Advanced Calculus of Several Variables

Devendra Kumar

Alpha Science

Advanced Calculus of Several Variables

Advanced Calculus of Several Variables

Devendra Kumar

Alpha Science International Ltd.
Oxford, U.K.

Advanced Calculus of Several Variables
222 pgs. | 33 figs.

Devendra Kumar
Department of Mathematics
Faculty of Science, Al Baha University
Al Baha, Saudi Arabia

ALPHA SCIENCE INTERNATIONAL LTD.
7200 The Quorum, Oxford Business Park North
Garsington Road, Oxford OX4 2JZ, U.K.

www.alphasci.com

Printed from the camera-ready copy provided by the Author.

ISBN 978-1-84265-916-8
E-ISBN 978-1-78332-027-1

Printed in India

Preface

This text is an outcome of the lectures delivered by the author at Addis Ababa University, Addis Ababa, Ethiopia. While dealing with this course, I realized there is a need for text book on "ADVANCED CALCULUS OF SEVERAL VARIABLES" which may give comprehensive idea of basic concepts and may also function as a companion for several variables. The primary purpose of this text is to bring this new mathematical formalism into the education system, not merely for its own sake, but as a basic framework for characterizing the full scope of modern mathematics.

The contents have been selected with the intention of meeting all requirement of graduate and undergraduate students in Science and Engineering disciplines. The author has tried all the contents of the book teaching four credit hours in one semester as a core course. Author's main aim is centralized around the theme how the reader may continue to advanced level text by self study and develop research oriented thoughts. Hence basic concepts and fundamental techniques have been emphasized while highly specialized topics and methods relegated to secondary one.

The author assumes that the reader is well acquainted with elementary calculus and algebra. Most of the chapters consist of unsolved problems just after relevant articles in the form of exercises. This book has been written with the intention of a giving reasons so that the reader may be able to work on the problems of Numerical Analysis, Operations Research, Differential Equations and Engineering applications. A bibliographic list has been incorporated which may help the reader for further studies.

I am very thankful to those writers whose direct or indirect help has been taken in this work by using their works as references.

Also, the author is thankful to Mr. S. Poothia (sachinpoothia @yahoo.com) for technical and copy editing of this book using Latex.

Devendra Kumar

Contents

Chapter 1

Euclidean $n-$Space R^n and Transformation

In this chapter we have discussed Euclidean n-space, norm, some inequalities, linear transformation and its boundedness in R^n, dual space, isometric transformation and proper rotations in R^2 and R^3.

1.1 Euclidean $n-$Space

Definition 1.1.1. The Euclidean $n-$space R^n is the set R^n of all $n-$tuples $(x^1, \ldots, x^n)$ of real numbers $x^i, i = 1, \ldots, n$ (called vectors or points) on which the two operations of vector addition and scaler multiplication are defined as follows:

(i) For any two points $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^n)$ in R^n

$$(x^1, \ldots, x^n) + (y^1, \ldots, y^n) = (x^1 + y^1, \ldots, x^n + y^n)$$

(ii) For any $a \in R$ and $(x^1, \ldots, x^n) \in R^n$

$$a.(x^1, \ldots, x^n) = (ax^1, \ldots, ax^n).$$

We can write $a.(x^1, \ldots, x^n) = a(x^1, \ldots, x^n)$. If $x \in R^n$ then $x = (x^1, \ldots, x^n)$ for some real numbers $x^1, \ldots, x^n$, and x^i is called the i^{th} component of x.

R^n is a vector space of dimension n. The vectors $e_1 = (1, 0, \ldots, 0)$, $\ldots e_n = (0, \ldots, 0, 1)$ in R^n are called the standard basis vectors of R^n. The vector $0 = (0, \ldots, 0)$ is called the zero vector in R^n.

1.2 Norm in R^n

Definition 1.2.1. The norm of a vector x in R^n denoted by $|x|$ is defined as:

$$|x| = \sqrt{(x^1)^2 + \cdots + (x^n)^2}.$$

Theorem 1.2.1. If $x \in R^n$ and $a \in R$, then

(i) $|x| \geq 0$ and $|x| = 0$ iff $x = 0$.

(ii) $|ax| = |a|.|x|$.

(iii) $|x| \leq \sum_{i=1}^n |x^i|$.

Proof. **(i)** Since $(x^1)^2 + \cdots + (x^n)^2 \geq 0$, using the definition, we get

$$\sqrt{(x^1)^2 + \cdots + (x^n)^2} \geq 0 \Rightarrow |x| \geq 0.$$

Now, if $|x| = 0$, then $(x^1)^2 + \cdots + (x^n)^2 = 0$, and hence $x^i = 0 \forall i = 1, \ldots, n$. Hence x is a zero vector in R^n.

(ii)

$$\begin{aligned} |ax| &= \sqrt{(ax^1)^2 + \cdots + (ax^n)^2} \\ &= |a|\sqrt{(x^1)^2 + \cdots + (x^n)^2} \\ &= |a||x|. \end{aligned}$$

(iii) We have

$$|x|^2 = (x^1)^2 + \cdots + (x^n)^2 = |x^1|^2 + \cdots + |x^n|^2.$$

-

But $|x^1|^2 + \cdots + |x^n|^2 \leq (|x^1| + \cdots + |x^n|)^2$, and hence

$$|x|^2 \leq (|x^1| + \cdots + |x^n|)^2.$$

Now taking square roots we get

$$|x| \leq \sum_{i=1}^{n} |x^i|.$$

Theorem 1.2.2. **(Cauchy - Schwartz Inequality)** If $x, y \in R^n$ then $|\sum x^i y^i| \leq |x||y|$, equality holds if and only if x and y are linearly dependent.

Proof. If x and y are linearly dependent then there exists $\lambda \in R$ such that $y = \lambda x$ or $\lambda y = x$. Without loss of generality assume $y = \lambda x$,

$$|x||y| = |\lambda||x|^2.$$

But

$$\left|\sum_{i=1}^{n} x^i y^i\right| = \left|\sum_{i=1}^{n} x^i \lambda x^i\right| = |x| \left|\sum_{i=1}^{n} (x^i)^2\right| = |\lambda||x|^2.$$

Hence

$$\left|\sum_{i=1}^{n} x^i y^i\right| = |x||y|.$$

Now suppose x and y are not linearly dependent. Then $\lambda y - x \neq 0$ for $\lambda \in R$. Hence $0 < |\lambda y - x|^2 = \sum_{i=1}^{n} (\lambda y^i - x^i)^2 = \lambda^2 \sum_{i=1}^{n} (y^i)^2 - 2\lambda \sum_{i=1}^{n} x^i y^i + \sum_{i=1}^{n} (x^i)^2$.

The last expression in the above equation is a quadratic in λ which never vanishes. Hence has no real root. The discriminant of the quadratic

$$4\left(\sum_{i=1}^{n} x^i y^i\right)^2 - 4\sum_{i=1}^{n} (y^i)^2 . \sum_{i=1}^{n} (x^i)^2 < 0.$$

Hence

$$\left|\sum_{i=1}^{n} x^i y^i\right|^2 < |x|^2|y|^2.$$

Taking square roots

$$\left|\sum_{i=1}^{n} x^i y^i\right| < |x||y|.$$

Theorem 1.2.3. (Minkowski Inequality) If $x, y \in R^n$ then

$|x+y| \le |x| + |y|$, equality holds iff $y = \lambda x$ where $\lambda \ge 0$ or $x = 0$.

Proof.

$$\begin{aligned}|x+y|^2 = \sum_{i=1}^{n}(x^i+y^i)^2 &= \sum_{i=1}^{n}(x^i)^2 + \sum_{i=1}^{n}(y^i)^2 + 2\sum_{i=1}^{n}x^iy^i \\ &\le |x|^2+|y|^2+2|x||y| \\ &= (|x|+|y|)^2\end{aligned}$$

$$\Rightarrow |x+y| \le |x| + |y|.$$

If $x = 0$, clearly equality holds. Hence suppose $\lambda x = y$, where $\lambda \ge 0$. Then by Cauchy-Schwartz Inequality

$$\left|\sum_{i=1}^{n} x^iy^i\right| = |x||y|.$$

But

$$\sum_{i=1}^{n} x^iy^i = \lambda \sum_{i=1}^{n}(x^i)^2 \ge 0.$$

Hence

$$\sum_{i=1}^{n} x^iy^i = \left|\sum_{i=1}^{n} x^iy^i\right|$$

and

$$\sum_{i=1}^{n} x^iy^i = |x||y|. \tag{1.2.1}$$

But

$$|x+y|^2 = \sum_{i=1}^{n}(x^i+y^i)^2 = \sum_{i=1}^{n}(x^i)^2 + \sum_{i=1}^{n}(y^i)^2 + 2\sum_{i=1}^{n}x^iy^i. \tag{1.2.2}$$

Hence in view of (1.2.1) and (1.2.2), we get

$$|x+y|^2 = |x|^2 + |y|^2 + 2|x||y| = (|x|+|y|)^2$$

or

$$|x+y| = |x| + |y|.$$

Conversely suppose that $|x+y| = |x| + |y|$ and $x \neq 0$. Then

$$|x+y|^2 = |x|^2 + |y|^2 + 2|x||y|.$$

But

$$|x+y|^2 = |x|^2 + |y|^2 + 2\sum_{i=1}^{n} x^i y^i.$$

Hence comparing the above equations, we get

$$\sum_{i=1}^{n} x^i y^i = |x||y|$$

or

$$\left|\sum_{i=1}^{n} x^i y^i\right| = |x||y|.$$

But by Cauchy-Schwartz Inequality there exists $\lambda \in R$ such that $\lambda x = y$ since $x \neq 0$. Hence

$$\lambda \sum_{i=1}^{n} (x^i)^2 = \sum_{i=1}^{n} (x^i)(\lambda x^i) = \sum_{i=1}^{n} x^i y^i = \left|\sum_{i=1}^{n} x^i y^i\right| \geq 0,$$

since $x \neq 0 \Rightarrow \lambda \geq 0$.

Corollary 1.2.1. If $x, y \in R^n$ then $|x-y| \leq |x| + |y|$, equality holds iff $y = \lambda x$ where $\lambda \leq 0$ or $x = 0$.

Proof. Let $z = -y$. Then by Minkowski inequality

$$|x+z| \leq |x| + |z|,$$

equality holds if and only if $z = \lambda x$ where $\lambda \geq 0$ or $x = 0$.

Substituting $z = -y$ and using the relation $|z| = |y|$ the above statement is equivalent to $|x-y| \leq |x| + |y|$ equality holds if and only if $y = \lambda x$ where $\lambda \leq 0$ or $x = 0$.

Corollary 1.2.2. If $x, y \in R^n$ then $|x| - |y| \leq |x-y|$.

Proof. Clearly $x = (x - y) + y$. By Minikowski inequality

$$|x| \leq |x - y| + |y|.$$

Hence

$$|x| - |y| \leq |x - y|.$$

Similarly $|y| - |x| \leq |y - x|$.

But $|x - y| = |y - x|$.

Hence

$$||x| - |y|| \leq |x - y|.$$

Distance between two points in R^n

Definition 1.2.2. Let $x, y \in R^n$. The distance between x and y denoted by $d(x, y)$ is defined as:

$$d(x, y) = |x - y|$$

d is also called a metric on R^n. The distance function d has the following properties:
For all $x, y, z \in R^n$

(i) $d(x, y) \geq 0$; equality holds if and only if $x = y$.

(ii) $d(x, y) = d(y, x)$.

(iii) $d(x, z) \leq d(x, y) + d(y, z)$.

Proof. The proof of (i) and (ii) is easy. For (iii) clearly, $x - z = (x - y) + (y - z)$. By Minkowski inequality

$$\begin{aligned} |x - z| &= |(x - y) + (y - z)| \leq |x - y| + |y - z| \\ \Rightarrow d(x, z) &\leq d(x, y) + d|y, z|. \end{aligned}$$

1.3 Inner Product in R^n

Definition 1.3.1. If $x, y \in R^n$, the inner product of x and y denoted by $< x, y >$ is defined as:

$$< x, y >= \sum_{i=1}^{n} x^i y^i.$$

Theorem 1.3.1. If x, x_1, x_2, y, y_1, y_2 are vectors in R^n and $a \in R$, then

(i) $< x, y >=< y, x >$ symmetry

(ii)

$$\left.\begin{array}{c} < ax, y >=< x, ay >= a < x, y > \\ < x_1 + x_2, y >=< x_1, y > + < x_2, y > \\ < x, y_1 + y_2 >=< x, y_1 > + < x, y_2 > \end{array}\right\} \quad \textit{bilinearity}$$

(iii) $< x, x >\geq 0$ and $< x, x >= 0$ if and only if $x = 0$ positive definiteness

(iv) $|x| = \sqrt{< x, x >}$ norm

(v) $< x, y >= \frac{|x+y|^2 - |x-y|^2}{4}$ polarization identity.

Proof.(i) $< x, y >= \sum_{i=1}^{n} x^i y^i = \sum_{i=1}^{n} y^i x^i =< y, x >$.

(ii) By (i) it suffices to prove that

$< ax, y >= a < x, y >$ and $< x_1 + x_2, y >=< x_1, y > + < x_2, y >$.

