, and $\boldsymbol{x}' = \boldsymbol{f}^{-1}(\boldsymbol{y} + t\boldsymbol{e_i})$. Then

$\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x}) = t\boldsymbol{e_i}$. Thus, $f_i(\boldsymbol{x}') - f_i(\boldsymbol{x}) = t$, and $f_j(\boldsymbol{x}') - f_j(\boldsymbol{x}) = 0$, when $\mathrm{i} \neq \mathrm{j}$.

By applying Mean Value Theorem (Remark 6.2 ii)), we can write

$\dfrac{f_m(x')-f_m(x)}{t} = \nabla f_m(\boldsymbol{x_m}) \bullet \dfrac{x'-x}{t}$, m = 1, 2, . . . , n. Here $\boldsymbol{x_m}$ is a point on the line segment joining $\boldsymbol{x}$ and $\boldsymbol{x}'$.

So, we get a system of n equations (for the n values of m). The left hand side of an equation in this system is 1, if m = i, otherwise it is 0. The right hand side is of the form

$$D_1 f_m(\boldsymbol{x_m}) \frac{x_1'-x_1}{t} + D_2 f_m(\boldsymbol{x_m}) \frac{x_2'-x_2}{t} + \ldots + D_n f_m(\boldsymbol{x_m}) \frac{x_n'-x_n}{t}, \quad m = 1, 2, \ldots, n.$$

The determinant of this system of linear equations is $j(\boldsymbol{X})$, which we know is non-zero. Hence we can solve it by Cramer's rule and get the variables $\dfrac{x_j'-x_j}{t}$ as the quotient of two determinants. Then, as t tends to zero, $\boldsymbol{x}'$ approaches $\boldsymbol{x}$, and hence, each $\boldsymbol{x_m}$ also approaches $\boldsymbol{x}$. The determinant in the denominator, $j(\boldsymbol{X}) = det[D_j f_i(\boldsymbol{x_i})]$ then approaches $J(\boldsymbol{x})$, the Jacobian of $\boldsymbol{f}$ at $\boldsymbol{x}$, which is again non-zero. Thus, as t tends to zero, the limit of $\dfrac{x_j'-x_j}{t}$ exists. That is, $\lim\limits_{t\to 0} \dfrac{f^{-1}(y+te_i)-f^{-1}(y)}{t}$ exists. Thus, $D_i\boldsymbol{f}^{-1}(\boldsymbol{y})$ exists for all i, and for all $\boldsymbol{y}$ in Y.

We have obtained the partial derivatives of the components of $\boldsymbol{f}^{-1}$ as quotients of two determinants. The entries in these determinants are partial derivatives of the components of $\boldsymbol{f}$, which are all continuous. Since a determinant is a polynomial of its entries, we conclude that the partial derivatives of $\boldsymbol{f}^{-1}$ are continuous on Y.

Example 8.2 : Show that the function $\boldsymbol{f}: \boldsymbol{R}^2 \to \boldsymbol{R}^2$, $\boldsymbol{f}(x, y) = (2xy, x^2 - y^2)$ is not invertible on $\boldsymbol{R}^2$, but is locally invertible at every point of E = {$(x, y) \mid x > 0$}. Also find the inverse function at one such point.

Solution : Here $\boldsymbol{f}(1, 1) = \boldsymbol{f}(-1, -1) = (2, 0)$. Therefore $\boldsymbol{f}$ is not injective, and hence is not invertible on $\boldsymbol{R}^2$. On the other hand, if $(x, y) \in$ E, then

$J(x, y) = \begin{vmatrix} 2y & 2x \\ 2x & -2y \end{vmatrix} = -4(x^2 + y^2) \neq 0$. Hence by the inverse function theorem, $\boldsymbol{f}$ is locally invertible.

Suppose $f(x, y) = (u, v)$. If $(x, y) \in$ E, then $y = \dfrac{u}{2x}$, and $v = x^2 - \dfrac{u^2}{4x^2}$. Therefore,

$4x^4 - 4x^2v - u^2 = 0$. Thus, $x^2 = \dfrac{v + \sqrt{v^2 + u^2}}{2}$, and $x = \left(\dfrac{v + \sqrt{v^2 + u^2}}{2}\right)^{1/2}$,

$y = \mathrm{u}(2\mathrm{v} + 2\sqrt{\mathrm{v}^2 + \mathrm{u}^2})^{-1/2}$

8.3 IMPLICIT FUNCTION THEOREM

If $x^2 + y^2 = 0$, find $\frac{dy}{dx}$. You must have done exercises like this in your under-graduate classes. Here, we take $f(x, y) = x^2 + y^2$, and find $f_x = 2x$, and $f_y = 2y$. Then $\dfrac{dy}{dx} = 2x/2y = x/y$. Of course, y cannot be zero.

While doing this exercise, actually we have used a theorem, the implicit function theorem. To recall, in this setting, a function which can be written as $y = g(x)$, is called an explicit function, and one which can be expressed only as $f(x, y) = 0$, is called an implicit function. The implicit function tells us that under certain conditions, we can express an implicit function as an explicit one, and then we can use this expression to find $\dfrac{dy}{dx}$.

In this section we are going to discuss this implicit function theorem for functions of several variables. Before we state and prove the general case, we first prove the case for functions involving only two variables, x and y.

Theorem 8.3 : Let f be a real-valued C^1 function, defined on the product $\mathrm{I}_1 \times \mathrm{I}_2$, where I_1 and I_2 are two intervals in $\boldsymbol{R}$. Let $(a, b) \in \mathrm{I}_1 \times \mathrm{I}_2$, and $f(a, b) = 0$, but $f_y(a, b) \neq 0$. Then there exists an interval I in $\boldsymbol{R}$, containing a, and a C^1 function $g : \mathrm{I} \to \boldsymbol{R}$, such that $g(a) = b$, and

$f(x, g(x)) = 0$ for all $x \in$ I.

Proof : We consider a function, $\boldsymbol{h}$: $I_1 \times I_2 \to \boldsymbol{R}^2$, given by $h(x, y) = (x, f(x, y))$. If we write

$\boldsymbol{h} = (h_1, h_2)$, the Jacobian matrix of $\boldsymbol{h}$ is

$J_h(x, y) = \begin{pmatrix} \frac{\partial h_1}{\partial x} & \frac{\partial h_1}{\partial y} \\ \frac{\partial h_2}{\partial x} & \frac{\partial h_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}$. The determinant of this matrix, $\frac{\partial f}{\partial x}$ is not zero at (a, b).

Thus, h is a C^1 function, with a non-zero Jacobian at (a, b). Therefore, by the inverse function theorem, Theorem 8.2 , we can conclude that h is locally invertible at (a, b). Let $u = (u_1, u_2)$ be the local inverse of h. You will agree that $u_1(x, y) = x$ for all x and y in $\boldsymbol{R}$. That is,

$u(x, y) = (x, u_2(x, y))$ for all x and y in $\boldsymbol{R}$. We now define g as, $g(x) = u_2(x, 0)$, and show that it has all the required properties.

Now, since $h(a, b) = (a, 0)$, $u(a, 0) = (a, b)$. This means, $u_2(a, 0) = b$. Thus, $g(a) = b$.

Also, $(x, 0) = h(u(x, 0)) = h(x, u_2(x, 0)) = h(x, g(x)) = (x, f(x, g(x)))$. This implies that

$f(x, g(x)) = 0$.

Since u is a C^1 function, g is also C^1. Differentiating $f(x, g(x)) = 0$ with respect to x using chain rule, we get $D_1f(x, g(x)) + D_2f(x, g(x))g'(x) = 0$, and thus,

$g'(x) = \frac{-D_1(f(x,g(x))}{D_2 f(x,g(x))}$, since $D_2f(x, g(x)) \neq 0$.

Basically, this theorem tells us that under certain conditions, the relation $f(x, y) = 0$, between x and y can be explicitly written as $y = g(x)$.

Remark 8.2 : If instead of $f_y(a, b) \neq 0$, we take the condition $f_x(a, b) \neq 0$, then we can express x as an explicit function of y.

Example 8.3 : Can $f(x, y) = x^3 + y^3 - 2xy$ be expressed by an explicit function $y = g(x)$ in a neighbourhood of the point (1, 1)?

Solution : Note that $f(1,1) = 0$, and $f_y = 3y^2 - 2x = 1$ at (1, 1). Further, f is a C^1 function on $\boldsymbol{R}^2$. Therefore, we can apply Theorem 8.3, and conclude that there exists a unique function g, defined on a neighbourhood of 1, such that $g(1) = 1$. Also, $g'(x) = \frac{3x^2 - 2y}{3y^2 - 2x}$ in this neighbourhood.

Example 8.4 : Check whether Theorem 8.3 can be applied at all points, where

$f(x, y) = x^2 - y^2 = 0$.

Solution : $x^2 - y^2 = 0$ is true at points (0, 0), (1, 1),(1, −1), (−1, 1), and (−1, −1). $f_y = -2y$, and $f_x = 2x$. At the point (0, 0), f_x and f_y are both zero, and hence we cannot apply the theorem. At all the remaining points, the function satisfies all the conditions of Theorem 8.3, and hence it can be applied. You will agree that at each of these points, we will get either

$g(x) = x$, or $g(x) = -x$.

We now go a step further, and consider a real-valued function of several variables.

Theorem 8.4 : Let f be a real-valued C^1 function, defined on an open set, U, in $\boldsymbol{R}^n$. Let

$\boldsymbol{a} = (a_1, a_2, \ldots, a_{n-1}) \in \boldsymbol{R}^{n-1}$, such that $(\boldsymbol{a}, b) \in$ U, $f(\boldsymbol{a}, b) = 0$, and $D_n f(\boldsymbol{a}, b) \neq 0$. Then there exists a unique C^1 function g, defined on a neighbourhood N of $\boldsymbol{a}$, such that $g(\boldsymbol{a}) = b$, and

$f(\boldsymbol{x}, g(\boldsymbol{x})) = 0$ for all $\boldsymbol{x} \in N$.