These follows from

$$< ax, y >= \sum_{i=1}^{n} (ax^i) y^i = a \sum_{i=1}^{n} x^i y^i = a < x, y >$$

and

$$\begin{aligned} < x_1 + x_2, y >= \sum_{i=1}^{n} (x_1^i + x_2^i) y^i &= \sum_{i=1}^{n} x_1^i y^i + \sum_{i=1}^{n} x_2^i y^i \\ &= < x_1, y > + < x_2, y > \end{aligned}$$

(iii)

$$< x, x >= \sum_{i=1}^{n} x^i x_i = \sum_{i=1}^{n} (x^i)^2 \geq 0;$$

but

$$\begin{aligned} < x, x >= 0 \qquad & \texttt{if and only if } \sum_{i=1}^{n} (x^i)^2 = 0, \\ & \texttt{if and only if } x^i = 0 \forall i = 1, \ldots, n. \\ & \texttt{if and only if } x = 0. \end{aligned}$$

(iv)

$$|x| = \sqrt{(x^1)^2 + \cdots + (x^n)^2} = \sqrt{< x, x >}.$$

(v)

$$\begin{aligned} & \frac{|x+y|^2 - |x-y|^2}{4} \\ = & \frac{1}{4} | < x+y, x+y > - < x-y, x-y > | \\ = & \frac{1}{4} | < x, x > +2 < x, y > + < y, y > \\ & - < x, x > +2 < x, y > - < y, y > | \\ = & < x, y > . \end{aligned}$$

Definition 1.3.2. If $x, y \in R^n$, then x and y are called $\perp$ or (orthogonal) if $< x, y >= 0$.

If vectors $x_1, \ldots, x_n$ in R^n form a basis in R^n and are pairwise perpendicular then the given basis is called an orthogonal basis of R^n. And if in addition each vector x_i has a unit norm, then the given basis is called an orthonormal basis in R^n.

i.e.;

$$< e_i, e_j >= \begin{cases} 0 & if \quad i \neq j \\ 1 & if \quad i = j \end{cases} \quad e_1, \ldots, e_n \in R^n.$$

Theorem 1.3.2. If $x_1, \ldots, x_n$ is an orthonormal basis of R^n, then for any vector x in R^n we have

(i) $x = \sum_{i=1}^n < x, x_i > x_i$.

(ii) $|x|^2 = \sum_{i=1}^n < x, x_i >^2$.

Proof. **(i)** Since $x_1, \ldots, x_n$ is a basis of R^n there exist constants $c_1, \ldots, c_n$ such that $x = \sum_{i=1}^n c_i x_i$.

Hence $< x, x_i >=< \sum_{i=1}^n c_i x_i, x_i >= \sum_{j=1}^n c_j < x_j, x_i >$.
But since $x_1, \ldots, x_n$ is an orthonormal basis it follows that

$$< x_i, x_j >= \begin{cases} 0 & if \quad i \neq j \\ 1 & if \quad i = j \end{cases},$$

$$\sum_{j=1}^n c_j < x_j, x_i >= c_i \Rightarrow < x, x_i > c_i \texttt{ for each } i = 1, \ldots, n.$$

Hence

$$x = \sum_{i=1}^n < x, x_i > x_i.$$

(ii) By (i) $x = \sum_{i=1}^n < x, x_i > x_i$.
Hence

$$\begin{aligned} |x|^2 =< x, x > &= \left\langle \sum_{i=1}^n < x, x_i > x_i, \sum_{i=1}^n < x, x_i > x_i \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n < x, x_i >< x, x_j >< x_i, x_j > . \end{aligned}$$

Again since the basis is orthonormal the above reduces to

$$\sum_{i=1}^n < x, x_i >^2 .$$

Theorem 1.3.3. If $x_1, \ldots, x_m$ are non zero vectors in R^n which are pairwise perpendicular then they are linearly independent.

Proof. Suppose $x_1, \ldots, x_m$ are not linearly independent. Then the exist real numbers $c_1, \ldots, c_m$ not all zero and such that

$$\sum_{i=1}^m c_i x_i = 0.$$

Hence

$$< \sum_{i=1}^{m} c_i x_i, \sum_{i=1}^{m} c_i x_i >= 0.$$

But

$$< \sum_{i=1}^{m} c_i x_i, \sum_{i=1}^{m} c_i x_i >= \sum_{i=1}^{m} c_i^2 |x_i|^2$$

$\Rightarrow$

$$\sum_{i=1}^{m} c_i^2 |x_i|^2 = 0 \Rightarrow c_i^2 |x_i|^2 = 0 \forall i = 1, \ldots, m.$$

But $x_i \neq 0 \Rightarrow c_i = 0 \forall i = 1, \ldots, m$, a contradiction $\Rightarrow x_1, \ldots, x_m$ are linearly independent.

The standard basis vectors in R^n form an orthonormal basis in R^n. If V is any subspace of R^n it is not evident that V has an orthonormal basis, and if it has how to find one such an orthonormal basis. For complete answer we have.

Theorem 1.3.4. Every subspace V of R^n has an orthogonal basis. (Gram-Schmidt Orthogonalziation Process).

Proof. Let $w_1 = v_1$ and hence $w_1 \neq 0$ (by induction). Define $w_2 = v_2 - \frac{<v_2, w_1>}{<w_1, w_1>} w_1$,
$w_2 \neq 0$ because v_1, v_2 are independent. Direct computation gives

$$\begin{aligned} < w_2, w_1 > &= \left\langle v_2 - \frac{< v_2, w_1 >}{< w_1, w_1 >} w_1, w_1 \right\rangle \\ &= < v_2, w_1 > - \frac{< v_2, w_1 >< w_1, w_1 >}{< w_1, w_1 >} = 0 \end{aligned}$$

$\Rightarrow w_2$ is orthogonal to w_1.

Continuing the process define $w_3, \ldots, w_m$ in R^n as follows

$$\begin{aligned} w_3 &= v_3 - \frac{< v_3, w_1 > w_1}{< w_1, w_1 >} - \frac{< v_3, w_2 >}{< w_2, w_2 >} w_2 \\ &\vdots \\ w_m &= v_m - \frac{< v_m, w_1 >}{< w_1, w_1 >} w_1 - v_m - \frac{< v_m, w_2 >}{< w_2, w_2 >} w_2 \\ &\quad - \cdots - \frac{< v_m, w_{m-1} >}{< w_{m-1}, w_{m-1} >} w_{m-1}. \end{aligned}$$

Suppose for $k : 2 \leq k \leq m-1, w_1, \ldots, w_k$ are non zero and pairwise orthogonal. Then we will show that w_{k+1} is orthogonal to all $w_i, i = 1, \ldots, k$ and $w_{k+1} \neq 0$. Hence suppose i is an integer such that $1 \leq i \leq k$. Then

$$\begin{aligned} < w_{k+1}, w_i > &= \left\langle v_{k+1} - \frac{< v_{k+1}, w_1 >}{< w_1, w_1 >} w_1 \right. \\ &\quad \left. - \cdots - \frac{< v_{k+1}, w_k >}{< w_k, w_k >} w_k, w_i \right\rangle \\ &= < v_{k+1}, w_i > - \frac{< v_{k+1}, w_1 >}{< w_i, w_i >} < w_i, w_i >= 0. \end{aligned}$$

And since w_{k+1} is a linear combination of $v_1, \ldots, v_{k+1}$ where at least the coefficient of v_{k+1} is different from zero it follows $w_{k+1} \neq 0$.

Hence $w_1, \ldots, w_m$ are pairwise orthogonal non zero vectors in V. Since $w_1, \ldots, w_m$ are independent and dimension of V is $m \Rightarrow w_1, \ldots, w_m$ is a basis of V.

1.4 Linear Transformation on R^n

Let $T : R^n \to R^m$ be a linear transformation.

$e_1, \ldots, e_n$ is the standard basis vectors in R^n, and

$e_1^{'}, \ldots, e_m^{'}$ is the standard basis vectors in R^m.

Suppose

$$T(e_i) = \sum_{j=1}^{m} a_{ji} e_j^{'}, i = 1, \ldots, n.$$

or

$$\begin{aligned} T(e_1) &= a_{11}e_1^{'} + \cdots + a_{m_1}e_m^{'} = (a_{11}, a_{21}, \ldots, a_{m_1}) \\ &\vdots \\ T(e_n) &= a_{1n}e_1^{'} + \cdots + a_{mn}e_m^{'} = (a_{1n}, \ldots, a_{mn}). \end{aligned}$$

Then the matrix $A = (a_{ij})$ is the matrix representation of T with respect to standard basis vectors in R^n and R^m.

Now if $y = T(x)$ for $x \in R^n$ and $y \in R^m$, then the matrix equation is

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

To see this

$$\begin{aligned} T(x) &= T(x^1e_1 + \cdots + x^ne_n) = x^1T(e_1) + \cdots + x^nT(e_n) \\ &= x^1(a_{11}, a_{21}, \ldots a_{m1}) + \cdots + x^n(a_{1n}, \ldots, a_{mn}) \\ &= \left(\sum_{i=1}^{n} a_{1i}x^i, \ldots, \sum_{i=1}^{n} a_{mi}x^i\right). \end{aligned}$$

Hence

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} & a_{1i}x^i \\ \vdots & \\ \sum_{i=1}^{n} & a_{mi}x^i \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}.$$

Boundedness of Linear Transformation

Theorem 1.4.1. If $T : R^n \to R^m$ is a linear transformation then there exists $M \geq 0$ such that $|T(h)| \leq M|h| \forall h \in R^n$.

Proof. Let the matrix representation T with respect to the standard basis vectors in R^n and R^m be

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}.$$

Then for any $h = (h^1, \ldots, h^n)$ in R^n

$$T(h) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h^n \end{pmatrix}.$$

Let $A_i = (a_{i1}, \ldots, a_{in}), i = 1, \ldots, m$,or

$$T(h) = \left(\sum_{i=1}^{n} a_{1i} h^i, \ldots, \sum_{i=1}^{n} a_{mi} h^i \right)$$

or

$$|T(h)| \leq \sum_{j=1}^{m} \left| \sum_{i=1}^{n} a_{ji} h^i \right| \leq \sum_{j=1}^{m} \sum_{i=1}^{n} |a_{ji} h^i|.$$

Using Schwartz inequality

$$\begin{aligned} \leq \sum_{j=1}^{m} |A_j||h| &= |h| \sum_{j=1}^{m} |A_j| \\ &= |h| M (M = \sum_{j=1}^{m} |A_j|) \end{aligned}$$

or

$$|T(h)| \leq M|h| \forall h \in R^n.$$

Corollary 1.4.1. If $T : R^n \to R^m$, where $n \leq m$ is a one to one linear transformation then there exists $m > 0$ such that $|T(x)| \geq m|x| \forall x \in R^n$.

Proof. Let the range of T in R^m be denoted by V. T^{-1} restricted to V is a linear transformation.

Extend: T^{-1} to R^m by defining $T^{-1}(x) = 0 \forall x \in R^m\ V$, and denote the extension also by T^{-1}. Using Theorem 1.4.1 there exists $M > 0$ such that $|T^{-1}(y)| \leq M|y| \forall y \in R^m$. But for all $y \in V$, there exists $x \in R^n$ such that $y = T(x)$, and conversely.

Hence

$$|T^{-1}(T(x))| \leq M|T(x)| \forall x \in R^n$$

or

$$\frac{1}{M}|T^{-1}(T(x))| \leq T(x) \forall x \in R^n.$$

Now if we set $m = \frac{1}{M}$, we have

$$m|x| \leq |T(x)| \forall x \in R^n.$$

1.5 Dual Space of R^n

Definition 1.5.1. The dual space of R^n is the set of all linear transformation from R^n into R^n denoted by $(R^n)^*$ with vector addition and scaler multiplication defined on $(R^n)^*$ as follows.

(i) If T, S are in $(R^n)^*$ then $(T+S)(x) = T(x) + S(x) \forall x \in R^n$.

(ii) If $T \in (R^n)^*$ and $a \in R$ then $(aT)(x) = a(Tx) \forall x \in R^n$,

$(R^n)^*$ is a vector space of dimension n.

Theorem 1.5.1. (Riesz Theorem). For each $x \in R^n$, let $\varphi_x \in (R^n)^*$ and given by $\varphi_x(y) =< x, y > \forall y \in R^n$. Define $T : R^n \to (R^n)^*$ by $T(x) = \varphi_x$ for each $x \in R^n$. Then

(i) T is a one to one linear transformation.

(ii) For each $\varphi \in (R^n)^*$ there is a unique $x \in R^n$ such that $\varphi = \varphi_x$.

Proof.(i). Let x and y be in R^n. Then $\forall z \in R^n$,

$$\varphi_{x+y}(z) =< x+y, z >=< x, z > + < y, z >= \varphi_x(z) + \varphi_y(z)$$

$\Rightarrow$

$$\varphi_{x+y} = \varphi_x + \varphi_y.$$

Hence $\forall x, y \in R^n$

$$T(x+y) = \varphi_{x+y} = \varphi_x + \varphi_y = T(x) + T(y).$$

Let $x \in R^n$ and $\lambda \in R$. Then $\forall z \in R^n$

$$\varphi_{\lambda x}(z) =< \lambda x, z >= \lambda < x, z >= \lambda\varphi_x(z)$$

$\Rightarrow$

$$\varphi_{\lambda x} = \lambda\varphi_x$$

or for $x \in R^n$ and $\lambda \in R$,

$T(\lambda x) = \varphi_{\lambda x} = \lambda\varphi_x$. Hence T is a linear transformation.

To show that T is one to one.

Let $x, y \in R^n$ such that $T(x - y) = 0$. Since T is linear it follows that $T(x) = T(y)$. But $T(x - y) = \varphi_{x-y}$. Hence for all $z \in R^n, \varphi_{x-y}(z) = 0 \Rightarrow< x - y, z >= 0 \forall z \in R^n$. If we take $z = x - y$, then $< x - y, x - y >= 0$. Hence $x - y = 0$, which is equivalent to $x = y \Rightarrow T$ is one to one.

(ii). Now since $T : R^n \to (R^n)^*$ is a one to one linear transformation where both the domain and range are spaces of dimension $n \Rightarrow T$ is onto $\Rightarrow T$ is isomorphism. Hence for each $\varphi \in (R^n)^*$ there is a unique x in R^n such that $\varphi = T(x) = \varphi_x$.

Definition 1.5.2. A linear transformation $T : R^n \to R^n$ is norm preserving if $|T(x)| = |x| \forall n \in R^n$, and inner product preserving if $< Tx, Ty >=< x, y > \forall x, y \in R^n$.

Theorem 1.5.2. Let $T : R^n \to R^n$ be a linear transformation. T is norm preserving if and only if T is a inner product preserving.

Proof. Suppose T is norm preserving. Let x and y be any two points in R^n. Then

$$|T(x + y)| = |x + y| \text{ and } |T(x - y)| = |x - y|.$$

By polarization identity

$$< x, y >= \frac{|x + y|^2 - |x - y|^2}{4}.$$

Hence by substitution, and using polarization identity,

$$\begin{aligned} < x, y > &= \frac{|T(x+y)|^2 - |T(x-y)|^2}{4} \\ &= \frac{|T(x)+T(y)|^2 - |T(x)-T(y)|^2}{4} =< Tx, Ty > . \end{aligned}$$

Conversely, suppose T is inner product preserving.

Let $x \in R^n$. Then

$$|T(x)|^2 =< Tx, Tx >=< x, x >= |x|^2$$

$\Rightarrow$

$$|T(x)| = |x|.$$

Theorem 1.5.3. Let $T : R^n \to R^n$ be a linear transformation. If T is norm preserving then T is one to one, and further more T^{-1} is a norm preserving linear transformation.

Proof. T is one to one. To see this let x and y be in R^n such that $x \neq y$. Then $|Tx - Ty| = |T(x - y)| = |x - y| \neq 0$.

Hence $Tx \neq Ty$. Since now T is a linear transformation from R^n into R^n i.e., one to one it must be onto. Hence T^{-1} exists. But then we know T^{-1} is also a linear transformation from R^n onto R^n. Now let $x \in R^n$ and $y \in R^n$ such that $Ty = x$. Then

$$|T^{-1}x| = |T^{-1}(Ty)| = |y|. \texttt{ But } |Ty| = |y|.$$

Hence

$|T^{-1}x| = |Ty| = |x|,$ `which shows` T^{-1} `is norm preserving.`

Definition 1.5.3. If $x, y \in R^n$ are nonzero vectors, the angle between x and y, denoted by $\angle(x, y) = arc\cos\frac{<x,y>}{|x||y|}$. A linear transformation $T : R^n \to R^n$ is angle preserving if T is one to one and $\forall x, y \in R^n$ and nonzero vectors we have

$$\angle(Tx, Ty) = \angle(x, y).$$

Clearly, if T is norm preserving then it is angle preserving. As we can see easily

$$\begin{aligned}
&arc\cos\frac{< x,y >}{|x||y|} = arc\cos\frac{< Tx,Ty >}{|Tx||Ty|} \\
&\Rightarrow \angle(Tx,Ty) = \angle(x,y).
\end{aligned}$$

The converse of the above statement is obviously false.

Lemma 1.5.1. If $T : R^n \to R^n$ is a linear transformation, $\alpha > 0$ and $|Tx| = \alpha|x| \forall x \in R^n$, then $< Tx, Ty >= \alpha^2 < x, y > \forall x, y \in R^n$.

Proof. Let $x, y \in R^n$. Then by polarization identity

$$\begin{aligned}
< Tx,Ty > &= \frac{|Tx+Ty|^2 - |Tx-Ty|^2}{4} \\
&= \frac{|T(x+y)|^2 - |T(x-y)|^2}{4} \\
&= \frac{\alpha^2|x+y|^2 - \alpha^2|x-y|^2}{4} \\
&= \alpha^2 < x,y > .
\end{aligned}$$

Lemma 1.5.2. Let $T : R^n \to R^n$ be angle preserving linear transformation. If $x, y \in R^n$ such that $|x| = |y|$ then $|Tx| = |Ty|$.

Proof. Suppose $|x| = |y|$. Then

$$|x| = |y| \Leftrightarrow < x,x >=< y,y > \Leftrightarrow < x,x > - < y,y >= 0.$$

But $< x,x > - < y,y >=< x+y, x-y > \Rightarrow < x+y, x-y >= 0$.

Since T is angle preserving $\Rightarrow < T(x+y), T(x-y) >= 0$.

But $< T(x+y), T(x-y) >=< Tx+Ty, Tx-Ty >=< Tx, Tx > - < Ty, Ty >$.

Hence $< Tx, Tx > - < Ty, Ty >= 0 \Rightarrow |Tx| = |Ty|$.

Theorem 1.5.4. Let T be a linear transformation from R^n into R^n. T is angle preserving iff $|Tx| = \alpha|x| \forall x \in R^n$ and some $\alpha > 0$.

Proof. Suppose T is angle preserving. Let $\{e_1, \ldots, e_n\}$ be the usual basis in R^n. Then $\{Te_1, \ldots, Te_n\}$ is an orthogonal basis of R^n. This can be seen easily as follows. Since $\{e_1, \ldots, e_n\}$ is a set of mutually orthogonal vectors, as T is angle preserving $\{Te_1, \ldots, Te_n\}$ is a set of mutually orthogonal vectors and by Theorem 1.5.3., $Te_1, \ldots, Te_n$ are linear independent.

Now using Lemma 1.5.1. $\exists \alpha > 0$ such that $|Te_i| = \alpha \forall i = 1, \ldots, n$. Let $x \in R^n$.

$$\begin{aligned}
|Tx|^2 = < Tx, Tx > &= \left\langle T\left(\sum_{i=1}^{n} x^i e_i\right), T\left(\sum_{i=1}^{n} x^i e_i\right)\right\rangle \\
&= \left\langle \sum_{i=1}^{n} x^i (Te_i), \sum_{i=1}^{n} x^i (Te_i)\right\rangle \\
&= \sum_{i=1}^{n} (x^i)^2 < Te_i, Te_i > \\
&= \alpha^2 |x|^2
\end{aligned}$$

$\Rightarrow |Tx| = \alpha|x|$.

Conversely suppose that $|Tx| = \alpha|x| \forall x \in R^n$ and some $\alpha > 0$. Now using Lemma 1.5.1., we can write

$$\angle(x, y) = arc\cos\frac{< x, y >}{|x||y|} = arc\cos\frac{\frac{< Tx, Ty >}{\alpha^2}}{\frac{|Tx|}{\alpha}\frac{|Ty|}{\alpha}} = arc\cos\frac{< Tx, Ty >}{|Tx||Ty|}$$

or

$$\angle(x, y) = \angle(Tx, Ty).$$

Furthermore, it is clear that $|Tx| = 0$ if and only if $x = 0$. Hence T is one to one. Therefore T is angle preserving.

1.6 Isometric Transformation

Definition 1.6.1. A function $T : R^n \to R^n$ is called isometric if $\forall x, y \in R^n, d(Tx, Ty) = d(x, y)$.

Obviously a norm preserving (or inner product preserving) linear transformation is isometric because,

$$d(Tx, Ty) = |Tx - Ty| = |T(x - y)| = |x - y| = d(x, y) \forall x, y \in R^n.$$

Conversely if $T : R^n \to R^n$ is an isometric function such that it keeps the origin fixed then T is norm preserving and inner product preserving. This can be proved as follows.

The preservation of norm follows immediately from $|Tx| = |Tx - T(0)| = d(Tx, T0) = d(x, 0) = |x - 0| = |x| \forall x \in R^n$. Now for all x and y in R^n,

$$|Tx - Ty| = d(Tx, Ty) = d(x, y) = |x - y|.$$

But

$$|Tx - Ty|^2 = |Tx|^2 + |Ty|^2 + 2 < Tx, Ty > .$$

Hence

$$|Tx|^2 + |Ty|^2 + 2 < Tx, Ty >= |x|^2 + |y|^2 + 2 < x, y > .$$

Since we have proved already that T preserves norm, it follows that $< Tx, Ty >=< x, y > \forall x, y \in R^n$.

The most interesting and surprising property of isometric functions from R^n into R^n is the one given by the following.

Theorem 1.6.1. If $T : R^n \to R^n$ is an isometric function which keeps the origin fixed, then T is a linear transformation.

Proof. Let x and y be two vectors in R^n.

We claim $T\left(\frac{x+y}{2}\right) = \frac{Tx+Ty}{2}$.

By triangle inequality

$$d(Tx, Ty) \leq d\left(Tx, T\frac{(x+y)}{2}\right) + d\left(\frac{T(x+y)}{2}, Ty\right).$$

Since T is isometric, we have

$$d(Tx, Ty) = |Tx - Ty| = |x - y|.$$

$$d\left(Tx, \frac{T(x+y)}{2}\right) = \left|Tx - T\left(\frac{x+y}{2}\right)\right| = \left|x - \frac{(x+y)}{2}\right| = \frac{1}{2}|x-y|,$$

and

$$d\left(\frac{T(x+y)}{2}, Ty\right) = \left|T\left(\frac{x+y}{2}\right) - Ty\right| = \left|\frac{x+y}{2} - y\right| = \frac{1}{2}|x-y|.$$

Hence

$$d(Tx,Ty) = d\left(Tx, T\left(\frac{x+y}{2}\right)\right) + d\left(T\left(\frac{x+y}{2}\right), y\right).$$

Hence $T\left(\frac{x+y}{2}\right)$ is the mid point of $Tx + Ty$.

Thus

$$T\left(\frac{x+y}{2}\right) = \frac{Tx+Ty}{2}.$$

$\Rightarrow T\left(\frac{x}{2}\right) = \frac{T(x)}{2}$ by taking y to be zero vector and $T(0) = 0$.

$\Rightarrow \frac{T(x+y)}{2} = T\left(\frac{x+y}{2}\right) = \frac{Tx+Ty}{2}$.

$\Rightarrow T(x+y) = Tx + Ty$.

To show that for any $\alpha \in R$ and $x \in R^n, T(\alpha x) = \alpha T(x)$. We proceed the case by case

(i) $\alpha = 0$, then $T(\alpha x) = 0 = \alpha T(x)$.

(ii) $\alpha = n \in N$, then using the additive property established above and induction $T(nx) = nT(x)$.

(iii) $\alpha = \frac{1}{n}, n \in N$. Let $y = \frac{x}{n}$ then $ny = x$.
By (ii) $T(x) = T(ny) = nT(y) = nT\left(\frac{x}{n}\right)$.

(iv) $\alpha = \frac{m}{n}, m, n \in N$. Using cases (iii) and (ii)

$$T\left(\frac{m}{n}x\right) = mT\left(\frac{x}{n}\right) = \frac{m}{n}T(x).$$

(v) $\alpha > 0$ and $\alpha \in R$. Let $\{r_i\}$ be a sequence of positive rational numbers converging to α. Then

$$\begin{aligned}|\alpha T(x) - T(\alpha x)| &\leq |\alpha T(x) - r_i T(x)| + |r_i T(x) - T(\alpha x)| \\ &= |\alpha T(x) - r_i T(x)| + |T(r_i x) - T(\alpha x)| \\ &= 2|\alpha - r_i||x|.\end{aligned}$$

Since the above is true for all r_i, it follows $\alpha T(x) = T(\alpha x)$.
Case (vi) $\alpha < 0, \alpha \in R$. First of all observe that by what we have proved already.

$$\frac{T(x) + T(-x)}{2} = \frac{T(x - x)}{2} = T(0) = 0.$$

$\Rightarrow$

$$-T(x) = T(-x)$$

$\Rightarrow$

$$\begin{aligned}T(\alpha x) &= T(-\alpha. - \alpha) \\ &= -\alpha T(-x) \\ &= -\alpha. - T(x) \\ &= \alpha T(x).\end{aligned}$$

Definition 1.6.2. An isometric linear transformation from R^n onto R^n is called a rotation. If $A = (a_{ij})$ is the matrix of T then $det A = \pm 1$. The rotation T is called proper if $det A = 1$, and improper if $det A = -1$.

1.7 Proper Rotation in R^n

Let T be a proper rotation of R^n, and e_1, e_2 the standard basis of R^n. Suppose $Te_1 = ae_1 + be_2$. By applying the pythagorean theorem and the isometry of T we have $a^2 + b^2 = 1$. Since $Te_2 \perp Te_1 \Rightarrow Te_2 = \pm(-be_1 + ae_2)$.

Now T is a rotation implies

$$det \;\; T = \pm \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \pm(a^2 + b^2).$$

Since the rotation T is proper $detT = a^2 + b^2$. Hence $Te_2 = -be_1 + ae_2$.