Proof : We consider a function $\boldsymbol{h}$: U $\rightarrow \boldsymbol{R}^{n-1} \times \boldsymbol{R}$, defined by $\boldsymbol{h}(\boldsymbol{x}, y) = (\boldsymbol{x}, f(\boldsymbol{x}, y))$. If we write $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$, then $h_i(\boldsymbol{x}, y) = x_i$, for $1 \le i \le n-1$, and $h_n(\boldsymbol{x}, y) = f(\boldsymbol{x}, y)$. Therefore, the Jacobian matrix of $\boldsymbol{h}$ is given by

$$J_{\boldsymbol{h}} = \begin{pmatrix} 1 & 0 & \cdot\cdot & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ . & & & & & \\ . & & & & & \\ 0 & & & & 1 & 0 \\ D_1 f & D_2 f & . & . & . & D_n f \end{pmatrix}.$$

The determinant of this matrix is $D_n\boldsymbol{f}$, which is non-zero. Therefore, we can apply the inverse function theorem (Theorem 8.2), and conclude that $\boldsymbol{h}$ is locally invertible at $(\boldsymbol{a}, b)$. If $\boldsymbol{u}$ is the local inverse of $\boldsymbol{h}$, and we write $\boldsymbol{u} = (\boldsymbol{u_1}, u_2)$, then you will see that $\boldsymbol{u_1}(\boldsymbol{x}, y) = \boldsymbol{x}$ for all $(\boldsymbol{x}, y)$. Thus, $\boldsymbol{u}(\boldsymbol{x}, y) = (\boldsymbol{x}, u_2(\boldsymbol{x}, y))$ for all $(\boldsymbol{x}, y)$. We now define $g(\boldsymbol{x}) = u_2(\boldsymbol{x}, 0)$, and show that this has the required properties.

Now, $\boldsymbol{u}(\boldsymbol{a}, 0) = (\boldsymbol{a}, b)$. This gives $g(\boldsymbol{a}) = u_2(\boldsymbol{a}, 0) = b$.

Also, $(\boldsymbol{x}, 0) = \boldsymbol{h}(\boldsymbol{u}(\boldsymbol{x}, 0)) = \boldsymbol{h}(\boldsymbol{x}, u_2(\boldsymbol{x}, 0)) = \boldsymbol{h}(\boldsymbol{x}, g(\boldsymbol{x})) = (\boldsymbol{x}, f(\boldsymbol{x}, g(\boldsymbol{x})))$. This implies that

$f(\boldsymbol{x}, g(\boldsymbol{x})) = 0$.

Example 8.5 : Examine whether the function $f(x, y, z) = x^2 + y^2 - 4$ can be expressed as a function $y = g(x, z)$ in a neighbourhood of the point (0, -2, 0).

Solution : We note that $f(0, -2, 0) = 0$, and $D_2 f = 2y = -4$ at (0, −2, 0). So, applying the implicit function theorem, there exists the required neighbourhood of (0, −2, 0). In fact, you can check that in the neighbourhood, $N = B((0, -2, 0), 1)$, we can express the function as

$y = -(4 - x^2)^{1/2}$.

Here are some exercises that you should try :

1) Determine whether the following functions are locally invertible at the given points :

i) $f(x, y) = (x^3y + 3, y^2)$ at (1, 3)

ii) $f(x, y, z) = (e^x cosy, e^x sinz, z)$ at (1, 1, 1).

2) For each of the following functions, show that the equation $f(x, y, z) = 0$ defines a continuously differentiable function $z = g(x, y)$, in a neighbourhood of the given point:

i) $f(x, y, z) = x^3 + y^3 + z^3 – xyz – 2$, (1, 1, 1)

ii) $f(x, y, z) = x^2 + y^3 – xysinz$, (1, - 1, 0).

That brings us to the end of this chapter. We hope you have studied the concepts carefully, and have understood them.

8.4 LET US SUM UP

In this chapter we have discussed two very important theorems: the inverse function theorem, and the implicit function theorem. The proofs of these theorems are a little complicated. So we have tried to go step by step from functions of one variable to functions of many variables.

The Inverse Function Theorem: gives the conditions under which a function, even though not invertible on its domain, is seen to be locally invertible. The Jacobian of the function being non-zero at a point ensures the local invertibility of the function in a neighbourhood of that point.

The Implicit Function Theorem: gives the conditions, under which an implicit relationship between variables can be expressed in an explicit manner. Here, again, the Jacobian plays an important role.

1

RIEMANN INTEGRAL - I

Unit Structure :

1.1 INTRODUCTION

The Riemann integral dealt with in calculus courses, is well suited for computations but less suited for dealing with limit processes.

Bernhard Riemann in 1868 introduced Riemann integral. He need to prove some new result about Fourier and trigonometric series. Riemann integral is based on idea of dividing. The domain of function into small units over each such unit or sub-interval we erect an approximation rectangle. The sum of the area of these rectangles approximates the area under the curve.

As the partition of the interval becomes thinner, the number of sub-interval becomes greater. The approximating rectangles become narrower and more precise. Hence area under the curve is more accurate. As limits of sub-interval tends to zero, the values of the sum of the areas of the rectangles tends to the value of an integral. Hence the area under curve to be equal to the value of the integral.

Before going for exact definition of Riemann explained the following definitions.

1.2 PARTITION

A closed rectangle in $\mathbb{R}^n$ is a subset A of $\mathbb{R}^n$ of the forms.

$A=[a_1,b_1]\times[a_2,b_2]\times....\times[a_n,b_n]$ where $a_i < b_i \in \mathbb{R}$. Note that $(x_1,x_2,....,x_n)\in A$ iff $a_i \le x_i \le b_i \forall i$.

The points $x_1, x_2, ..., x_n$ are called the partition points.

The closed interval $I_1 = [x_0, x_1], I_2 = [x_1, x_2], ..., I_n = [x_{n-1}, x_n]$ are called the component internal of $[a, b]$.

Norm : The norm of a portion P is the length of the largest sub-internal of P and is denoted by $\|P\|$.

For example : Suppose that $P_1 = t_0, t_1, ...t_k$ is a partition of $[a_1, b_1]$ and $P_2 = S_0, ..., S_r$ is a partition of $[a_2, b_2]$. Then the partition $P = (P_1.P_2)$ of $[a_1, b_1] \times [a_2, b_2]$ divides the closed rectangle $[a_1, b_1] \times [a_2, b_2]$ into Kr-gub rectangles.

In general if P_i divides $[a_i, b_i]$ into k_i sub-interval then $P = (P_1, ...P_n)\ [a_1, b_1] \times .... \times [a_n, b_n]$ into $K = k_1 k_2 .....k_n$ sub-rectangle. These sub-rectangles are called sub-rectangles of the partition p.

Refinement :

Definition : Let A be a rectangle in $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ be a bounded function and P be partition of A for each sub-rectangles of the partition.

$$ms(f) = \inf\{f(x) : x \in S\}$$
$$= g.l.b. of\ \ f\ on [x_{s-1}, x_s]$$

$$Ms(f) = \sup\{f(x) : x \in S\}$$
$$= l.u.b. of\ \ f\ on [x_{s-1}, x_s]$$

where $S = 1, 2, ..., n$

The lower and upper sums of f for 'p' are defined by
$L(f, p) = \sum_s m_s(f) v(s)$ and $U(f, p) = \sum_s M_s(f) v(s)$

Since $m_s < M_s$ we have $L(f, p) \leq U(f, p)$

Refinement of a partition : Let $P = (P_1, P_2, ..., P_n)$ and $P^* = (P_1^*, ..., P_n^*)$ be partition of a rectangle A in $\mathbb{R}^n$. We say that a partition P^* is a refinement of P if $P \subset P^*$.

If P_1 and P_2 are two partition of A then $P = P_1 \cup P_2$ is also a partition of A is called the common refinement of P_1 and P_2.

A function $f : A \to \mathbb{R}$ is called integrable on the rectangle A in $\mathbb{R}^n$ if $'f'$ is bounded $\therefore g.l.b$ of the set of all upper sum of $'f'$ and $l.u.b$ of the set of all lower sum of $'f'$ exist.

Let $U(f) = \inf\{U(f,p)\}$
$L(f) = \sup\{L(f,p)\}$

If $U(f) = L(f)$ is called $'f'$ is R-integrable over A.
$\therefore$ if can be written as $U(f) = L(f) = \int_A f$.

Theorem :

Let P and P' be partitions of a rectangle A in $\mathbb{R}^n$. If P' refines P then show that $L(f,p) \le L(f,P')$ and $U(f,P') \le U(f,p)$.

Proof :

Let a function $f : A \to \mathbb{R}$ is bounded on A P & P^* are two partition of A and P' is retinement to P.

Any subrectangle S of P' is union of some subrectangles $s_1, s_2, ..., s_k$ of P' and $V(S) = V(s_1) + V(s_2) + ... + V(s_k)$.

Now $m_s(f) = \inf\{f(x); x \in s\} \le \inf\{f(x); x \in s_i\}$
$\therefore m_s(f) \le m_{s_i}(f) \qquad \forall\, i = 1, ..., k$

$$L(f,p) = \sum_{s \in p} m_s(f) V(s)$$

$$\therefore m_s(f)V(s) = m_s(f)\big(V(s_1) + ... + V(s_k)\big)$$
$$\le m_{s_1}(f)V(s_1) + ... + m_{s_k}(f)V(s_k)$$

The sum of LHS for all subrectangle s_i of P' will get $L(f,P')$.

$\therefore L(f,p) \le L(f,p^1)$
Now, $M_s(f) = \sup\{f(x); x \in S\}$
$\ge \sup\{f(x); x \in S_i\}$
$M_s(f) \ge M_{s_i}(f) \qquad \forall\, i = 1, ..., K$

$$U(f,p)=\sum_{s\in p} m_s(f)V(s)$$

Now, $Ms_i(f)V(S)=Ms(f)\left(V(S_1)+V(S_2)+....+V(S_k)\right)$
$\leq Ms(f)V(s_1)+.....+M_s(f)V(s_2)+....+M_s(f)V(s_k)$

Taking the of L.H.S. for all subrectangle S_i of P' will get $U(f,P')\therefore U(f,P)\geq U(f,P')$.

Theorem :

Let P_1 & P_2 be partitions of rectangle A & $f:A\to\mathbb{R}$ be bounded function. Show that $L(f,P_2)\leq U(f,P_1)$ & $L(f,P_1)\leq \bigcup(f,P_2)$.

Proof :

Let a function $f:A\to\mathbb{R}$ be a bounded find P_1 & P_2 are any two partition of A.

Let $P=P_1\bigcup P_2$

$\therefore P$ is a refinement of both P_1 & P_2

$U(f,P)\leq U(f,P_1)$ ……….. (I)

$U(f,P)\leq U(f,P_2)$ ……….. (II)

$L(f,P)\geq L(f,P_1)$ ……….. (III)

$L(f,P)\geq L(f,P_2)$ ……….. (IV)

$\therefore$ We get $U(f,P_1)\geq U(f,P)\geq L(f,P)\geq L(f,P_2)$.