Now define an angle $\theta, 0 \leq \theta < \pi$ such that $a = \cos\theta$. Hence $b = \pm\sin\theta$.

Thus T is given by

$$\begin{aligned} Te_1 &= \cos\theta e_1 + \sin\theta e_2 \\ Te_2 &= -\sin\theta e_1 + \cos\theta e_2 \end{aligned}$$

or

$$\begin{aligned} Te_1 &= \cos\theta e_1 + (-\sin\theta) e_2 \\ Te_2 &= \sin\theta e_1 + \cos\theta e_2 \end{aligned}$$

$\Rightarrow$

$$T(x^1, x^2) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$$

or

$$T(x^1, x^2) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}.$$

The geometric interpretation of the above two transformation is given below.

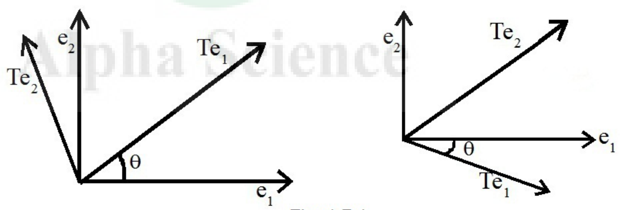

Fig. 1.7.1

θ is the angle of rotation of the plane.

1.8 Rotation in R^3

If $T : R^3 \to R^3$ is a rotation then there always exists a vector x in R^3 and a real number $|\lambda| = 1$ such that $Tx = \lambda x$.

In other words a rotation in R^3 always has at least one eigenvalue. Moreover if T is not the identity transformation, then there is exactly one eigenvalue. If T is proper rotation then the eigenvalue of T is 1. And in this case the line generated by the eigenvectors of T is called axis of rotation.

The rotation angle of T is defined as the rotation of the plane perpendicular to the axis of rotation.

Exercises 1

1. Prove that if $A : R^n \to R^m$ is a linear transformation, then $|A| < \infty$ and A is uniformly continuous mapping from R^n to R^m.

2. Prove that if $A, B : R^n \to R^m$ and $\alpha \in R$, then

$$|A + B| \leq |A| + |B|, |\alpha A| = |\alpha||A|.$$

3. If $|A - B| = d(A, B)$ then linear transformation becomes a normal space.

4. Prove that if $A : R^n \to R^m$,$B : R^m \to R^p$ then $B.A : R^n \to R^p$ and

$$|B.A| \leq |B| + |A|.$$

5. Let u and ν be vectors in R^n. Prove that if $u = 0$ or $\nu = \alpha u$ for some number α, then the Cauchy-Schwartz Inequality becomes an equality.

   Then prove the converse: If $| < u, \nu > | = \|u\|\|\nu\|$, then either $u = 0$ or $\nu = \alpha u$ for some number α.

6. Let u be a point in R^n, and suppose that $\|u\| < 1$. Show that if ν is in R^n and $\|v - u\| < 1 - \|u\|$, then $\|\nu\| < 1$.

7. Let u be a point in R^n and let r be a positive number. Suppose that the points in R^n are at a distance less than r from the point u. Prove that if $0 \leq t \leq 1$, then the point $t\nu + (1 - t)u)$ is also at a distance less than r from u.

8. Consider the two points $u = (1, 3, -2)$ and $\nu = (2, 2, 4)$ in R^3. Find the norm of u and the norm of ν, and show that u and ν are perpendicular. Show that

$$\|u + \nu\|2 = \|u\|^2 + \|\nu\|^2.$$

9. Find a linear transformation $T : R^3 \to R^3$ that has the property that $T(1, 1, 1) = (0, 2, 0), T(1, 1, -1) = (1, 2, 0)$ and $T(2, 0, 0) = (1, 1, 1)$.

10. Find the 2×2 matrix that is associated with the mapping in the plane that rotates points 90^o counterclockwise about the origin.

11. For a point (x, y) in the plane R^2, define $T(x, y)$ to be the point on the line $l = \{(x, y) \text{ in } R^2 : y = 2x\}$ that is closest to (x, y). Show that the mapping $T : R^2 \to R^2$ is linear and find the 2×2 matrix that is associated with this mapping.

12. Suppose that the mapping $T : R^n \to R^n$ has the property that there is another mapping $S : R^n \to R^n$ such that

$$T(S(x)) = S(T(x)) = x \texttt{ for all } x \texttt{ in } R^n.$$

Prove that $T : R^n \to R^n$ is invertible and that its inverse is the mapping $S : R^n \to R^n$.

Chapter 2

Topology on the Euclidean $n-$space R^n

In this chapter we have discussed open and closed sets in R^n, product of sets, compactness, dense and nowhere dense sets, sequence in R^n and Lebesgue covering theorem.

2.1 Open and Closed Sets

Definition 2.1.1. Let $a, b \in R^n$ such that $a^i < b^i, i = 1, \ldots, n$. The closed rectangle in R^n denoted by $[a, b]$ is defined as the cartesian product $[a^1, b^1] \times \cdots \times [a^n, b^n]$ of the closed intervals $[a^1, b^1], \ldots, [a^n, b^n]$ in R^1. Similarly the open rectangle in R^n denoted by (a, b) is defined as the cartesian product

$$(a^1, b^1) \times \cdots \times (a^n, b^n),$$

of the open intervals $(a^1, b^1), \ldots, (a^n, b^n)$ in R^1.

Definition 2.1.2. (Open and closed sets in R^n). A set U in R^n is said to be an open set in R^n if for each $x \in R^n$ there is an open rectangle in R^n contained x and contained in U. A set V in R^n is closed in R^n if its complement is an open set in R^n.

Proposition 2.1.1. An open rectangle is an open set.

Proof. Suppose (a, b) is an open rectangle in R^n. Let $x \in (a, b)$. Then the open rectangle $\left(\frac{x+a}{2}, \frac{b+x}{2}\right)$ contains x and is contained by (a, b). Hence (a, b) is an open set.

Proposition 2.1.2. A closed rectangle is a closed set.

Proof. Consider a point $x \in [a, b]^c$, and set $r = \min\{|x^i - b^i|, |x^i -$

$a^i|, i = 1, \ldots, n\}$. Denote by α the point $\left(\frac{r}{2}, \ldots, \frac{r}{2}\right)$ in R^n. Then the open rectangle $(x - \alpha, x + \alpha) \subset [a, b]^c$. Hence $[a, b]^c$ is open, and therefore $[a, b]$ is closed.

Remark 2.1.1. The two sets R^n and ϕ are both open and closed sets in R^n. There is no other subsets of R^n that are both open and closed.

Theorem 2.1.1.

(i) The union of any collection of open sets in R^n is open.

(ii) The intersection of a finite collection of open sets in R^n is open.

Proof.

(i) Let $\{0_\alpha : \alpha \in A\}$ be a family of open sets indexed by the set A. Suppose $x \in \cup_{\alpha \in A} 0_\alpha$, then $x \in 0_{\alpha_0}$ for some $\alpha_0 \in A$. Since 0_{α_0} is open there is an open rectangle $w \subset R^n$ such that $x \in w \subset 0_{\alpha_0}$ But then $w \subset \cup_{\alpha \in A} 0_\alpha$. Hence $\cup_{\alpha \in A} 0_\alpha$ is open.

(ii) Let $\{0_\alpha : \alpha \in A\}$ be a family of open sets indexed by the set A containing a finite number of elements. Suppose $x \in \cap_{\alpha \in A} 0_\alpha$; then $x \in 0_\alpha$ for every $\alpha \in A$. Since for each $\alpha \in A, 0_\alpha$ is open, there is an open rectangle $w_\alpha = (a_\alpha, b_\alpha) \subset R^n \ni x \in w_\alpha \subset 0_\alpha$ for each $\alpha \in A$. Set $r_i = \min\{|x^i - a^i_\alpha|, |x^i - b^i_\alpha|; \ \ i = 1, \ldots, n$ and $\alpha \in A\}$ $\beta = \left(\frac{r_1}{\alpha}, \ldots, \frac{r_n}{\alpha}\right)$. Then the open rectangle $\{x - \beta, x + \beta\} \subset (a_\alpha, b_\alpha) \subset 0_\alpha$ for each $\alpha \in A$. Hence $\{x - \beta, x + \beta\} \subset \cap_{\alpha \in A} 0_\alpha \Rightarrow \cap_{\alpha \in A} 0_\alpha$ is an open set.

Now using, De Morgan's laws, which states that for any family $\{\cup_\alpha, \alpha \in A\}$ of sets

$$(\cup_{\alpha \in A} u_\alpha)^c = \cap_{\alpha \in A} u^c_\alpha \text{ and } (\cap_{\alpha \in A} u_\alpha)^c = \cup_{\alpha \in A} u^c_\alpha.$$

Corollary 2.1.1.

(i) The intersection of any collection of closed sets in R^n is closed.

(ii) The union of finite collection of closed sets is closed.

Let J be the collection of all subsets of R^n that are open. Then J satisfies the following conditions.

(i) $\phi \in J$ and $R^n \in J$.

(ii) The union of any number of sets in J is in J.

(iii) The intersection of a finite number of sets in J is in J.

R^n together with J whose elements satisfy (i),(ii),(iii) form a topology space (R^n, J).

Note: The above topology J is called usual topology of R^n.

Alternative definition of open sets in R^n

Definition 2.1.3. An open ball in R^n with centre a in R^n and radius $r > 0$ is denoted by $B_r(a) = \{x \in R^n : |x - a| < r\}$.

Similarly closed ball $\overline{B}_r(a) = \{x \in R^n : |x - a| \leq r\}$.

A set U in R^n is open if for every $x \in U \exists$ an open ball B containing x such that $x \in B \subset U$.

The open sets in R^n generated by using the above definitions are identical with the open sets in R^n which were generated by previous definition. Which can be seen by the following theorem:

Theorem 2.1.2.

(i) If w is any open rectangle in R^n then for each x in w there is an open ball B containing x and contained in w i.e., $x \in B \subset w$.

(ii) If B is any open ball in R^n then for each x in B there is an open rectangle w containing x and contained in B i.e., $x \in w \subset B$.

Proof.

(i) Suppose w is any open rectangle in R^n, and x be any point in w. Since w is an open rectangle in R^n there are two points a, b in R^n such that $w = (a, b)$. Hence let

$$r = \min\{|x^i - a^i|, |x^i - b^i|, i = 1, \ldots, n\}.$$

Then the open ball $B_r(x)$ contains x and is contained in w.

(ii) Suppose B is an open ball in R^n. Then there is $a \in R^n$ and $r > 0$ such that B is an open ball with center a and radius r i.e., $B = B_r(a)$. Let $x \in B_r(a), t = |x - a|$, and $s = r - t$. Then $B_s(x) \subset B_r(a)$, or which is the same thing

$$\{y \in R^n : |y - x| < s\} \subset \{y \in R^n : |y - a| < r\}.$$

To see this suppose

$$z \in \{y \in R^n : |y - x| < s\}.$$

By triangle inequality

$$|z - a| \leq |z - x| + x - a| < s + t.$$

Hence

$$z \in \{y \in R^n : |y - a| < r\}.$$

Now define the open rectangle w in R^n by

$$w = \left(x^1 - \frac{s}{2}, x^1 + \frac{s}{2}\right) \times \cdots \times \left(x^n - \frac{s}{2}, x^n + \frac{s}{2}\right).$$

Clearly, $w \subset B_s(x) \subset B_r(a)$ and $x \in w$.

Each of the two families of sets, namely the collection of open rectangles and the collection of open balls, which generates the open sets is called a basis (for the usual topology), and each elements of a basis is called a basic open set.

Theorem 2.1.3. A subset A of R^n is open iff A is the union of a countable collection of open rectangles (balls).

Proof. Suppose A is a countable union of open rectangles $\Rightarrow A$ is open.

Conversely assume A is an open set. Let Q denote the set of all rational numbers. Then clearly $Q^n \cap A$ is a countable set. Hence suppose $\{r_k, k \in N\}$ be an enumeration of $Q^n \cap A$. For each

$k \in N$, let m_k denote the smallest positive integer such that the open rectangle $w \subset R^n$ defined by

$$\begin{aligned} w = w_{1/mk}(r_k) \;\; &= \;\; \left(r_k^1 - \frac{1}{m_k}, r_k^1 + \frac{1}{m_k}\right) \times \cdots \times \\ &\qquad \left(r_k^n - \frac{1}{m_k}, r_k^n + \frac{1}{m_k}\right) \subset A. \end{aligned}$$

$\Rightarrow \cup_{k \in N} w_{1/m_k}(r_k) \subset A$.

Now take any $x \in A, \exists m \in N \ni w_{2/m}(x) \subset A$. Clearly there is a point $y \in R^n$ with rational components $\ni y \in w_{1/m}(x)$. i.e., $y = r_k$ for some $k \in N$. Hence $x \in w_{1/m}(r_k) \subset w_{2/m}(x) \subset A$. But $w_{1/m}(r_k) \subset w_{1/m_k}(r_k)$. Consequently $x \in w_{1/m_k}(r_k)$. Hence $A \subset \cup_{k \in N} w_{1/m_k}(r_k)$.

Corollary 2.1.2. A subset A of R^n is closed iff A is the intersection of a countable collection of closed sets.