Hence $U(f,P_1)\geq L(f,P_2)$

Similarly, $U(f_2,P_2)\geq U(f,P)\geq L(f,P)\geq L(f,P_1)$.

Hence, $U(f,P_2)\geq L(f,P_1)$

Theorem :

Let a function $f:A\to\mathbb{R}$ be bounded on A then for any $\in>0,\exists$a partition P on A such that $U(f,P)<U(f)+\in$ and $L(f,P)>L(f)-\in$

Proof :

Let a function $f : A \to \mathbb{R}$ be bounded on A $U(f) = \inf\{U(f,P)\}$ and $L(f) = \sup\{L(f,P)\}$ for any $\in > 0, \exists$ partitions P_1 & P_2 of A such that $U(f,P_1) < U(f) + \in$ & $L(f,P_2) > L(f) - \in$.

Let $P = P_1 \cup P_2$ the common refinement of P_1 and P_2.

$$U(f,P) \le U(f,P_1) \le U(f) + \in$$
$$L(f,P) \ge L(f,P_2) > L(f) - \in$$
$$\therefore\ U(f,P) < U(f) + \in$$
$$L(f,P) > L(f) - \in$$

1.3 RIEMANN CRITERION

Let A be a rectangle in $\mathbb{R}^n$ A bounded function $f : A \to \mathbb{R}$ is integrable iff for every $\in > 0$, there is a partition P of A such that $U(f,P) - L(f,P) < \in$.

Proof :

Let a function $f : A \to \mathbb{R}$ is bounded.

$$U(f) = \inf\{U(f,P)\}$$
$$L(f) = \sup\{L(f,P)\}$$

Let f be integrable of A

$\therefore U(f) = L(f)$

for any $\in > 0,\ \exists$ a partition P on A such that $U(f,p) < U(f) + \in/2$ and $L(f,p) > L(f) - \in/2$.

$\therefore U(f,p) = U(f) + \in/2$ & $-L(f,p) < -L(f) + \in/2$.

$\therefore U(f,p) - L(f,P) < U(f) + \in/2 - L(f) + \in/2$.

$\therefore U(f,p) - L(f) < \in$

Conversely,

Let for any $\in > 0,\ \exists$ a partition P on A such that $U(f,p) - L(f,P) < \in$.

$$\left[U(P,f) - U(f)\right] + \left[U(f) - L(f)\right] + \left[L(f) - L(f,P)\right] < \in$$

Since $U(f,P)-U(f)\geq o,$

$U(f)-L(f)\geq o$

and $L(f)-L(f,P)\geq o$

$\therefore$ we have, $o\leq U(f)-L(f)<\in$

Since $\in$ is arbitrary, $U(f)=L(f)$

$\therefore f$ is integrable over A.

Example 1

Let A be a rectangle in $\mathbb{R}^n$ and $f:A\to\mathbb{R}$ be a constant function. Show that f is integrable and $\int_A f=C.V(A)$ for some $C\in\mathbb{R}$.

Solution :

$f(x)=C \quad \forall x\in A$

$\therefore f$ is bounded on A

Let P be a partition of A

$m_s(f)=\inf\{f(x);x\in s\}=C$

$M_s(f)=\sup\{f(x);x\in s\}=C$

$\therefore L(f,P)=\sum_S m_s(f)V(S)=C\sum_S V(S)=CV(A)$

$U(f,P)=\sum_S M_s(f)V(S)=C\sum_S V(S)=CV(A)$

$\therefore U(f)=L(f)=C\,V(A)$

$\therefore f$ is integrable over A.

$\therefore$ by Reimann criterion, $\in<0$ s.t.

$\int_A f=C.V(A)$ for some $C\in\mathbb{R}$.

Example 2 :

Let $F:[0,1]X[0,1]\to\mathbb{R}$

$$f(x,y)=\begin{cases} o & if\ x\,is\,rational \\ 1 & if\ x\,is\,irrational \end{cases}$$

Show that 'f' is not integrable.

Solution :

Let P be a partition of $[0,1]\times[0,1]$ into S subport of P.

Take any point $\exists(x_1, y_1) \in S$ such that x is rational.

$\therefore f(x,y) = o$ and $\exists(x_1, y_1) \in S$ such that x_1, is irrational

$\therefore f(x_1, y_1) = 1$

$\therefore m_s(f) = \inf\{f(x); x \in S\} = 0$

$M_s(f) = \sup\{f(x); x \in S\} = 1$

$$L(f,P) = \sum_S m_s(f)V(S) = 0$$

$$\therefore U(f,P) = \sum_S M_s(f)V(S) = 1$$

$\therefore U(f) = 1, L(f) = 0$

$\therefore U(f) \neq L(f)$

$\therefore f$ is not integrable $[0,1]\times[0,1]$

1.4 PROPERTIES OF RIEMANN INTEGRAL

1) Let $f : A \to \mathbb{R}$ be integrable and $g = f$ except at finitely many points show that g is integrable and $\int_A f = \int_A g$.

Proof :

Since f is integrable over A.

$\therefore$ by Riemann Criterion, $\exists$ a partition P of A.

Such that $U(f,P) - L(f,P) < \in$ ……… (I)

Let P' be a refinement of P, such that

1) $\forall x \in A$ with $f(x) \neq g(x)$, it belongs to 2^n subrectangles of P'

2) $V(S) < \dfrac{\in}{2^{n+1} d(u-\ell)}$

Where d = numbers of points in A at which $f \neq g$

$$u = \sup_{x\in A}\{g(x)\} - \inf_{x\in A}\{f(x)\}$$

$$\ell = \inf_{x\in A}\{g(x)\} - \sup_{x\in A}\{f(x)\}$$

$\therefore P'$ is refines P, we have

$$L(f,P) \leq L(f,P') \leq U(f,P') \leq U(f,P)$$

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

Now

$$U(g,P')-U(f,P')$$
$$=\sum_{i=1}^{d}\left(\sum\left(Ms_{ij}(g)-Ms_{ij}(f)\right)V(s_{ij})\right)$$

$\because$ On other rectangle, $f=g$ and so $Ms_{ij}(g)=Ms_{ij}(f)$.

$\because$ $Ms_{ij}(g)\le\sup_{x\in A}\{g(x)\}$ & $Ms_{ij}(f)\ge\inf_{x\in A}\{f(x)\}-Ms_{ij}(f)\le\inf_{x\in A}\{f(x)\}$

$$Ms_{ij}(g)-Ms_{ij}(f)\le u$$

$$\therefore U(g,P')-U(f,P')\le\sum_{i=1}^{d}\left(\sum_{j=1}^{2^n}u\right)V(S_{ij})$$

Let $V=\sup\left\{V(S_{ij})\right\}\le U(g,P^1)-U(f,P^1)\le\sum_{i=1}^{d}\sum_{j-1}^{2^n}uV\le d2^n u.v$ ……. (II)

Now similarly we get $L(g,P^1)-L(f,P^1)\ge d2^n\ell V$ ……... (III)
by (II) & (III) we get.

$$U(g,P^1)-L(g,P^1)\le U(f,P^1)+d2^n u\vartheta-L(f,P^1)-d2^n\ell\vartheta$$
$$\le\frac{\in}{2}+d2^n(u-\ell)V$$
$$\le\frac{\in}{2}+\frac{d2^n\in(u-\ell)}{d2^{n+1}(u-\ell)}\frac{\in}{2}+\frac{\in}{2}=\in$$
$$\therefore U(g,P^1)-L(g,P^1)<\in$$

By Reimann Criterion G is integrable by equation (II)

$$U(g,P^1)-U(f,P^1)\le d2^n uv$$
$$\therefore U(g,P^1)\le U(f,P^1)+d2^n u\vartheta$$

Note that $\int_A g\le U(g,P^1)\le U(f,P^1)+d2^n u\vartheta$

$$\le L(f,P^1)+\frac{\in}{2}+d2^n u\vartheta$$
$$<L(f,P^1)+\in\!\!/2+\frac{d2^n u\in}{d2^{n+1}(u+\ell)}$$

$$< L\left(f, P^1\right) + \in\!/2 + \in\!/2$$
$$< L\left(f, P^1\right) + \in$$
$$< \int_A f + \in$$

This is true for any $\in > 0$

$\int_A g \leq \int_A f$ …………………….. (IV)

Now $\int_A g \geq L(g, P') \geq L(f, P') + \in\!/2$

$$\geq U(f, P')$$
$$\geq \int_A f > \int_A f - \in\!/2$$

$\therefore \int_A f = \inf\{U(f, P)\}$

$\therefore \int_A g > \int_A f - \in\!/2$

$\therefore$ This is true for any $\in > 0$

$\therefore \int_A g \geq \int_A f$ ……… (V)

$\therefore$ from (IV) & (V) we get

$\int_A g = \int_A f$

2) Let $f : A \to \mathbb{R}$ be integrable, for any partition P of A and sub-rectangle S, show that

i) $m_s(f) + m_s(g) \leq m_s(f + g)$ and

ii) $M_s(f) + M_s(g) \geq M_s(f + g)$

Deduce that

$L(f, P) + L(g, P) \leq L(f + g, P)$ and

$U(f + g, P) \leq U(f, P) + U(g, P)$

Solution :

Let P be a partition of A and S be a Subrectangle

$\therefore m_s(f) = \inf\{f(x); x \in S\}$

$\Rightarrow m_s(f) \leq f(x) \forall x \in S$

Similarly $m_s(g) \le g(x) \forall x \in S$

$\therefore m_s(f) + m_s(g) \le f(x) + g(x) \forall x \in S$

$\Rightarrow m_s(f) + m_s(g)$ is lower bound of

$\{f(x) + g(x); x \in S\} = \{(f+g)(x); x \in S\}$

$\Rightarrow m_s(f) + m_s(g)$ is lower bound of

$\{f(x) + g(x); x \in S\} = \{(f+g)(x); x \in S\}$

$\Rightarrow m_s(f) + m_s(g) \le \inf\{(f+g)(x); x \in S\}$

$= m_s(f+g)$

$\therefore m_s(f) + m_s(g) \le m_s(f+g)$

ii) $Ms(f) = sub\{f(x); x \in s\}$

$\Rightarrow Ms(f) \ge f(x) \ \forall \ x \in s$

Similarly $Ms(g) \ge g(x) \forall x \in S$

$\therefore Ms(f) + Ms(g) \ge f(x) + g(x) \forall x \in S$

$\Rightarrow Ms(f) + Ms(g)$ is upper bound of

$\{f(x) + g(x); x \in S\} = \{(f+g)(x); x \in S\}$

$\Rightarrow Ms(f) + Ms(g) \ge \sup\{(f+g)(x); x \in S\} = Ms(f+g)$

$\therefore Ms(f) + Ms(g) \ge Ms(f+g)$

Hence,

$$\begin{aligned} L(f,P) + L(g,P) &= \sum_{s \in p} \big(Ms(f) + Ms(g)\big) V(S) \\ &\le \sum_{s \in p} \big(Ms(f+g)\big) V(S) \\ &< L(f+g,P) \end{aligned}$$

$\therefore L(f,P) + L(g,P) \le L(f+g,P)$

$$\begin{aligned} U(f,P) + U(g,P) &= \sum_{s} \big(Ms(f) + Ms(g)\big) V(S) \\ &\ge \sum_{s} \big(Ms(f+g)\big) V(S) \\ &\ge U(f+g,P) \end{aligned}$$

$U(f,P) + U(g,P) \ge U(f+g,P)$ Proved.