Exercises 2.1

1. Determine which of the following subsets A of R^2 are open in R^2, closed in R^2, or neither open nor closed in R^2.

   (i) $A = \{\nu = (x, y) : x^2 > y\}$.

   (ii) $A = \{\nu = (x, y) : x \geq 0, y \geq 0\}$.

   (iii) $A = \{\nu = (x, y) : x \text{ is rational }\}$.

   (iv) $A = \{\nu = (x, y) : x^2 + y^2 = 1\}$.

2. Let r be a positive number, and define $0 = \{\nu \in R^n : |\nu| > r\}$. Prove that 0 is open in R^n by showing that its complement is closed in R^n.

3. Let A be a subset of R^n and w be a point in R^n. The translate of A by w is denoted by $w + A$ and is defined by

$$w + A \equiv \{w + \nu : \nu \in A\}.$$

   (i) Show that A is open if and only if $w + A$ is open.

   (ii) Show that A is closed if and only if $w + A$ is closed.

4. Let $y \subset X$. Give an example where A is open in y but not open in X. Give an example where A is closed in y but not closed in X.

5. Let $A \subset X$. Show that if C is a closed set of X and C contains A, then C contains $\overline{A}$.

2.2 Interior, Exterior and Boundary Points of a Set

Definition 2.2.1. Let A be a subset of R^n. A point $x \in R^n$ is called

(i) an interior point of A if there exists an open rectangle containing x and contained in A, i.e., $x \in B \subset A$.

(ii) an exterior point of A if there exists an open rectangle containing x and contained in A^c, i.e., $x \in B \subset A^c$ or $x \in B \subset R^n - A$.

(iii) a boundary point of A if every open rectangle containing. x has a nonempty intersection ball with A and A^c, i.e., $x \in B, B \cap A \neq \phi, B \cap A^c \neq \phi$.

The above three sets $intA, extA$ and $boundA$- are mutually disjoint and their union is the set R^n.

Theorem 2.2.1. For any subset A of R^n, $intA$ and $extA$ are open and $boundA$ is closed.

Proof. We will prove the result by contradiction. Suppose $intA$ is not open. Then $\exists x \in intA \ni$ for every open rectangle $w \subset R^n$ containing x, we have $w \cap (intA)^c \neq \phi$. Hence $w \cap (extA) \neq \phi$ or $w\cap$ bound $A) \neq \phi$ for every open rectangle $w \subset R^n$ containing x.

If for $w \subset R^n, w \cap (extA) \neq \phi \Rightarrow w \cap A^c \neq \phi$. Hence suppose $w \cap$ bounded $A \neq \phi$. Then $\exists y \in w \wedge$ bound A. Since y is a boundary point, and w is an open rectangle containing $y, \Rightarrow w \cap A^c \neq \phi$.

Hence in any case $w \cap A^c \neq \phi \forall$ open rectangle $w \subset R^n$ which contain x. Hence x is not an interior point of A contradiction.

Similarly, we can prove that $extA$ is open. Now we have (bound $A)^c = (intA \cup extA) \Rightarrow$ bound A is closed.

Cluster Point. A point $x \in R^n$ is called a cluster point of $A \subset R^n$ if every open rectangle containing x has a nonempty intersection with $A \setminus \{x\}$. If $x \in A$ but it is not a cluster point of A then x is called an isolated point of A, i.e., $x \in w \cap A \setminus \{x\} \neq \phi$, $x \in A$, $w \cap A = \{x\}$.

A closed set every point of which is a cluster point of the set is called a perfect set. A closed rectangle in R^n and a closed ball in R^n are both perfect sets. A subset of R^n whose complement contains no isolated point is called a dense set. The set Q of rational number is dense set in R, and the set $Q \times \cdots \times Q$, n product of Q is dense set in R^n.

Theorem 2.2.2. A subset of R^n is closed iff the subset contains all its cluster points.

Proof. Suppose $A \subset R^n$ is closed, and x is a cluster point of A. If $x \in A$ then $x \in A^c$. Since A^c is open $\exists$ an open rectangle $w \ni x \in w$ and $w \subset A^c$. Hence x is not a cluster point of A. Contradiction $\Rightarrow x \in A$.
Conversely suppose $A \subset R^n$ contains all its cluster points and $x \in A^c$. Then $\exists$ an open rectangle w such that $x \in w$ and $w \subset A^c$. Hence A^c is open, $\rightarrow A$ is closed.
The set of all cluster points of the set $A \subset R^n$ is denoted by A', and called the derived set of A. The set $A \cup A'$ is called the closure of the set A and is denoted by $\overline{A}$.

$$A = (a,b) \text{ or } [a,b) \text{ or } (a,b], [a,b]$$
$$\overline{A} = [a,b], A^0 = (a,b), \quad b(A) = [a,b]$$
$$A = \text{ rational points in } (0,1).$$
$$int(\overline{A}) = int[0,1] = \phi, \quad A \text{ is nowhere dense}.$$

Exercises 2.2

1. Let r be a positive number, and define $O = \{\nu \in R^n : |\nu| > r\}$. Prove that O is open in R^n by showing that every point in O is an interior point of O.

2. For a subset A of R^n, the closure of A, denoted by $cl\ A = intA \cup bdA$.Prove that $A \subseteq clA$, and that $A = cl\ A$ if and only if A is closed in R^n.

3. Let A and B be subsets of R^n with $A \subseteq B$.

   (i) Prove that $int\ A \subseteq\ int\ B$.

   (ii) Is it necessarily true that $b(A) \subseteq b(B)$?

4. (i) Show that if Q is a rectangle, then Q equals the closure of $int\ Q$.

   (ii) If D is a closed set, what is the relation in general between the set D and the closure of $int\ D$?

   (iii) If U is an open set, what is the relation in general between the set U and the interior of $\overline{U}$?

5. If we denote the general point of R^2 by (x, y), determine $int\ A$, $ext\ A$ and $b(A)$ for the subset A of R^2 specified by each of the following conditions :

   (a) $x = 0$.

   (b) x and y are rational.

   (c) $0 \leq x < 1$.

   (d) x is rational and $y > 0$.

   (e) $0 < x^2 + y^2 < 1$.

   (f) $y < x^2$.

   (g) $0 \leq x < 1$ and $0 \leq y < 1$.

   (h) $y \leq x^2$.

2.3 Product of Sets

Definition 2.3.1. Let $A \subset R^n$ and $B \subset R^n$. The product of $A\&B$ denoted by $A \times B$ is defined as

$$\begin{aligned} A \times B &= \big\{(x^1, \dots, x^n, y^1, \dots, y^m) : (x^1, \dots, x^n) \\ &\quad \in A\&(y^1, \dots, y^m) \in B\big\}. \end{aligned}$$

If $x \in A\&y \in B$ then the point $(x^1, \dots, x^n, y^1, \dots, y^m)$ in $A \times B$ is denoted by (x, y).

This product is different from the cartesian product of two sets. From definition it is evident that $A \times B \subset R^{n+m}$, and in particular $R^n \times R^m = R^{n+m}$. This is not true if the product was cartesian.

Theorem 2.3.1. Let $A \subset R^n$ and $B \subset R^m$. Then

(i) $A\&B$ are open sets in R^n and R^m respectively iff $A \times B$ is open in R^{n+m}.

(ii) A and B are closed sets in R^n and R^m respectively iff $A \times B$ is closed in R^{n+m}.

Proof.

(i) Suppose $A\&B$ are open. Let $x \in A \times B$. Then $\exists$ an open rectangle $w_1 \subset R^n \ni (x^1, \dots, x^n) \in w_1 \subset A$, and an open rectangle $w_2 \subset R^m \ni (x^{n+1}, \dots, x^{n+m}) \in w_2 \subset B$. Hence $w_1 \times w_2$ is an open rectangle $\ni x \in w_1 \times w_2 \subset A \times B$. Hence $A \times B$ is open in R^{n+m}.

Conversely suppose $A \times B$ is on open set, and $x \in A$. Then pick $y \in B$. Now $(x^1, \dots, x^n, y^1, \dots, y^m) \in A \times B$. Since $A \times B$ is open $\exists$ a rectangle $w \subset R^n \ni (x, y) \in w \subset A \times B$. Let u be the open rectangle in R^n, and v be the open rectangle in R^m, such that $u \times v = w$. Clearly $x \in u \subset A$. Hence A is open. Similarly B is open.

(ii) Suppose A and B are closed, and $x \in R^{n+m}$ is not in $A \times B$. Then $x \in (A \times B)^c$. Clearly $(x^1, \dots, x^n) \notin A$ or $(x^{n+1}, \dots, x^{n+m}) \notin B$. Suppose the first alternative holds. Since A

is closed, A^c is open. But then $\exists$ an open rectangle $u \subset R^n \ni (x^1, \ldots, x^n) \in u \subset A^c$. Hence the open rectangle $u \times R^m \subset (A \times B)^c$, and contains x. Hence $(A \times B)^c$ is open and consequently $A \times B$ is closed. If the second alternative holds similar argument shows $A \times B$ is closed.

Conversely suppose $A \times B$ is closed but A or B is not closed. Assume the first alternative holds. Then $\exists x \in R^n, x$ is cluster point of A but it does not belong to A. Pick $y \in B$. Clearly $(x, y) \notin A \times B$. Now let w be an open rectangle in R^{n+m} containing (x, y). But $w = u \times \nu$ where u is an open rectangle in R^n containing x and ν is an open rectangle in R^m containing y. Since x is cluster point of $A \exists z \neq x \ni z \in A \cap u$. Hence $(z, y) \neq (x, y)$ and $(z, y) \in w \cap (A \times B)$. Hence (x, y) is a cluster point of $A \times B$. Hence $A \times B$ is not closed. Contradiction, hence A is closed. Similarly if B is not closed then $A \times B$ is also not closed. Therefore B is also closed.

Exercises 2.3

1. For each index i with $1 \leq i \leq n$, let F_i be a closed subset of R. Prove that the cartesian product

$$F_1 \times F_2 \times \cdots \times F_n$$

is closed subset of R^n.

2. For each index i with $1 \leq i \leq n$, let 0_i be an open subset of R. Prove that the cartesian product

$$0_1 \times 0_2 \times \cdots \times 0_n$$

is open subset of R^n.

3. If $(x, y) \in R^n$ and $(z, w) \in R^n$, then

   (i) $<(x, z), (y, w)> = <x, y> + <z, w>$.

   (ii) $|(x, z)| = \sqrt{|x|^2 + |z|^2}$.

4. Prove that:

   (a) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.

(b) $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$.

5. Prove that:

   (a) If $A \subset R^n$ and $B \subset R^n$, and x is a cluster point of $A \cap B$ then x is a cluster point of both $A \& B$.

   (b) If $A \subset R^n, B \subset R^n$, and x is a cluster point of $A \subset B$ then x is a cluster point of A or of B.

2.4 Compact Sets

The most important single concept in the topology of Euclidean space that is crucial to the study of calculus is that of the concept of compactness.

Definition 2.4.1. A collection O of open sets is an open cover of A if every point x in A is in some open set in the collection O.

Example 2.4.1. $(1/n, n), n \in N$ is an open cover of the set $(0, 1)$ in R.
$(-n, n), n \in N$ is an open cover of R.

Definition 2.4.2. A set $A \subset R^n$ is called compact if every open cover O of A has a finite sub cover of A.

Theorem 2.4.1. Heine-Borel Theorem. A closed and bounded interval in R^n is compact.

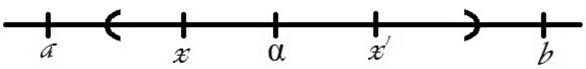

Fig. 2.4.1

Proof. Let $[a, b]$ be a closed and bounded interval in R^1, and O be an open cover of $[a, b]$. Define the set A by $A = \{x : x \in [a, b]$ and $[a, x]$ is covered by a finite sub collection of $O\}$. Clearly A is not empty because $a \in A$ and is bounded by b. Hence A has an upper bound. Call it α. We claim $\alpha \in A$ and $\alpha = b$. Suppose $\alpha \notin A$. Since O covers $[a, \alpha]$ there exists an open set $U \in O \ni \alpha \in U$. Let $a \leq x < \alpha \ni [x, \alpha] \subset U$. But then $\exists$ a finite sub collection O' of O

that covers $[a, x]$. Consequently $O' \cup \{U\}$ cover $[a, \alpha]$. Hence $\alpha \in A$ contradicting the fact $\Rightarrow \alpha \in A$.

Suppose $\alpha < b$. Let $U \in O \ni \alpha \in A$. Since U is open $\exists x' \ni \alpha < x' < b$ and $[\alpha, x'] \subset U$. But then the finite sub collection $O' \cup \{U\}$ cover $[\alpha, x'] \Rightarrow x' \in A \Rightarrow x' \leq \alpha$. Contradicting the fact. Hence $\alpha = b$. Therefore $[a, b]$ is compact.

Lemma 2.4.1. If $B \subset R^m$ is compact and $x \in R^n$ then $\{x\} \times B$ is compact in R^{n+m}.

Proof. Suppose O is an open cover of $\{x\} \times B$. Each $(x, y) \in \{x\} \times B$ is contained in some open set $O_y \in O$. Hence $\exists$ an open rectangle w_y in $R^{n+m} \exists (x, y) \in w_y \subset O_y$. But then $w_y = u_y \times \nu_y$ where u_y is an open rectangle in R^n containing x and ν_y is an open rectangle in R^m containing y. Consequently the collection on $\{\nu_y : y \in B\}$ is an open cover of B and since B is compact $\exists$ a finite sub collection $\{\nu_{y_1}, \ldots, \nu_{y_k}\}$ covering B. But then the open rectangles

$$u_{y_1} \times \nu_{y_1}, \ldots, u_{y_k} \times \nu_{y_k}$$

cover $\{x\} \times B$. Since each $u_{y_i} \times \nu_{y_i} \subset O_{y_i} \Rightarrow O_{y_1}, \ldots, O_{y_k}$ cover $\{x\} \times B \Rightarrow \{x\} \times B$ is compact.

Lemma 2.4.2. Suppose $B \subset R^m$ and $x \in R^n$. If B is compact and O is an open cover of $\{x\} \times B$, then there is an open set $u \subset R^n$ and containing x and such that $u \times B$ can be covered by a finite sub collection of O.

Proof. Each $(x, y) \in \{x\} \times B$ is contained in some open set $O_y \in O$. Hence $\exists$ a rectangle $w_y \in R^{n+m} \ni (x, y) \in w_y \subset O_y$. But then each $w_y = u_y \times \nu_y$, where $u_y \subset R^n$ is an open rectangle containing x and $\nu_y \subset R^m$ is open rectangle containing y. The collection $\{\nu_y : y \in B\}$ is an open cover of B. Since B is compact there exists a finite sub collection $\{\nu_y, \ldots, \nu_{y_k}\}$ that covers B. Now let $u = u_{y_1} \cap \cdots \cap u_{y_k}$. Clearly $x \in u$, and u is open. Hence the collection of open sets $\{u \times \nu_{y_1}, \ldots, u \times \nu_{y_k}\}$ cover $u \times B$. But $u \times \nu_{y_i} \subset O_{y_i}, i = 1, \ldots, k$. Hence the finite sub collection $O' = \{O_{y_1}, \ldots, O_{y_k}\}$ cover $u \times B$.

Theorem 2.4.2. If $A \subset R^n$ and $B \subset R^m$ are compact then $A \times B \subset R^{n+m}$ is compact.

Proof. Let O be an open cover of $A \times B$. For each $x \in A$. By Lemma 2.4.2 there exists an open set $u_x \subset R^n$ and such that $u_x \times B$ is covered by a finite sub collection O_x of O. But the collection $\{u_x : x \in A\}$ of open sets is a cover of A. And since A is compact there is a finite sub collection $\{u_{x_1}, \ldots, u_{x_k}\}$ that covers A. But then the union of the sub collections $O_{x_1}, \ldots, O_{x_k}$ of O cover $A \times B$. Since each sub collection O_{x_i} is finite their union is also finite. Hence a finite sub collection of O cover $A \times B$. Hence $A \times B$ is compact.

Corollary 2.4.1. If $A_i \subset R^{n_i}, i = 1, \ldots, k$ are compact then $A_1 \times A_2 \times \cdots \times \subset R^{n_1 + \cdots + n_k}$ is compact.

In particular any closed rectangle $w \subset R^k, k \in N$ is compact.

Theorem 2.4.3. $A \subset R^n$ is compact iff A is closed and bounded.

Proof. Suppose A is compact. Consider the open sets

$$u_n = \{x \in R^n : |x| < n\}, n \in N.$$

The collection $\{u_n : n \in N\}$ is an open cover of A. Since A is compact $\exists$ a finite sub collection $\{u_1, \ldots, u_n\}$ that cover A. But then since the open sets are increasing $A \subset u_n$. Hence A is bounded.

To show that A is closed. Let $c \notin A$. Consider the collection of closed sets $F_k = \{x \in R^n : |c - x| \le 1/k\}, k \in N$. Clearly $\cap_{k \in N} F_k = \{c\}$. Hence $A \subset (\cap_{k \in N} F_k)^c$. For each $k \in N$, define $H_k = F_k^c$. Each H_k is an open set and furthermore $A \subset (\cap_{k \in N} F_k)^c = \cup_{k \in N} F_k^c = \cup_{k \in N} H_k$. Hence the collection $\{H_k : k \in N\}$ is an open cover of A. Since A is compact $\exists$ a finite sub collection $\{H_{k_1}, \ldots, H_{k_i}\}$ that covers A. Let $k_0 = \max\{k_1, \ldots, k_i\}$. Obviously then $A \subset H_{k_0}$, and hence the interior of $F_{k_0} = \{x \in R^n : |x - c| \le 1/k_0\}$ contains c and does not intersect A. Hence c is not a cluster point of A. Hence A is closed.

Conversely, suppose A is closed and bounded. Then $\exists$ rectangle

$w \subset R^n \ni A \subset w$. Now suppose O is an open cover of A. Since A^c is open $\Rightarrow O \cup \{A^c\}$ is an open cover of w. Since w is compact $\exists$ a finite sub collection O' of O such that $O' \subset \{A^c\}$ cover w. But then O' covers A. Hence A is compact.

Applications of Compactness

Theorem 2.4.4. (Bolzano-Weierstrass Theorem)

Every bounded infinite subset of R^n has a cluster point.

Proof. Let A be an infinite subset of R^n. Suppose A has no cluster points. Hence each point of A is an isolated point. Consequently for each x in $A \exists$ an open set $u_x \ni u_x \cap A = \{x\}$. The collection $\{u_x : x \in A\}$ is an open cover of A. But A is closed because it contains all of its cluster points, and also bounded by hypothesis. Hence it is compact. So there is a finite sub collection of $\{u_x : x \in A\}$ that covers A. But this is impossible. Hence A has a cluster point.

Theorem 2.4.5. (Cantor Intersection Theorem)

Suppose F_1 is a nonempty, closed and bounded subset of R^n, and $\{F_i\}_{i \in N}$ is a sequence of nonempty closed sets in R^n such that $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \ldots$. Then $\cap_{i \in N} F_i \neq \phi$.

Proof. Suppose $\cap_{i \in N} F_i$ is empty. Consider the collection of open sets defined by $\{F_i^c : i \in N\}$. Since $(\cap_{i \in N} F_i)^c = \cup_{i \in N} F_i^c \Rightarrow \cup_{i \in N} F_i^c = R^n$. Hence $F_1 \subset \cup_{i \in N} F_i^c$. But F_1 is closed and bounded $\Rightarrow$ it is compact. Hence there is a finite number of open sets $F_1^c, \ldots, F_k^c$ that cover F_1. Now using the hypothesis, we know that $F_1^c \subset \cdots \subset F_k^c$. Hence $F_1^c \cup \cdots \cup F_k^c = F_k^c$. Hence $F_1 \subset F_k^c$. But $F_1 \cap F_k = \phi$. However $F_k \subset F_1$ by hypothesis and hence $F_1 \cap F_k = F_k \Rightarrow F_k$ is non empty set. Contradiction. Hence $\cap_{i \in N} F_i$ is non empty.

Exercises 2.4

1. Suppose that the function $f : R^n \to R$ is continuous, and $f(u) \geq |u|$ for every point $u \in R^n$. Prove that $f^{-1}([0,1])$ is compact.

2. Determine which of the following subsets of R is compact.

   (i) $\{x \text{ in } [0,1] : x \text{ is rational } \}$.

   (ii) $\{x \text{ in } R : x^2 - x\}$.

   (iii) $\{x \text{ in } R : e^x - x^2 \leq 0\}$.

3. Let A and B be compact subsets of R. Define $K = \{(x, y) \text{ in } R^2 : x \in A, y \in B\}$. Prove that K is compact.

4. Let u be a point in R^n, and let r be a positive number. Prove that the set $\{\nu \in R^n : d(u, \nu) \leq r\}$ is compact.

2.5 Dense and Nowhere Dense Sets in R^n

Definition 2.5.1. A set $A \subset R^n$ is said to be dense in $B \subset R^n$ if $\overline{A} \supset B$. If A is dense in R^n it is simply called a dense set.

Example.

1. Let A be the set of all rational numbers in [0,1] and B be the closed interval [0,1], then A is dense in B.

2. Let $A = Q$, the set of all rational numbers and $B = R^1$, then A is dense in B.

3. Let $A = Z$, set of all integers and $B = Q$, then A is not dense in B.

4. Let $A = \{1/n; n \in N\}$ and $B = [0, 1]$, then A is not dense in B.

5. Let $A = \{(x^1, \ldots, x^n) \in R^n : x^i \in Q\}$ and $B = R^n$ then A is dense in B.

Definition 2.5.2. A set $A \subset R^n$ is nowhere dense in $B \subset R^n$ if $B \cap (\overline{A})^c$ is dense in B. If A is nowhere dense in $B = R^n$ then A is called simply a nowhere dense set.

Example.

1. Let $A = \{1/n : n \in N\}$ and $B = [0, 1]$.
   $\overline{A} = A \cup \{0\}$
   $B \cap (\overline{A})^c = [0, 1] - [\{1/n : n \in N\} \cup \{0\}]$
   Hence $\overline{B \cap (\overline{A})^c} = [0, 1] \Rightarrow A$ is nowhere dense in B.

2. Let $A = Z$ and $B = R$.
   $B \cap (\overline{A})^c = \cup_{n \in Z}(n, n+1)$
   $\overline{B \cap (\overline{A})^c} = R$
   $\Rightarrow Z$ is nowhere dense in R.

Theorem 2.5.1. $A \subset R^n$ is nowhere dense iff $int(\overline{A}) = \phi$.

Proof. Suppose A is nowhere dense. Then by defenition $(\overline{A})^c$ is dense $\Rightarrow \overline{(\overline{A})^c} = R^n$. Now let $x \in R^n$. Then $x \in (\overline{A})^c$ or $x \in (\overline{A}^c)'$. If $x \in (\overline{A})^c$ then $x \notin \overline{A}$. Hence $x \notin int\overline{A}$. If $x \in (\overline{A}^c)'$ then for any ball $B(x), B(x) \cap \overline{A}^c \neq \phi$. Hence $x \notin int\overline{A}$. Hence $int\overline{A} = \phi$.

Conversely suppose that $int(\overline{A}) = \phi$. If $x \in \overline{A}$, then for any ball $B(x), B(x) \cap \overline{A}^c \neq \phi$. Hence $x \in (\overline{A})^c \Rightarrow \overline{A} \subset (\overline{A})^c$. But then $\overline{\overline{A}^c} = \overline{A}^c \cup (\overline{A}^c)' \supset \overline{A}^c \cup \overline{A} = R^n \Rightarrow \overline{A}^c$ is dense set $\Rightarrow A$ is nowhere dense.

Corollary 2.5.1. Let F be a closed set in R^n. F is nowhere dense iff it contains no open sets that is nonempty.

Proof. Suppose F is nowhere dense. By above theorem $int\overline{F} = \phi \Rightarrow F$ contains no open set that is nonempty. Conversely suppose F contains no nonempty open set $\Rightarrow int\overline{F}$ contains no nonempty open set. Hence F is nowhere dense.

Theorem 2.5.2. $A \subset R^n$ is nowhere dense iff for every nonempty open set O there is an open ball contained in $O \backslash A$.

Proof. Suppose A is nowhere dense $\Rightarrow int\overline{A} = \phi$. Now if O is nonempty open set, then $O \cap (\overline{A})^c \neq \phi$. Otherwise $O \subset \overline{A}$

which contradicts $int\overline{A} = \phi$. Therefore let $x \in O \cap \overline{A}^c$. But then since $\overline{A}^c$ is open $\exists$ an open ball $B(x)$ such that $B(x) \subset O \cap \overline{A}^c \Rightarrow B(x) \subset O \backslash \overline{A} \Rightarrow B(x) \subset O \backslash A$. Conversely suppose for every open set $O \neq \phi$ there is an open ball contained in $O \backslash A$. Now take any $x \in \overline{A}$, and let U be any open set containing x. Then there is a ball $B \subset U \backslash A$. Hence $B \subset A^c \Rightarrow B \backslash U$ bound $A = \phi$. Otherwise $B \cap A \neq \phi$ contradict $B \subset A^c$. Hence $B \subset \overline{A}^c$. But then $U \nsubseteq int(\overline{A}) \Rightarrow x \notin int\overline{A} \Rightarrow int\overline{A} = \phi$.

Theorem 2.5.3. Let $\{c_n : n \in N\}$ be a collection of nowhere dense sets in R^n. Then $\cup_{n \in N} c_n$ does not contain a nonempty open set.

Proof. Let O be any nonempty open set. Since c_1 is nowhere dense $\Rightarrow$ there is an open ball $B_1 \subset O - c_1$. Let $\overline{D}_1$ be a closed ball with same center as B_1 and having half of the radius of B_1. Now consider the open ball D_1. Again c_2 is nowhere dense $\Rightarrow$ there is an open ball $B_2 \subset D_1 - c_2$. Let $\overline{D}_2$ be a closed ball same center as B_2 and having half of the radius of B_2. Following the same procedure we construct by induction closed ball $\overline{D}_1, \overline{D}_2, \ldots, \overline{D}_n, \ldots$ such that each $\overline{D}_{n+1} \subset \overline{D}_n - c_{n+1}$ and $\overline{D}_{n+1} \subset \overline{D}_n \forall n \in N$. Since each closed ball $\overline{D}_n, n \in N$ is bounded by Cantor Intersection Theorem $\exists x_0 \in \cap_{n \in N} \overline{D}_n$. But then $x_0 \notin \cup_{n \in N} c_n$. However since each $\overline{D}_n \subset O, x_0 \in O$. Hence $O \nsubseteq \cup_{n \in N} c_n$.