3) Let $f : A \to \mathbb{R}$ be integrable, & $g : A \to \mathbb{R}$ integrable than show that $f+g$ is integrable and $\int_A (f+g) = \int_A f + \int_A g$.

Proof :

Let P be any partition of A then

$U(f+g,P) - L(f+g,P) \le U(f,P) + U(g,P) - [L(f,P) + L(g,P)]$

$\le U(f,P) + U(g,P) - L(f,P) - L(g,P)$ …………………….. (I)

$\therefore f$ is integrable.

By Rieman interion for given $\in > 0, \exists$ a partition P, of A such that $U(f,P_1) - L(f,P_1) < \in/2$ ………………………………..… (II)

Similarly $\because g$ is integrable for $\in > 0, \exists$ a partition P_2 of A such that $U(g,P_2) - L(f,P_2) < \in/2$ ……………………………………… (III)

Then $P^* = P_1 \cup P_2$ is a refinement of both P_1 & P_2.

$\therefore L(f,P_1) \le L(f,P^*)$; $U(f,P_1) \ge U(f,P^*)$ & $L(g,P_2) \le L(f,P^*)$; $U(g,P_2) \ge U(g,P^*)$ ………………………………………….. (IV)

$\therefore \in/2 > U(f,P_1) - L(f,P_1) \ge U(f,P^*) - L(f,P^*)$

$\in/2 > U(g,P_2) - L(g,P_2) \ge U(g,P^*) - L(g,P^*)$ ……………….. (V)

The equation I is true for any partition P of A.

In general, it is true for partition P^* of A

$\therefore U(f+g,P^*) - L(f+g,P^*)$

$\le U(f,P^*) - L(f,P^*) + U(g,P^*) - L(g,P^*)$

$< \in/2 + \in/2 = \in$

$\therefore U(f+g,P^*) - L(f+g,P^*) < \in$

By Riemann Criterian $f+g$ is integrable.

Let $\in/0$ since $\int_A f = \sup\{f,P\}$ so $\exists a$ partition P such that $\int_A f < (f,P_1) + \in/2$.

Similarly $\exists a$ partition $P_2, P_3, ...P_n$ of A S

$$\int_A g < L(g, P_2) + \frac{\in}{2}$$

$$U(f, P_3) < \int_A f + \frac{\in}{2}$$

$$U(g, P_4) < \int_A g + \frac{\in}{2}$$

Let $P = P_1 \cup P_2 \cup P_3 \cup P_4$.

Then $\int_A f < (f, P_1) + \frac{\in}{2} \leq L(f, P) + \frac{\in}{2}$

Similarly $\int_A g < L(g, P) + \frac{\in}{2}$

$U(f, P) < \int_A f + \frac{\in}{2}$ and $U(g, P) < \int_A g + \frac{\in}{2}$

$$\int_A f + \int_A g - \in < L(f, P) + L(g, P) \leq L(f + g, P) \leq \int_A f + g$$

$$\leq U(f + g, P)$$

$$\leq U(f, P) + U(g, P)$$

$$< \int_A f + \frac{\in}{2} + \int_A g + \frac{\in}{2}$$

$$< \int_A f + \int_A g + \in$$

$$\therefore \int_A f + \int_A g - \in \int_A f + g < \int_A f + \int_A g - \in$$

This is true for any $\in > 0$

$$\therefore \int_A f + \int_A g \leq \int_A f + g \leq \int_A f + \int_A g \Rightarrow \int_A f + g = \int_A f + \int_A g$$

4) Let $f : A \to \mathbb{R}$ be integrable for any constant C, show that $\int_A (Cf) = C \int_A f$.

Proof :

Let $C \in \mathbb{R}$

Case 1

Let $\in > 0$ and suppose $C > 0$.

Let P be a partition of A and S be a subrectangle of P.

$$M_s(Cf) = \sup\{(Cf)(x); x \in S\}$$
$$= \sup\{Cf(x); x \in S\}$$
$$= C\sup\{f(x); x \in S\}$$
$$= CMs(f)$$

Similarly,

$$ms(Cf) = Cm_s(f)$$
$$\therefore U(Cf, P) = \sum_S Ms(Cf)v(S) = C\sum_S Ms(f)v(S)$$
$$= C\ U(f, P)$$

Similarly $L(Cf, P) = C\,L(f, P)$

$\therefore f$ is integrable for above $\in < 0, \exists$ a partition P of A such that $U(f, P) - L(f, P) < \in/C$

$$\therefore U(Cf, P) - L(Cf, P) = C\,U(f, P) - C\,L(f, P)$$
$$= C\left[U(f, P) - L(f, P)\right]$$
$$= C \times \in/C = C$$

By Riemann Criteria.

(Cf) is integrable

for $\in > 0, \exists a$ partition P of A such that

$$C\int_A f - \in = C\left(\int_A f - \in/C\right) < C\,L(f, P) = L(Cf, P)$$
$$\le \int_A Cf \le U(Cf, P)$$
$$< C\,U(f, P) < C\left(\int_A f + \in/C\right)$$
$$\therefore \left(\int_A f - \in/C\right) < \int_A Cf < C\left(\int_A f + \in/C\right) = C\int_A f + \in$$

This is true for any $\in < 0$

$$C\int_A f \le \int_A (Cf) \le C\int_A f$$
$$\therefore \int_A Cf = C\int_A f$$

Case II

Now suppose $C < 0$

Let P be a partition of A and S be any subrectangle in P.

$\therefore Ms(Cf) = C\,Ms(f)$ and

$m_s(Cf) = C\,Ms(f)$

$\therefore L(Cf,P) = C\,U(f,P)$ and

$U(Cf,P) = C\,L(f,P)$

$\therefore f$ is integrable for above $\in > 0, \exists$ a partition P of A such that

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

$\therefore U(Cf,P) - L(Cf,P) = C\,L(f,P) - C\,U(f,P)$

$= -C\left[U(f,P) - L(f,P)\right]$

$< -C\, \in\!/-C$

$< \in$

By Riemann Criteria (Cf) is integrable.

for $\in > 0, \exists$ a partition P of A such that $C\int_A f - \in < \int_A Cf < C\int_A f + \in$.

This is true for every $\in > 0$

$C\int_A f < \int_A Cf \le -C\int_A f$

$\therefore \int_A Cf = C\int_A f$

Example 3:

Let $f,g : A \to R$ be integrable & suppose $f \le g$ show that $\int_A f \le \int_A g$.

Solution :

By definition $\int_A f = \inf\{U(f,P)\}$ and $\int_A g = \inf\{U(g,P)\}$.

Let P be any partition of A & S be any subrectangle in P as $f \le g$

$m_s(f) \le m_s(g)$

$\therefore U(f,P) \le U(g,P)$

$\inf\{U(f,P)\} \le \inf\{U(g,P)\}$

This is true for any partition

$\therefore \int_A f \le \int_A g$

Example 4:

If $f : A \to \mathbb{R}$ is integrable show that if is integrable and $\left|\int_A f\right| \le \int_A |f|$.

Solution :

$\Rightarrow$ Suppose *f* is integrable first we have to show that $|f|$ is integrable.

Let P be a partition of A & S be subrectangle of P then

$$Ms(|f|) = \sup\{|f(x)|; x \in S\}$$
$$= \sup\{|f|(x); x \in S\}$$
$$= \left|\sup\{f(x); x \in S\}\right|$$
$$= |Ms(f)|$$

Similarly

$$Ms(|f|) = |Ms(f)|$$

$$U(|f|, P) = \sum_S M_s(|f|)V(S) = \sum_S |M_s(f)|V(S)$$

$$L(|f|, P) = \sum_S |m_s(f)|V(S)$$

$$\therefore \sum_P \left(|M_s(f)| - |m_s(f)|\right)V(S) \le \sum_P \left(|M_s(f)| - |m_s(f)|\right)V(S)$$
$$\le U(f, P) - L(f, P)$$

$\therefore f$ is integrable, for $\in > 0, \exists$ a partition P such that $U(f, P) - L(f, P) < \in$.

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

$\therefore$ By Riemann criteria

$|f|$ is integrable over $\mathbb{R}$.

Now $\left|\int_A F\right| = \left|\inf_P \{U(f, P)\}\right|$

$$= \left|\inf_P \sum_{S \in P} M_s(f)V(S)\right|$$
$$= \left|\inf_P \sum M_s(f)V(S)\right|$$
$$= \left|\inf_P \sum_P M_s|f|V(S)\right|$$

$$\leq \inf \sum_{P} M_s |f| V(S)$$
$$= \inf \{U(|f|, P)\}$$

$$\therefore \left| \int_A f \right| = \int_A |f|$$

Example 5:

Let $f : A \to \mathbb{R}$ and P be a partition of A show that f is integrable iff for each sub-rectangle S the function $f/_S$ which consist of f restricted to S is integrable and that in this case $\int_A f = \sum_S \int_S f/_S$.

$\Rightarrow$ Suppose $f : A \to \mathbb{R}$ is integrable.
Let P be a partition of A & S be a sub-rectangle in P.