Corollary 2.5.2. If $\{E_n : n \in N\}$ is a collection of dense open sets, then $\cap_{n \in N} E_n$ is nonempty.

Proof. For each n let $c_n = E_n^c$. Hence $c_n^c = E_n$. Since c_n is closed $\Rightarrow$ c_n is a nowhere dense set. Hence $\cup_{n \in N} c_n$ does not contain a nonempty open set. Hence $R^n \nsubseteq \cup_{n \in N} c_n \Rightarrow$ there exists $x \notin \cup_{n \in N} c_n$. But then $x \in (\cup_{n \in N} c_n)^c = \cap_{n \in N} c_n^c = \cap_{n \in N} E_n$.

Corollary 2.5.3. (Baire Theorem) R^n can not be the union of a countable nowhere dense sets.

Proof. Let $\{c_n : n \in N\}$ be a collection of nowhere dense sets in $R^n \Rightarrow \cup_{n \in N} c_n$ does not contain a nonempty open set $\Rightarrow R^n \nsubseteq \cup_{n \in N} c_n$.

Proposition 2.5.1. The set Q of rational numbers is not the intersection of a countable collection of open sets.

Proof. Assume $Q = \cap_{n\in N} U_n$, where U_n is open in R. Let $c_n = U_n^c$. Then $c_n \cap Q = \phi$ and c_n is closed. Hence c_n does not contain a nonempty open set $\Rightarrow c_n$ is nowhere dense. Since Q is countable let $\{x_n\}^\infty$, be an enumeration of Q. Now let $D_n = c_n \cup \{x_n\}, x \in N$. Each D_n is nowhere dense. But $R = U_{n\in N} D_n$. Contradicting the Bair Theorem. Hence Q is not the intersection of a countable collection of open sets.

Proof. 2.5.2. The set of all irrational numbers Q^c is not the union of a countable collection of closed sets in R.

Proof. Suppose $Q^c = \cup_{n\in N} c_n$ where c_n is a closed set. For each $n \in N$ define $O_n = c_n^c$. Then O_n is open, and furthermore $Q = (\cup_{n\in N} c_n)^c = \cap_{n\in N} c_n^c = \cap_{n\in N} O_n$.Contradicting to Proposition 2.5.1. Hence Q^c is not the union of a countable collection of closed sets in R.

2.6 Sequence in R^n

Definition 2.6.1. A function $f : N \to R^n$ is called a sequence in R^n.

As in the case of real numbers f is usually denoted or represented by $\{x_k\}$ where $x_k = f(k) \forall k \in N$.

Convergence of a sequence in R^n. A sequence $\{x_k\}$ in R^n converges to a limit x in R^n if for every $\varepsilon > 0 \exists$ a positive integer k_0 such that

$$|x - x_k| < \varepsilon \texttt{ for } k \geq k_0.$$

or $\lim_{k\to\infty} x_k = x$ or $x_k \to x$.

In the case of a sequence of real numbers a convergent sequence in R^n has a unique limit.

Theorem 2.6.1. Let $\{x_k\}$ be a sequence in $R^n . x_k \to x$ iff $x_k^i \to x^i$ for each $i = 1, \ldots, n$.

Proof. Suppose $\{x_k\}$ converges to x. Let $\varepsilon > 0$ then $\exists$ a positive integer k_0 such that $|x_k - x| < \varepsilon, k \geq k_0$.

But $|x_k^i - x^i| < |x_k - x| < \varepsilon \forall \ \ i = 1, \ldots, n$.

$\Rightarrow \{x_k^i\}$ converges to x^i for each $i = 1, \ldots, n$.

Conversely suppose $\{x_k^i\}$ converges to x^i for each $i = 1, \ldots, n$. Let $\varepsilon > 0$. Then for each $i = 1, \ldots, n \exists$ a positive integer k_i such that $|x_k^i - x^i| < \frac{\varepsilon}{n}$ whenever $k \geq k_i$.

Let $k_0 = \max\{k_1, \ldots, k_n\}$. Then for all $k \geq k_0$.

$|x_k - x| \leq |x_k^1 - x^1| + |x_k^2 - x^2| + \cdots + |x_k^n - x_k^n| < \frac{\varepsilon}{n} + \cdots + \frac{\varepsilon}{n} = \varepsilon$. $\Rightarrow x_k \to x$.

Theorem 2.6.2.

(i) A set $A \subset R^n$ is closed iff every convergent sequence $\{x_k\}$ in A has its limit in A.

(ii) For set $B \subset R^n, x \in \overline{B}$ iff $\exists$ a sequence $\{x_k\}$ in B which converges to x.

Proof.

(i) Suppose A is closed, and the sequence $\{x_k\}$ in A is convergent to x. Now if U is any open set containing x then $\exists k_0 \in N$ such that $\forall k \geq k_0 : x_k \in U$. Hence x is a cluster point of A. But since A is closed $\Rightarrow x \in A$.

Conversely suppose every convergent sequence in A has a limit in A. Let x be a cluster point of A. Let $x_1 \in B(x, 1) \cap A$ distinct from x and for each positive integer $k > 1$ let

$x_k \in B(x, 1/k) \cap A$ `and distinct from` $x_1, \ldots, x_{k-1}$.

Then x_k is a sequence in A with limit $x \Rightarrow x \in A \Rightarrow A$ is closed.

(ii) Proof is similar to (i).

Definition 2.6.2. A sequence $\{x_k\}$ in R^n is a Cauchy sequence if for every $\varepsilon > 0, \exists$ positive integer k_0 such that

$$|x_m - x_l| < \varepsilon \texttt{ whenever } m, l \geq k_0.$$

Theorem 2.6.3. A sequence $\{x_k\}$ in R^n converges to a point in R^n iff it is a Cauchy sequence.

Proof. Suppose $x_k \to x$, and $\varepsilon > 0$. There exists $k_0 \in N$ such that $\forall k \geq k_0, |x_k - x| < \frac{\varepsilon}{2}$. Hence $\forall x, l \geq k_0$

$$|x_k - x_l| \leq |x_k - x| + |x_l - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $\{x_k\}$ is a Cauchy sequence.

Conversely, suppose $\{x_k\}$ is a Cauchy sequence. For each $i = 1, \ldots, n$. $|x_k^i - x_l^i| \leq |x_k - x_l| \Rightarrow$ for each $i = 1, \ldots, n, \{x_k^i\}$ is a Cauchy sequence. Hence for each $i = 1, \ldots, n, \{x_k^i\}$ converges to some real numbers x^i. Hence $x_k \to x$.

Example 1. $f : N \to R^3$

$f(n) = \left(-\frac{1}{n}, \frac{1}{n^2}, 2n\right)$

$x_k = \left(-\frac{1}{k}, \frac{1}{k^2}, 2k\right)$

$\{x_k\}_{k=0}^{\infty}$ is a sequence.

Example 2. $x_k = \left(-\frac{1}{k}, \frac{2}{k^2}\right)$

$\lim_{k\to\infty} x_k = (0, 0)$

then $\{x_k\} = \left\{-\frac{1}{k}, \frac{2}{k^2}\right\}_{k=1}^{\infty} \to (0, 0)$

Example 3. Let $\{x_k\} = \left(-\frac{1}{k}, 1 - \frac{1}{k}\right)$, show that $\{x_k\} \to (0, 1)$.

Solution. $\forall \varepsilon > 0, \exists k_0 \in N$ such that $\left|\left(-\frac{1}{k}, 1 - \frac{1}{k}\right) - (0, 1)\right| < \varepsilon \forall k \geq k_0$

$$\begin{aligned}\left|\left(-\frac{1}{k}, 1 - \frac{1}{k}\right) - (0, 1)\right| &= \sqrt{(1/k)^2 + (-\frac{1}{k})^2} = \sqrt{2}\left(\frac{1}{k}\right) < \varepsilon \\ &\Rightarrow \; k > \frac{\sqrt{2}}{\varepsilon} \quad \forall \varepsilon > 0, k_0 > \frac{\sqrt{2}}{\varepsilon}.\end{aligned}$$

Then $\left|\left(-\frac{1}{k}, 1-\frac{1}{k}\right)-(0,1)\right|<\varepsilon \ \ \forall k \geq k_0 \geq \frac{\sqrt{2}}{\varepsilon}$.

Theorem 2.6.4. A sequence $\{x_k\} \rightarrow x$ iff for every open set U containing x, there is a positive integer k_0 such that $x_k \in U \forall k \geq k_0$.

Proof. Suppose $\{x_k\} \rightarrow x$ and let U be any open set containing $x, x \in U$ and U is open $\Rightarrow \exists \varepsilon>0, \ni B_{\varepsilon}(x) \subseteq U$.

$\{x_k\} \rightarrow x \Leftrightarrow for \varepsilon>0 \exists k_0 \in N$ `such that` $|x_k-x|<\varepsilon \forall k \geq k_0$

$\Leftrightarrow x_k \in B_{\varepsilon}(x) \forall k \geq k_0$. Thus $\forall$ open set $U \exists k_0$ such that $\forall k \geq k_0, x_k \in U$.

Conversely, suppose that $\forall$ open set U containing $x \exists k_0 \in N$ such that $\forall k \geq k_0, x_k \in U$. We have to prove that $\{x_k\} \rightarrow x$. Let $\varepsilon>0$, consider $B_{\varepsilon}(x), B_{\varepsilon}(x)$ is an open set then $\exists k_0 \in N$ such that $\forall k \geq k_0, \quad x_k \in B_{\varepsilon}(x) \quad \Leftrightarrow |x_k-x|<\varepsilon, \forall k \geq k_0 \Rightarrow \{x_k\} \rightarrow x$.

Theorem 2.6.5. Given a set $A \subset R^n, x \in \overline{A}$ iff there exists a sequence $\{x_k\}$ in A which converge to x.

Proof. Let $A \subset R^n$ and $x \in \overline{A}$ we have to prove that $\exists$ a sequence $\{x_k\}$ in $A \rightarrow x$.

$x \in \overline{A} \Rightarrow x \in A$ or $x \in A'$

If $x \in A$, then consider the constant sequence $x_k=x \forall k$. Then $\{x_k\} \rightarrow x$.

Suppose $x \in A'$, construct a sequence of terms in A that converges to x.

Let $\{x_n\}$ be a sequence in A and $\{x_n\} \rightarrow x$. Then we have to show that $x \in A$.

$\{x_n\} \rightarrow x \Rightarrow \forall \varepsilon>0, \exists k_0 \in N$ `such that` $x_k \in B_{\varepsilon}(x) \forall k \geq k_0$

$$B_{\varepsilon}(x) \cap A \neq \phi \quad \Rightarrow \quad \begin{array}{l}(i) \ \ B_{\varepsilon}(x) \cap A \backslash\{x\} \neq \phi \\ (ii) B_{\varepsilon}(x) \cap A \backslash\{x\}=\phi .\end{array}$$

If (i) then $x \in A' \subseteq A \Rightarrow x \in A$.

(ii) then $x_n = x \ \ \forall k \geq k_0$ but $x_n \in A \Rightarrow x \in A$.

Conversely suppose that $\{x_n\}$ is a sequence in A that converges to x. Then we have to show that $x \in \overline{A} = A \cup \overline{A}$ .If $x \in A$, then nothing to prove.

Suppose $x \notin A, \{x_n\} \rightarrow x \Rightarrow \forall \varepsilon > 0, \exists k_0$ such that $|x_k - x| < \varepsilon \forall k \geq k_0$, in particular $\forall r > 0, B_r(x)$ contains infinitely many points $\exists \{x_n\}$ have infinitely many points of A.

i.e., $B_r(x) \cap A \backslash \{x\} \neq \phi \Rightarrow A^{'} \Rightarrow x \in \overline{A}$.

Lebsegue Covering Theorem.

Let A be a compact set in R^n and O an open cover of A. $\exists$ a positive number β such that $\forall x, y \in A$, if $|x - y| < \beta$ then there is an open set in O containing $x \& y$.

Proof. For each $x \in A$, let B_x be a ball with center x and contained in some open set in O. Let c_x be an open ball with center x and radius half of that B_x. The collection $\{c_x : x \in A\}$ of open balls covers A. Since A is compact $\exists$ a finite sub collection $\{c_{x_1}, \ldots, c_{x_k}\}$ covering A. Let $\beta = \min\{r(c_{x_1}), \ldots, r(c_{x_k})\}.r(c_{x_i})$ means radius of c_{x_i}. Now suppose $x \& y \in A$ such that $|x - y| < \beta$. Then $x \in c_{x_i}$. But then both $x \& y \in B_{x_i}$. Hence $x \& y$ are both contained by some open set in O that contains the ball B_{x_i}.

Problem 1. Prove that

If $A \subset R^n$ then $A \cup b(A) = \overline{A} = intA \cup b(A)$.