Now to show that $f/_S ; S \to \mathbb{R}$ is integrable.
Let $\in > 0, \exists$ a partition P' of A such that $U(f, P) - L(f, P') < \in$ ($\therefore f$ is integrable)
Let $P' = P \bigcup P'$ then P_1 is refinement of both P & P'.
$\therefore U(f, P') \geq U(f, P_1)$ & $L(f, P') \leq L(f, P_1)$
$\therefore U(f, P_1) - L(f, P_1) \leq U(f, P') - L(f, P') < \in$ ………………… (I)
$\because P_1$ is refinement of P
$\because S$ is union of some subrectangle of P_1 say $S = \underset{i=1}{U} si$.
$\therefore \in > U(f, P_1) - L(f, P_1) = \sum_{S \in P_1} (M_s(f) - m_s(f)) V(S)$ for all rectangle.

$$\geq \sum_{i=1}^{k} (Ms_i(f) - m_{s_i}(f)) V(S)$$
$$= U\left(f/_S, P\right) - L\left(f/_S, P\right)$$

$\therefore$ By Riemann Criterion

$\therefore f/_S$ is integrable.

Conversely, Suppose $f/_S$ is integrable for each $S \in P$.

To show that f is integrable.
Let $\in > 0, \exists$ partition P_S of S such that

$U\left(f/_S, P_S\right) - L\left(f/_S, P_S\right) < \in/k$ .... (II)

$\therefore f/_S$ is integrable for each $S \in P$ where K is number of rectangle in P.

Let P^1 be the partition of A obtained by taking all the subrectangle defined in the partition P_S.

There is a refinement P_S^1 of P_S containing subrectangles in P^1.

$\therefore U\left(f/s, P_S^1\right) - L\left(f/s, P_S^1\right) < \in/k$ .... (III)

$$\therefore U\left(f, P^1\right) - L\left(f, P^1\right) = \sum_{S^1 \in P^1} \left(M_{S^1}(f) - m_{S^1}(f)\right) V\left(S^1\right)$$

$$= \sum_{S \in P} \left(\sum_{S^1 \in P_S^1} \left(Ms^1(f) - m_{s^1}(f)\right) V\left(S^1\right) \right)$$

$$= \sum_{S \in P} \left(U\left(f/s, P_S^1\right) - L\left(f/s, P_S^1\right)\right)$$

$$< \sum_{S \in P} \in/k$$

$$< k, \in/k < \in$$

$\therefore$ By Riemann Criterian f is integrable.

Let $\in > 0$

$$\sum_{S \in P} \left(\int_S f/S - \in/k \right) < \sum_{S \in P} L\left(f/S, P_S\right)$$

$$< \sum_{S \in P} \left(\sum_{S^1 \in P_S^1} m_s^1(f) V(S') \right)$$

Let P^1 be a partition of A, obtained by taking allthe subrectangle defined in P_S.

$$\therefore \sum_{S\in P}\left(\int_S f/S - \in/k\right) < \sum_{S^1\in P^1}\left(m_{s^1}(f)\right)V\left(S^1\right)$$

$$< L\left(f, P^1\right) < \int_A f < U\left(f, P^1\right)$$

$$= \sum_{S^1\in P^1} M_{s^1}(f)V\left(S^1\right)$$

$$= \sum_{S\in P}\left(\sum_{S^1\in P^1} M_{s^1}(f)V\left(S^1\right)\right)$$

$$\therefore \sum_{S\in P}\left(U(f/S, P_S)\right) < \sum_{S\in P}\left(\int_S f/S + C\!\!\diagup_k\right)$$

$$\therefore \sum_{S\in P}\int f/S - \in C\int_A f < \sum_{S\in P}\int f/S + \in$$

This is true for all $\in > 0$

$$\therefore \sum_{S\in P}\int f/S \le \int f \le \sum_{S\in P}\int_S f/S$$

$$\therefore \int_A f = \sum_{S\in P}\int_S f/S$$

Example 6:

Let $f : A \to \mathbb{R}$ be a continues function show that f is integrable on A.

Solution :

Let $f : A \to \mathbb{R}$ be a continuous function to show that f is integrable.

Let $\in > 0$, since A is closed rectangle it is closed and bounded in $\mathbb{R}^n$.

$\therefore A$ is compact.

$\because f$ is continuous function on compact set $\Rightarrow f$ is uniformly continuously on $\mathbb{R}$.

$\therefore$ for the above $\in > 0, \exists \delta > 0$ such that $\forall x, g \in A,$ $\|x - y\| < \delta \Rightarrow \left|f(x) - f(y)\right| < \in/V(A)$.

Let P be a partition of A such that side length of each subrectangle is less than $\delta/\sqrt{n}$.

If $x, y \in S$ for some subrectangles S then

$$\|x - y\| = \sqrt{\left(x_1 - y_1\right)^2 + ... + \left(x_n - y_n\right)^2}$$

$$< \sqrt{n\left(S \middle/ \sqrt{n} \right)^2} = \delta$$

$$\left| f(x) - f(y) \right| < \in / V(A)$$

$\because S$ is compact
$\therefore f$ is continuous
$\therefore f$ attains its bound in S.

Let $S_1, S_2, \ldots\ldots, S_k$ be the subrectangle in A. Then for $1 < i < k, \exists\, x_i, y_i \in S_i$ such that $Ms_i(f) = f(x_i) m_{si}(f) = f(y_i)$.

$$\therefore U(f,P) - L(f,P) = \sum_{i=1}^{k} \left(Ms_i(f) - m_{s_i}(f) \right) V(S_i)$$

$$= \sum_{i=1}^{k} \left(f(x_i) - f(y_i) \right) V(S_i)$$

$$< \sum_{i=1}^{k} \frac{\in}{V(A)} V(S_i) < \frac{\in}{V(A)} \sum_{V(A)}^{k} V(S_i)$$

$$< \frac{\in}{V(A)} V(A) < \in$$

$\therefore$ By Riemann Criterion *f* is integrable.

1.5 REVIEW

After reading this chapter you would be knowing.

- ❖ Defining R-integral over a rectangle in $\mathbb{R}^n$
- ❖ Properties of R-integrals
- ❖ R-integrabal functions
- ❖ Continuity of functions using $\mathbb{R}$ -intervals.

1.6 UNIT END EXERCISE

I) Let $f; [0,1] \times [0,1] \to \mathbb{R}$ be defined by

$$f(x,y) = 0 \; if \; 0 \le y \le 1/3$$
$$= 3 \; if \; 1/3 \le y \le 1$$

show that *f* is integrable.

II) Let Q be rectangle in $\mathbb{R}^n$ & $f; Q \to \mathbb{R}$ be any bounded function.

a) Show that for any partition P of Q $L(f,P) < U(f,P)$

b) Show that upper integral of function *f* exit.

III) Let f be a continuous non-negative function on $[0,1]$ and suppose there exist $x_0 \in [a,b]$ such that $f(x_0) > 0$ show that $\int_0 f(x)\,dx > a$.

IV) Let f be integrable on $[a,b]$ and $F:[a,b] \to \mathbb{R}$ and $F^1(x) = f(x)$ then prove that $\int_a f(x)\,dx = F(b) - F(a)$

V) Which of the following functions are Riemann integrable over $[0,1]$. Justify your answer.

a) The characteristic function of the set of rational number in $[0,1]$.

b) $f(x) = x\sin y_x$ for $0 < x < 1$

$f(0) = 3$

VI) Prove that if *f* is $\mathbb{R}$-integrable then $|f|$ is also R-integrable is the converse true? Justify your answer.

VII) Show that a monotone function defined on an interval $[a,b]$ is R-inegrable.

VIII) A function $f;[0,1] \to \mathbb{R}$ is defined as $f(x) = \frac{1}{3^{n-1}}\ \forall \frac{1}{3^n} < x \le \frac{1}{3^{n-1}}$

where $n \in \mathbb{N}$

$f(0) = 0$

show that *f* is R-integrable on $[0,1]$ & calculate $-\int_0^1 f(x)\,dx$.

IX) $f(x) = x\lfloor x \rfloor\ \ \forall x \in [1,3]$ where $\lfloor x \rfloor$ denotes the greatest integer not greater than x show that *f* is R-integrable on $[1,3]$.

X) A function $f;[a,b] \to \mathbb{R}$ is continuous on $[a,b]$ $f(x) \ge 0$ $\forall x \in [a,b]$ and $\int_a^b f(x)\,dx = 0$ show that $f(x) = 0\ \ \forall x \in [a,b]$.

❖❖❖❖

2

MEASURE ZERO SET

Unit Structure :

2.1 INTRODUCTION

As we have seen, we cannot tell if a function is Riemann integrable or not merely by counting its discontinuities one possible alternative is to look at how much space the discontinuities take up. Our question then becomes : (i) How can one tell rigorously, how much space a set takes up. Is there a useful definition that will concide with our intuitive understanding of volume or area?

At the same time we will develop a general measure theory which serves as the basis of contemporary analysis.

In this introductory chapter we set for the some basic concepts of measure theory.

2.2 MEASURE ZERO SET

Definition :

A subset 'A' of $\mathbb{R}^n$ said to have measure 'O' if for every $\in > 0$ there is a cover $\{U_1, U_2 ....\}$ of A by closed rectangles such that the total volume $\sum_{i=1}^{\infty} v(Ui) < \in$.

Theorem :

A function 'f' is Riemann integrable iff 'f' is discontinuous on a set of Measure zero.

A function is said to have a property of Continuous almost everywhere if the set on which the property does not hold has measure zero. Thus, the statement of the theorem is that 'f' is Riemann integrable if and only if it is continuous atmost everywhere.

Recall positive measure : A measure function $u : M \to [0, \infty]$ such that $V\left(\bigcup_{i=1}^{\infty} u_i\right) = \sum_{i=1}^{\infty} V(u_i)$.

Example 1:

1) "Counting Measure" : Let X be any set and $M = P(X)$ the set of all subsets : If $E \subset X$ is finite, then $\mu(E) = \eta(E)$ if $E \subset X$ is infinite, then $\mu(E) = \infty$
2) "Unit mass to x_0 - Dirac delta function" : Let X be any set and $M = P(X)$ choose $x_0 \in X$ set.
$\mu(E) = 1 \;\; if \;\; x_0 \in E$
$\quad\quad = 0 \;\; if \;\; x_0 \notin E$

Example 2:

Show that A has measure zero if and only if there is countable collection of open rectangle $V_1, V_2, ...$ such that $A \subseteq \bigcup V_i$ and $\sum V(v_i) < \in$.

Solution :

Suppose A has measure zero.

For $\in > 0, \exists$ countable collection of closed rectangle $V_1, V_2, ...$ such that $A \subseteq \bigcup_{i=1}^{\infty} V_i$ and $\sum_{i=1}^{\infty} V(V_i) < \frac{\in}{2}$.

For each i, choose a rectangle u_i such that $u_i \supseteq v_i$ and $V(u_i) \le 2V(v_i)$.

Then $A \subseteq \bigcup_{i=1}^{\infty} v_i \subseteq \bigcup_{i=1}^{\infty} u_i$ and $\sum_{i=1}^{\infty} V(u_i) \le \sum_{i=1}^{\infty} V(u_i) \le \sum_{i=1}^{\infty} 2V(v_i)$
$\le 2\sum_{i=1}^{\infty} v(u_i) < 2\frac{\in}{2} = \in$

Note that : u_i are open rectangles in $\mathbb{R}^n$ conversely, Suppose for $\in > 0, \exists$ countable collection of open rectangles $u_1, u_2, ...$ such that $A \subseteq \bigcup_{i=1}^{\infty} u_i$ and $\sum_{i=1}^{\infty} V(u_i) < \in$.

For each i, consider $V_i = \overline{u_i}$ then V_i is a closed rectangle and $V(v_i) = V(u_i)$.

Then $A \subseteq \bigcup_{i=1}^{\infty} u_i \subseteq \bigcup_{i=1}^{\infty} v_i$ and $\sum_{i=1}^{\infty} V(v_i) = \sum_{i=1}^{\infty} V(u_i) < \in$.

A has measure zero.

Note : Therefore we can replace closed rectangle with open rectangles in definition of measure zero sets.

Example 3:

Show that a set with finitely many points has measure zero.

Solution :

Let $A = \{a_1, ..., a_m\}$ be finite subset of $\mathbb{R}^n$.

Let $\in > 0, a_i = (a_{i1}, a_{i2}, ..., a_{in})$ and

$$Vi = \left[a_{i1} - \frac{1}{2}\left(\frac{\in}{2^{i+1}}\right)^{1/n}, a_{i1} + \frac{1}{2}\left(\frac{\in}{2^{i+1}}\right)^{1/n}\right] \times ...$$

$$... \times \left[a_{in} - \frac{1}{2}\left(\frac{\in}{2^{i+1}}\right)^{1/n}, a_{in} + \frac{1}{2}\left(\frac{\in}{2^{i+1}}\right)^{1/n}\right]$$

Then $V(Vi) = \prod_{i=1}^{n}\left(\frac{\in}{2^{i+1}}\right)^{\frac{1}{n}} = \frac{\in}{2^{i+1}}$

Clearly $a_i \in Vi$ for $1 \le i \le m$

$\therefore A \subseteq \bigcup_{i=1}^{m} Vi$ and $\sum_{i=1}^{m} V(Vi) = \sum_{i=1}^{m} \frac{\in}{2^{i+1}} < \in \cdot \sum_{i=1}^{\infty} \frac{1}{2^{i+1}} < \in \cdot \frac{1}{2} < \in$

$\therefore$ By definition of measure of zero

$\therefore$ A has measure of zero.

Example 4:

If $A = A_1 \cup A_2 \cup A_3 \cup ....$ and each Ai has measure zero, then show that A has measure zero.

Solution :

Let $\in > 0$ and $A = A_1 \cup A_2 \cup ....$ with each Ai has measure zero.

$\because$ Each Ai has measure zero for $i=1,2,....$ $\exists$ a cover $\{u_{i1},U_{i2},....,U_{in}\}$ of Ai

By closed rectangle such that $\sum_{i=1}^{\infty} V(u_{ii}) < \frac{\in}{2^i}, i=1,2,....$

Then the collection of U_{ii} is cover A

$\therefore \sum_{i=1}^{\infty} V(V_i) < \sum_{i=1}^{\infty} \frac{\in}{2^i} < \in$

Thus $A = A_1 \cup A_2 \cup A_n....$ has measure zero.

Example 5:

Let $A \subset \mathbb{R}^n$ be a Rectangle show that A does not have measure zero. But ∂A has measure zero.

Proof :

Suppose A has measure zero.

$\because$ A is a rectangle in $\mathbb{R}^n$

$\therefore V(A) > 0$

Choose $\in > 0$ such that $\in < V(A)$ …………………….. (I)

$\because$ A has measure zero

$\exists$ countable collection of open rectangle $\{u_i\}$ such that $A \subseteq \bigcup_{i=1}^{\infty} u_i$ and $\sum V(u_i) < \in$.

$\because$ A is compact

This open cover has a finite subcover after renaming. We may assume that $\{u_1, u_2, ...u_k\}$ is subcover of the cover $\{u_i\}$.

$\therefore A \subseteq \bigcup_{i=1}^{\infty} u_i$.

Let P be partition of A that contains all the vertices all u_i's $i=1$ to k. Let $S_1, S_2, ..., S_n$ denote the subrectangle of partitions.

$\therefore V(A) = \sum_{j=1}^{n} V(S_j) \le \sum_{i=1}^{k} V(u_i) < \sum_{i=1}^{\infty} V(u_i) < \in$

which is a contradiction to (I)

$\therefore$ A does not have measure zero.

Note that ∂A is a finite union of set of the form $B = [a_1, b_1] \times [a_i, b_i] \times ..... \times [a_n, b_n], \forall$. B can be covered by are closed rectangle. $B_\delta = [a_1, b_1] \times ..... \times [a_i, a_{i+\delta}] \times ..... \times [a_n, b_n]$.

Then $V(B_\delta)$ depend on δ and $V(B_\delta) \to 0$ as $\delta \to 0$.

$\therefore B_\delta$ has measure zero

$\therefore$ Boundary of A (∂A) is finite union of measure zero.

$\therefore \partial A$ has measur5e zero.

Example 6:

Let $A \subset \mathbb{R}^n$ with $A^\circ \neq \varnothing$. Show that A does not measure zero.

Solution :

Let $A \subset \mathbb{R}^n$, with $A^\circ \neq \varnothing$

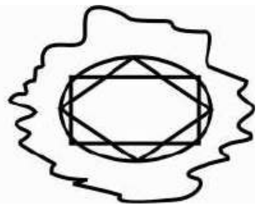
Let $x \in A^\circ$

$\therefore \exists\ r > 0$, such that $B(x, r) \leq A$, But

$B(x, r) = \{y \in A; \|y - x\| < r\}$

$$= \left\{ y \in A; \sum_{i=1}^{n} |y_i - x_i| < r \right\}$$

If f A has measure zero
then $B(x, r)$ has measure zero
which is not possible as
$B(x, r)$ is Rectangle



$\therefore A$ does not have measure zero.

Example 7:

Show that the closed interval $[a, b]$ does not have measure zero.

Solution :

Suppose $\{u_i\}_{i=1}$ be a cover of $[a, b]$ by open intervals.

$\because [a, b]$ is compact this open cover has a finite subcover.

After renaming, we may assume $\{u_1, u_2, ..., u_n\}$ is the subcover of $\{u_i\}$ of $[a, b]$.

We may assume each u_i intersect $[a,b]$ (otherwise replace u_i with $u_i \cap [a,b]$)

Let $u = \bigcup_{i=1}^{n} u_i$

If u is not connected then $[a,b]$ is contained in one of connected component of u.

$\Rightarrow [a,b] \subseteq u_i$ for some i
$\therefore [a,b] \cap u_j = \varnothing$ for $i \neq j$
Which is not possible
$\therefore u$ is connected
$\Rightarrow u$ is an open interval say $u = (c,d)$ Then as $[a,b) \subseteq u = (c,d)$
$\Rightarrow \sum V(u_i) = d - c > b - a$

In particular we cannot find an open cover of $[a,b]$ with total length of the cover $< \frac{b-a}{2}$.
$\therefore [a,b]$ does not have measure zero.

Example 8:

If $A \subseteq [0,1]$ is the union of all open intervals (a_i, b_i) such that each rational number in (0,1) is contained in some (a_i, b_i). If $T = \sum_{i=1}^{\infty} (bi - ai) < 1$ then show that the boundary of A does not have measure zero.

Solution :

We first show that $\partial A = [0,1] \setminus A$
Note that $\partial A = \bar{A} \setminus A^\circ$
$\because A$ is open $\Rightarrow A^\circ = A$
Also $Q \cap [0,1] \subseteq A$
$\therefore \bar{Q} \cap \overline{[0,1]} \subseteq \bar{A}$
$\therefore [0,1] \subseteq \bar{A}$
But $A \subseteq [0,1] \Rightarrow \bar{A} \subseteq [0,1]$
$\therefore \bar{A} = [0,1]$
$\therefore \partial A = [0,1] \setminus A$

Let $\in = 1 - T > 0$

If ∂A has measure zero then since $\in > 0, \exists$ a cover of ∂A with open intervals such that sum of length of intervals $< 1 - T$
$\because \partial A$ is closed and bounded
$\Rightarrow \partial A$ is compact
$\Rightarrow \exists$ finite subcover $\{u_i\}_{i=1}^{n}$ for ∂A
$\therefore \sum \ell(u_i) < 1 - T$

Note that $\left\{u_i; 1 \le i \le n; (a_i, b_i)_{i=1}^{\infty}\right\}$ cover $[0,1]$ and sum of lengths of these open intervals is less than $1 - T + T = 1$ which is not possible as $[0,1] \subseteq \cup \left\{u_i;\ 1 \le i \le n;\ (a_i, b_i)_{i=1}^{\infty}\right\} \therefore \partial A$ does not have measure zero.

2.3 DEFINITION

A subset 'A' of $\mathbb{R}^n$ has content 'O' if for every $\in > 0$, there is a finite cover $\{u_1, u_2, ..., u_n\}$ of A by closed rectangles such that $\sum_{i=1}^{n} V(u_i) < \in$

Remark :
1) If A has content O, then A clearly has measure O.
2) Open rectangles can be used instead of closed rectangles in the definition.

Example 9:
If A is compact and has measure zero then show that A has content zero.

Solution :
Let A be a compact set in $\mathbb{R}^n$
Suppose that A has measure zero

$\therefore \exists$ a cover $\{u_1, u_2, ...\}$ of A such that $\sum_{i=1}^{\infty} V(u_i) < \in$ for every $\in > 0$.
$\because A$ is compact, a finite number $u_1, u_2, ..., u_n$ of u_i also covers A and $\sum_{i=1}^{n} V(u_i) < \sum_{i=1}^{\infty} V(u_i) < \in$
$\therefore A$ has content zero.

Example 10 :

Give one example that a set A has measure zero but A does not have content zero.

Solution :

Let $A = [0,1] \cap Q$

Then A is countable

$\Rightarrow A$ has measure zero

Now to show that A does not have content zero.

Let $\{[a_i, b_i); 1 \le i \le n\}$ be cover of A

$\therefore A \subseteq [a_i, b_i] \cup .... \cup [a_n, b_n]$

$\therefore \bar{A} \subseteq [a_1, b_1] \cup .... \cup [a_n, b_n]$

But $\bar{A} = [0,1]$

$\therefore \sum_{i=1}^{n} \ell([a_i, b_i)) > 1$

In particular, we cannot find a finite cover for A such that $\sum_{i=1}^{n} \ell(a_i, b_i) < 1/2$

$\therefore A$ does not have content zero.

Example 11:

Show that an unbounded set cannot have content zero.

Solution :

Let $A \subseteq \mathbb{R}^n$ be an unbounded set.

To show that A does not have content zero

Suppose A has content zero for $\in > 0, \exists$ finite cover of closed rectangles $\{u_i\}_{i=1}^{k}$ of A such that $A \subseteq \bigcup_{i=1}^{k} u_i$ and $\sum_{i=1}^{k} V(u_i) < \in$.

Let $u_i = [a_{i1}, b_{i1}] \times .... \times [a_{in}, b_{in}]$

Let $a_i = \min\{a_{1i}, a_{2i}, ...a_{ki}\}$

$b_i = \max\{b_{1i}, b_{2i}, ...b_{ki}\}$

then $\cup u_i \subseteq [a_1, b_1] \times .... \times [a_n, b_n]$

$\therefore A \subseteq [a_1, b_1] \times .... \times [a_n, b_n]$

$\therefore A$ is bounded

Which is contradiction

$\therefore A$ does not have content zero.

Example 12:

$f : A \to \mathbb{R}$ is non-negative and $\int_A f = 0$ where A is rectangle, then show that $\{x \in A; f(x) \neq 0\}$ has measure zero.

Solution :

For $n \in \mathbb{N}$, $A_n = \{x \in A; f(x) < 1/n\}$

Note that $\{x \in A, f(x) \neq 0\} = \{x \in A; F(x) > 0\}$

$\{\because f$ is non-negative$\}$

$= \bigcup_{n=1}^{\infty} \{x \in A; f(x) > 1/n\} = \bigcup_{n=1}^{\infty} A_n$

We have to show that A_n has measure zero

$\because \int_A f = 0$ and $\int_A f = \inf_P \{U(f, P)\} = 0$ for $\in > 0, \exists$ a partition P such that $U(f, P) < \in/n$

Let S be a subrectangle in P

if $S \cap A_n \neq \varnothing \Rightarrow M_s(f) \leq 1/n$

clearly $\{S \in P; S \cap A_n \neq \varnothing\}$ covers A_n and

$\sum_{S \in P} \frac{1}{n} V(S) < \sum_{S \in P} M_s(f) V(S) \left(\because M_s(f) > \frac{1}{n} \right)$

$< \bigcup(f, P) < \in/n$

$\therefore \sum_{\substack{S \cap A_n \neq \varnothing \\ s \in p}} V(S) < \in$

By definition A_n has content zero

$\Rightarrow A_n$ has measure zero

$\therefore \{x \in A, f(x) \neq 0\}$ is countable union of measure zero set.

$\therefore \{x \in A; f(x) \neq 0\}$ has measure zero.

* Oscillation $o(f, a)$ of 'f' at a

$\therefore$ for $\delta > 0$, Let $M(a, f, \delta) = \sup\{f(x); x \in A \,\&\, |x - a| < \delta\}$

$m(a, f, \delta) = \inf\{f(x); x \in A \,\&\, |x - a| < \delta\}$

The oscillation $o(f, a)$ of f at a defined by $o(f, a) = \lim_{\delta \to o} \left(M(a, f, \delta) - m(a, f, \delta)\right)$

This limit always exist since $M(a,f,\delta)-m(a,f,\delta)$ decreases as δ decreases.

Theorem :

Let A be a closed rectangle and let $f: A \to \mathbb{R}$ be a bounded function such that $O(f,x) < \in$ for all $x \in A$ show that there is a partition P of A with $U(f,P) - L(f,P) < \in \cdot V(A)$.

Proof :

Let $x \in A \Rightarrow U(f,x) < \in \Rightarrow \lim_{\delta \to O}\left(M(x,f,\delta) - m(x,f,\delta)\right) < \in$

$\therefore \exists$ a closed rectangle u_x containing x in its interior such that $M_{u_x} - M_{u_x} < \in$ by definition of oscillation.

$\therefore \{u_x ; x \in A\}$ is a cover of A

$\therefore A$ is compact

$\Rightarrow$ This cover has a finite subcover say $\{u_{x1}, u_{x2}, ..., u_{xk}\}$

$\therefore A \subseteq \bigcup_{i=1}^{k} u_{x_i}$.

Let P be a partition for A such that there each subrectangle 'S' of P is contained in some u_{x_i} then $M_s(f) - m_s(f) < \in$ for each subrectangle 'S' in f

$$\therefore U(f,P) - L(f,P) = \sum_{S \in P}\left(M_s(f) - m_s(f)\right)V(S)$$

$$< \in \sum_{S \in P} V(S)$$

$$< \in \cdot V(A)$$

2.4 LEBESGUE THEOREM (ONLY STATEMENT)

Let A be a closed rectangle and $f: A \to \mathbb{R}$ is bounded function. Let $B = \{x ; f$ is not continuous at x$\}$. Then f is integrable iff B is a set of measure zero

2.5 CHARACTERISTIC FUNCTION

Let $C \subseteq \mathbb{R}^n$. The characteristics function χ_c of C is defined by

$$\chi_c(x) = 1 \quad if \ x \in C$$
$$= 0 \quad if \ x \notin C$$

If $C \subset A$ where A is a closed rectangle and $f : A \to \mathbb{R}$ is bounded then $\int_C f$ is defined as $\int_C f\chi_c$ provided $\int f \cdot \chi_c$ is integrable [i.e. if f and χ_c are integrable]

Theorem :

Let A be a closed rectangle and $C \subset A$. Show that the function $\chi_c : A \to$ □ is integrable if and only if ∂C has measure zero.

Proof :

To show that $\chi_C : A \to \mathbb{R}$ is integrable iff ∂C has measure zero.

By Lebesgue theorem, it is enough to show that $\partial C = \{x \in A : \chi_c$ is discontinuous}

Let $a \in C^\circ \Rightarrow \exists$ an open rectangle 'u' containing a such that $u \subseteq C$
$\therefore \chi_c(n) = 1 \ \forall n \in U$
$\Rightarrow \chi_c$ is continuous at a.

Let $a \in Ext(c) =$ Exterior of C
[By definition union of all open sets disjoints from C]
Ext (C) is an open set
$\exists$ an open rectangle u containing such that $U \subseteq Ext(c)$
$\therefore \chi_c(n) = 0 \ \forall n \in u$
$\Rightarrow \chi_c$ is continuous at a
If $a \notin \partial c$ then χ_c is continous at a ……………………. (I)

Let $a \in \partial c \Rightarrow$ for any open rectangle U with a in its interior contains a point $y \in C^\circ$ & a point $z \in \mathbb{R}^n | c$
$\therefore \chi_c(y) = 1 \,\&\, \chi_c(z) = 0$
$\therefore \chi_c$ is not continuous at a
$\therefore \partial c = \{x \in A : \chi_c$ is discontinuous at x }
$\therefore$ By Lebesgue Theorem.
χ_c is interrable if and only if ∂c has measure zero.

Theorem :

Let A be a closed rectangle and $C \subset A$

If C is bounded set of measure zero and $\int_A \chi_c$ exist then show that $\int_A \chi_c = 0$.

Proof :

$C \subseteq A$ be a bounded set with measure zero.

Suppose $\int_A \chi_c$ exist $\Rightarrow \chi_c$ is integral

To show that $\int_A \chi_c = 0$

Let P be a partition of A and S be a subrectangle in P.

$\because S$ does not have measure zero

$\Rightarrow S \not\subseteq C$

$\Rightarrow \exists x \in S$ but $x \notin C$

$\therefore \chi_c(x) = 0$

$\Rightarrow m_s(\chi_c) = 0$

This is true for any subrectangle S in P

$\therefore L(\chi_c, P) = \sum m_s(\chi_c) V(C) = 0$

This is true for any partition P

$\therefore \int_A \chi_c = \sup\{L(\chi_c, P); P$ is partition of}

$\int_A \chi_c = O$

2.6 FUBINI'S THEOREM

Fubini's Theorem reduces the computation of integrals over closed rectangles in $\mathbb{R}^n, n > 1$to the computation of integrals over closed intervals in $\mathbb{R}$. Fubini's Theorem is critically important as it gives us a method to evaluate double integrals over rectangles without having to use the definition of a double integral directly.

If $f : A \rightarrow R$ is a bounded function on a closed rectangle then the least upper bound of all lower sum and the greatest lower bound of all upper sums exist. They are called the lower integral and upper integral of f and is denoted by $L\int_A F$ and $U\int_A F$ respectively.

Fubini's Theorem

Statement : Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^n$ be closed rectangles and let $f : A \times B \to \mathbb{R}$ be integrable for $x \in A$, Let $g_x : B \to \mathbb{R}$ be defined by $g_x(y) = F(x, y)$ and let

$$\ell(x) = L\int_B g_x = L\int_B f(x, y)\,dy$$

$$u(x) = U\int_B g_x = U\int_B f(x, y)\,dy$$

Then ℓ and μ are integable on A and $\int_{A\times B} f = \int_A L = \int_A \left(L\int_B f(x)\,dy \right) dx$

$$\int_{A\times B} f = \int_A u(x)\,dx = \int_A \left(U\int_B f(x, y)\,dy \right) dx$$

Proof :

Let P_A be a partition of A and P_B be a partition of B. Then $P = (P_A, P_B)$ is a partition of $A \times B$

Let S_A be a subrectangle in P_A and S_B be a subrectangle in P_B

Then by definition,

$S = S_A \times S_B$ is a subrectangle in P

$$L(f_1 P) = \sum_{S \in P} m_s(f) V(S)$$

$$= \sum_{S_B \in P_B} m_{s_A \times s_B}(f) V(S_A \times S_B)$$

$$= \sum_{S_A \in P_A} \left(\sum_{S_B \in P_B} m_{s_A \times s_B}(f) V(S_B) \right) V(S_A) \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(I)}$$

For $x \in S_A, m_{s_A \times s_B}(f) \subseteq M_{s_B}(g_x)$

$\therefore$ For $x \in S_{A,}$

$$\therefore \sum_{S_B \in P_B} m_{s_A \times s_B} V(S_A) \cdot V(S_B) \le \sum m_{s_B}(g_x) V(S_B)$$

$$= L(g_x, P_B) \le L\int_B g_x = L(x)$$

This is true for any $x \in A$

$$\therefore L(f, P) = \sum_{S_A \in P_A} \left(\sum_{S_B \in P_B} m_{s_A \times s_B}(f) V(S_B) \right) V(S_A)$$

$$\le \sum_{S_A \in P_A} m_{s_A}(L(x)) V(S_A)$$

$$= L(\ell(x), P_A) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(II)}$$

$\therefore$ From (I) & (II)

$L(f,P) \le (L(x),P_A)$ ………………………………………… (III)

Now $U(f,P) = \sum_{S\in P} M_S(f)V(s)$

$$= \sum_{\substack{S_A\in P_A\\ S_B\in P_B}} M_{S_A\times S_B}(f)V(S_A\times S_B)$$

$$= \sum_{S_A\in P_A}\left(\sum_{S_B\in P_B} M_{S_A\times S_B}(f)V(S_B)\right)V(S_A) \quad \text{…………….. (IV)}$$

For $x\in S_A, M_{S_A\times S_B}(f) \ge M_{S_B}(g_x)$

$\therefore$ For $x\in S_A$,

$$\sum_{S_B\in P_B} M_{S_A\times S_B}(f)V(S_B) \ge \sum_{S_B\in P_B} M_{S_B}(g_x)V(S_B)$$

$$= u(g_x,P_B) \ge u\int_B g_x = \mu(x)$$

This is true for any $x\in A$.

$$\sum_{S_A\in P_A}\left(\sum_{S_B\in P_B} M_{S_A\times S_B}(f)V(S_B)\right)V(S_A)$$

$$\ge \sum_{S_A\in P_A} M_{S_A}(u(x))V(S_A)$$

$$= (u(x),P_A) \quad \text{……………………………………….. (V)}$$

from (IV) & (V)

$$U(f,P) \ge U(u(x),P_A) \quad \text{……………………………. (VI)}$$

$\therefore$ By (III) & (VI)

$$L(f,P) \le L(\ell(x),P_A) \le u(L(x),P_A)$$

$$\le u(\ell(x),P_A) \le U(f,P) \quad \text{………………………… (VII)}$$

Also

$$L(f,P) \le L(\ell(x),P_A) \le L(\mu(x),P_A) \le u(\ell(x),P_A) \quad \text{…………. (VIII)}$$

$\therefore f$ is integrable

$$\sup_P\{L(f,P)\} = \inf_P\{U(f,P)\} = \int_{A\times B} f$$

$$\Rightarrow \sup_{P_A}\{L(\ell(x),P_A)\} = \inf_{P_B}\{u(\ell(x),P_A)\} = \int_{A\times B} f$$

$\therefore \ell(x)$ is integrable

$$\int_{A\times B} f = \int_A \ell(x) = \int_A \left(L\int_B f(x,y) \right) dx \text{ .... (IX)}$$

Also by (VIII) & (IX)

$$\sup_{P_A}\left\{L\left(L(x),P_A\right)\right\} = \inf_{P_A}\left\{U\left(u(x),P_A\right)\right\} = \int_{A\times B} f$$

$\therefore u(x)$ is integrable.

$$\Rightarrow \int_{A\times B} f = \int_A u(x)\,dx = \int_A \left(U\int_B f(x,y) \right) dx$$

Hence Proved

Remark :

The Fubini's theorem is a result which gives conditions under which it is possible to compute a double integral using interated integrals, As a consequence if allows the under integration to be changed in iterated integrals.

$$\int_{A\times B} f = \int_B \left(L\int_B f(x,y)\,dx \right) dy$$

$$= \int_B \left(U\int_A f(x,y)\,dx \right) dy$$

These integrals are called iterated integrals.

Example 13:

Using Fubini's theorem show that $D_{12}f = D_{21}f$ if $D_{12}(f)$ and $D_{21}(f)$ are continuous.

Solution :

$\Rightarrow$ Let $A \subseteq R$ and $f : A \to \mathbb{R}$ continuous

T.P.T $D_{12}f = D_{21}f$

Suppose $D_{12}f \neq D_{21}f$

$\therefore \exists\, x_0, y_0$ in domain of f such that

$\left(D_{12}f(a) - D_{21}f(a)\right) \neq 0$

without loss of generality, $\left(D_{12}f(a) - D_{21}f(a)\right) > 0$ or

$\left(D_{12}f - D_{21}f\right)(a) > 0$ .... (I)

$$\therefore \int_A (D_{12} f - D_{21} f)(x,g) > 0$$

Let $A = [a,b] \times [c,d]$

$\therefore$ By Fubini's Theorem

$$\int_A D_{21} f(x,y) = \int_c^d \int_a^b D_{21} f(x,y)\, dx\, dy$$

$$= \int_c^d \left(D_2 f(b,y) - D_2 f(g,y) \right) dy$$

$$= f(b,d) - f(b,c) - f(a,d) + f(a,c)$$

Similarly,

$$\int_A D_{12} f(x,y) = f(b,d) - f(b,c) - f(a,d) + f(a,c)$$

$$\therefore \int_A D_{21} f(x,y) = \int_A D_{12} f(x,y)$$

$$\Rightarrow \int_A (D_{21} f - D_{12} f)(x,y) = 0$$

Which is contradiction to (I)

$\boxed{D_{12} f = D_{21} f}$ proved

Example 14:

Use Fubini's Theorem to compute the following integrals.

1) $$I = \int_0^1 \int_0^{\sqrt{1+x^2}} \frac{dy.dx}{1+x^2+y^2}$$

Solution :

$$I = \int_0^1 \int_0^{\sqrt{1+x^2}} \frac{dy.dx}{1+x^2+y^2}$$

$$= \int_0^1 dx \int_0^{\sqrt{1+x^2}} \frac{dy}{1+x^2+y^2}$$

$$= \int_0^1 dx \left[\frac{1}{\sqrt{1+x^2}} \tan^{-1} \frac{y}{\sqrt{1+x^2}} \right]_0^{\sqrt{1+x^2}}$$

$$= \int_0^1 dx. \frac{1}{\sqrt{1+x^2}} . \frac{\pi}{4}$$

$$= \frac{\pi}{4}\int_0^1 \frac{dx}{\sqrt{1+x^2}}$$

$$= \frac{\pi}{4}\left[\log\left(x+\sqrt{1+x^2}\right]_0^1\right.$$

$$= \frac{\pi}{4}\log\left[\sqrt{x}+1\right]$$

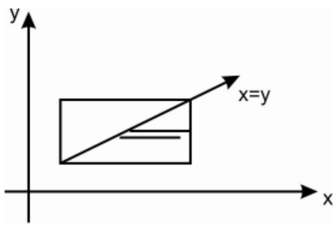
ii) $$I = \int_0^1 dy \int_y^1 \sin\left(\frac{\pi x^2}{2}\right) dx$$

Solution :

$$C = \{(x, y); y \le x < 1, 0 \le y \le 1\}$$

By Fubini's Theorem

$$I = \int_0^1 \int_y^1 \sin\left(\frac{\pi x^2}{2}\right) dxdy$$



$$= \int_0^1 \int_0^x \sin\left(\frac{\pi x^2}{2}\right) dxdy$$

$$= \int_0^1 \sin\left(\frac{\pi x^2}{2}\right) [y]_0^x \, dx$$

$$= \int_0^1 x \sin\left(\frac{\pi x^2}{2}\right) dx$$

Put $\frac{\pi x^2}{2} = t,$

x	θ	1
t	0	$\pi/2$

$$\frac{2\pi x}{2} dx = dt$$

$$xdx = \frac{dt}{\pi}$$

$$I = \int_0^{\pi/2} \sin t \frac{dt}{\pi} = \frac{1}{\pi}\int_0^{\pi/2} \sin t \, dt \, \frac{1}{\pi}(-\cos t)_0^{\pi/2}$$

$$= \frac{1}{\pi}[-0+1] = \frac{1}{\pi}$$

2.7 REVIEWS

After reading this chapter you would be knowing.

- Definition of Measure zero set and content zero set.
- Oscillation $O(f,a)$
- Find set contain measure zero on content zero
- Statement of Lebesgue Theorem
- Definition of characteristic function & its properties.
- Fubini's Theorem & its examples.

2.8 UNIT END EXERCISES

1. If $B \subseteq A$ and A has measure zero then show that & has measure zero.
2. Show that countable set has measure zero.
3. If A is non-empty open set, then show that A is not of measure zero.
4. Give an example of a bounded set C if measure zero but ∂C does not have measure zero.
5. Show by an example that a set A has measure zero but A does not have content zero.
6. Prove that $[a_1,b_1]\times....\times[a_n,b_n]$ does not have content zero if $a_i < b_i$ for each i.
7. If C is a set of content zero show that the boundary of C has content zero.
8. Give an example of a set A and a bounded subset C of A measure zero such that $\int_A \chi_c$ does not exist.
9. If f & g are integrable, then show that f_g is integrable.
10. Let $U=[0,1]$be the union of all open intervals (a_i,b_i)such that each rational number in $(0,1)$ is contained in some (a_i,b_i). Show that if $f=\chi_c$except on a set of measure zero, then f is not integrable on $[0,1]$.
11. If $f:[a,b]\times[a,b]\to\mathbb{R}$ is continuous; then show that $\int_a^b\int_x^b f(x,y)\,dx\,dy=\int_a^b\int_x^b f(x,y)\,dy\,dx$
12. Use Fubini's theorem, to compute $\int_0^{\pi/2} dy\int_0^{\pi/2}\frac{\sin x}{x+y}dx$

13. Let $A = [-1,1] \times [0, \pi/2]$ and $f : A \to \mathbb{R}$ defined by $f(x, y) = x \sin y - ye^x$ compute $\int_A f$

14. Let $f(x, y, z) = z \sin(x + y)$ and $A = [0, \pi] \times \left[-\pi/2, \pi/2\right] \times [0,1]$ computer $\int_A f$.

❖❖❖❖