Solution.

$$\begin{aligned} b(A) &= \overline{A} \cap (\overline{A}^c) \\ A \cup b(A) &= A \cup (\overline{A} \cap \overline{A}^c) \\ &= (A \cup \overline{A}) \cap (A \cup \overline{A}^c), \quad A \subset \overline{A} \\ &= \overline{A} \cap (A \cup A^c), \quad \overline{A}^c \subset A^c \\ &= \overline{A} \cap R^n = \overline{A}. \end{aligned}$$

$$\begin{aligned} intA \cup b(A) &= intA \cup (\overline{A} \cap \overline{A}^c) \\ &= (intA \cup \overline{A}) \cap (intA \cup \overline{A}^c), \quad intA \subset A \subset \overline{A} \\ &= \overline{A} \cap \{intA \subset (intA)^c) \\ &= \overline{A} \cap R^n = \overline{A}. \end{aligned}$$

Exercises 2.6

1. Prove Pythagorean Theorem.

   If $x \& y$ are in R^n, and are perpendicular then
   $$|x+y|^2 = |x|^2 + |y|^2.$$

2. Prove Parallelogram law.

   If $x \& y$ are in R^n, then
   $$|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2.$$

3. If $x \& y$ are nonzero vectors in $R^n, < x, y >= 0$, and α, β are real numbers such that $\beta > \alpha > 0$ then prove that
   $$|\beta x + \beta y| > |\alpha x + \beta y| > |\alpha x + \alpha y|.$$

4. Suppose $T : R^n \to R^n$ is a l.t. and there exist an orthonormal vectors $x_1, \ldots, x_n$ in R^n and real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $T_x = \lambda_i x_i \forall i$. Then T is angle preserving iff all $|\lambda_i|$ are equal.

5. Let $T : R^n \to R^n$ be a function that preserve inner product. Prove that T is additive. That is $\forall x, y$ in R^n
   $$T(x+y) = T(x) + T(y).$$

6. Show that $|x+y||x-y| \leq |x|^2 + |y|^2$.
   $$(|x+y| - |x-y|)^2 \geq 0.$$

7. If $T : R^m \to R^m$ is a l.t.. Show that there is a number M s.t. $|T(h)| \leq M|h|$ for $h \in R^m$.

8. Let $(R^n)^*$ denote the dual space of the vector space R^n. If $x \in R^n$ define $\varphi_x \in (R^n)^*$ by $\varphi_x(y) =< x, y >$. Define $T : R^n \to (R^n)^*$ by $T(x) = \varphi_x$. Show that T is a $1-1$ l.t. and conclude that every $\varphi \in (R^n)^*$ is φ_x for a unique $x \in R^n$.

9. Let $A \subset R^n$. Prove that

   (a) $int\ A \subset A$.

   (b) $int\ A(int\ A) = int\ A$.

   (c) $int\ (A \cap B) = int\ A \cap int\ B$.

Chapter 3

Functions of Several Variables

So far we have discussed the calculus of functions of single variable. But, in the real world, physical quantities usually depends on two or more variables, so in this chapter we turn our attention for functions of several variables and extend the basic ideas of differential calculus of such functions.

3.1 Definitions and Properties

The temperature T at a point on the surface of the earth at any given time depends on the longitude x and latitude y of the point. We think T as being a function of the two variables x and y or as a function of the point (x, y) denoted by $T = f(x, y)$.

The volume V of a circular cylinder depends on its radius r and height h. Here $V = \pi r^2 h$ and V is a function of r and h. Denoted by $V = \pi r^2 h = V(r, h)$.

Definition 3.1.1. Let $D \subseteq R^2 = R \times R$. A function f of two variables is a rule that assigns to each ordered pair (x, y) in D a unique real number denoted by $f(x, y)$. The set D is the domain of f and its range is the set of values that f takes on i.e., range $= \{f(x, y) : (x, y) \in D\}$.

Remark 3.1.1. $f : D \subseteq R^2 \to R$. For $(x, y) \in D \subseteq R^2, f(x, y) \in R$.

Remark 3.1.2. We write $z = f(x, y)$. The variables x and y are independent variables and z is dependent variable.

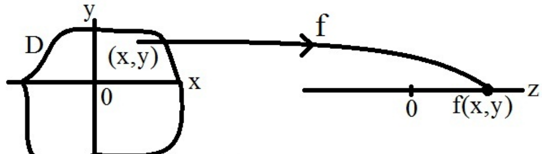

Fig. : 3.1.1

Remark 3.1.3. The value of a function f of three variables at a point (x, y, z) is denoted by $f(x, y, z)$ i.e., $f : D \subseteq R^3 \rightarrow R$ such that $(x, y, z) \in D \Rightarrow (x, y, z) \in R^3$ and $f(x, y, z) \in R$.

- x, y, z are independent variables.
- $U = f(x, y, z)$ is dependent variables.

Example 3.1.1.

(a) $f(x, y) = xy$ for $x, y \geq 0$, area of a rectangle.

(b) $f(x, y, z) = xyz$ for $x, y, z \geq 0$, volume of a rectangular parallelopiped.

Example 3.1.2. Find the domain of each of the functions and evaluate $f(3, 2)$

(a) $f(x, y) = \frac{\sqrt{x+y+1}}{x-1}$ (b) $f(x, y) = x\ln(y^2 - x)$.

Solution.

(a)

$$\begin{aligned} \texttt{Domain } D &= \{(x, y) : x + y + 1 \geq 0 \texttt{ and } x \neq 1\} \\ &= \{(x, y) : y = -x - 1 \texttt{ and } x \neq 1\} \\ f(3, 2) &= \frac{\sqrt{3 + 2 + 1}}{3 - 1} = \frac{\sqrt{6}}{2}. \end{aligned}$$

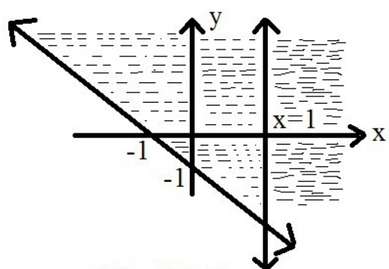

Fig. 3.1.2

(b)

$$\begin{aligned} f(x,y) = xln(y^2 - x).\ \texttt{Domain} &= \{(x,y) : y^2 - x > 0\} \\ &= \{(x,y) : x < y^2\} \\ f(3,2) = 3ln(2^2 - 3) = 3ln1 &= 0. \end{aligned}$$

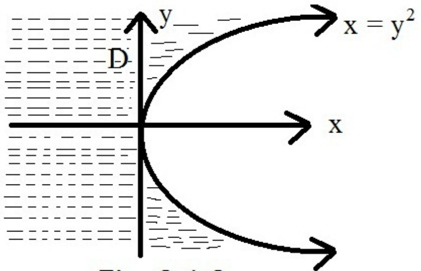

Fig. 3.1.3

3.2 Graphs and Level Curves

Definition 3.2.1. If f is a function of two variables with domain D, the graph of f is the set $S = \{(x,y,z) \in R^3 : z = f(x,y)$ for $x, y \in D\}$. This is a surface.

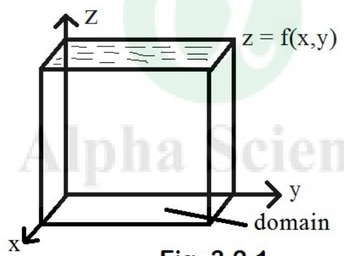

Fig. 3.2.1

Remark 3.2.1. Here we can consider combinations of functions of several variables.

(i) For functions f and g of two variables, the following holds.

(a) $(f \pm g)(x,y) = f(x,y) \pm g(x,y)$.

(b) $(fg)(x,y) = f(x,y).g(x,y)$.

(c) $(f/g)(x,y) = \frac{f(x,y)}{g(x,y)}$ where $g(x,y) \neq 0$.

(ii) The formulas for functions of three variables analogous.

$*$ In both cases domain $(f \pm g)$ = domain (fg) = domain $f \cap$ domain g and domain (f/g) = (domain $f \cap$ domain g) : $\{(x,y) : g(x,y) \neq 0\}$.

Remark 3.2.2. If f is a function of two (three) variables and g is a function of one variable then the composition gof is defined and

$*$ $(gof)(x,y) = g(f(x,y)), (x,y) \in$ domain f and $f(x,y) \in$ domain g.

$*$ $(gof)(x,y,z) = g(f(x,y,z)), (x,y,z) \in$ domain f and $f(x,y,z) \in$ domain g.

Example 3.2.1. Let $f(x,y) = x^2 - y^2, g(x,y) = 2x^2 + y$ and $h(t) = 2t^2 + t$. Then

1. $(f+g)(x+y) = f(x,y) + g(x,y) = 3x^2 - y^2 + y.$
2. $(f-g)(x,y) = f(x,y) - g(x,y) = -x^2 - y^2 - y.$
3. $(f.g)(x,y) = f(x,y).g(x,y) = 2x^4 - 2x^2y^2 + x^2y - y^3.$
4. $(f/g)(x,y) = \frac{f(x,y)}{g(x,y)} = \frac{x^2-y^2}{2x^2+y} \quad y \neq -2x^2.$
5. $(hof)(x,y) = h(f(x,y)) = h(x^2-y^2) = 2(x^2-y^2)^2+(x^2-y^2).$
6. $(hog)(x,y) = h(g(x,y)) = h(2x^2+y) = 2(2x^2+y)^2+(2x^2+y).$

Example 3.2.2. Find the domain, range and sketch the graph of $f(x,y) = \sqrt{9 - x^2 - y^2}$.

Solution. Domain of $f = \{(x,y) : 9 - x^2 - y^2 \geq 0\} = \{(x,y) : x^2 + y^2 \leq 9\}$ which is the disc with center $(0,0)$ and radius 3.

Range of $f = \{z : z = \sqrt{9 - x^2 - y^2}, (x,y) \in \text{domain } f\}$.

Here $z \geq 0$ and $0 \leq z = \sqrt{9 - x^2 - y^2} \leq 3 \Rightarrow$ range $= \{z : 0 \leq z \leq 3\} = [0,3]$.

Example 3.2.3. Sketch the graph of function $h(x,y) = x^2 + y^2$.

Solution. $h(x,y) = x^2 + y^2$, domain $= R^2$,

Range $= \{z : z = h(x,y) = x^2 + y^2\}, 0 \leq z < \infty$,

Definition 3.2.2. If a surface is the graph in three space of an equation of second degree. It is called a quadratic surface. The

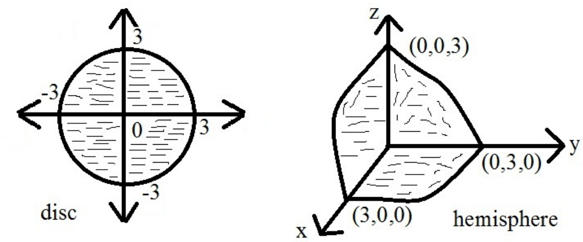

Fig. 3.2.3

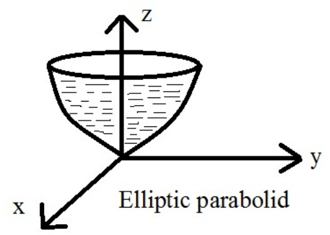

Fig. 3.2.3

second degree equation has the form

$$Ax^2 + By^2 + cz^2 + Dxy + Exz + Fyz + Hy + Iz + J = 0.$$

$*$ Ellipsoid, elliptic parabolid, hyperbolic parabolid and elliptic cone are quadratic surfaces.

Level Curves. In general it is not simple to sketch the graph of functions of several variables. To overcome this problem we can sketch what we call, its level curves.

Definition 3.2.3. The level curves of a function f of two variables are the curve with equation $f(x, y) = K$, where K is a constant (in the range).

Remark 3.2.3. For a function f of three variables $f(x, y, z) = K$ where $k \in$ range f is a level curve (contour line).

Example 3.2.4. Sketch the level curves of the function $f(x, y) = 6 - 3x - 2y$ for the values $K = -6, 0, 6, 12$.

Solution. The level curves are $6 - 3x - 2y = K \Rightarrow 3x + 2y +$

$(K-6)=0$. Now

$$K=-6 \Rightarrow 3x+2y-12=0 \Rightarrow y=-\frac{3}{2}x+6.$$
$$K=0 \Rightarrow 3x+2y-6=0 \Rightarrow y=-\frac{3}{2}x+3.$$
$$K=6 \Rightarrow 3x+2y=0 \Rightarrow y=-\frac{3}{2}x$$
$$K=12 \Rightarrow 3x+2y+6=0 \Rightarrow y=-\frac{3}{2}x-2.$$

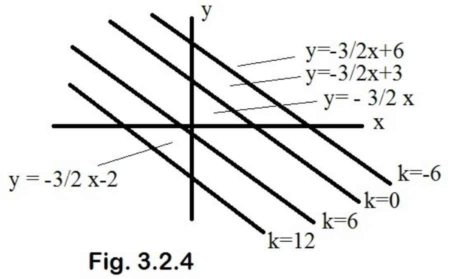

Fig. 3.2.4

Remark 3.2.4. The contour line (level curve) where f assume the value K shows exactly where the graph of f intersects the plane $\{(x,y,z) : z=K\}$.
Remark 3.2.5. If $f(x,y,z)$ denotes the temperature at any point (x,y,z) in a space, then the level surface $f(x,y,z)=K$ is the surface on which the temperature is constantly K and is called an isothermal surface.

Exercises 3.2

1. Let $f:[a,b]\to R$. The graph of f is the subset
$$G_f=\{(x,y) : y=f(x)\}$$
of R^2. Show that if f is continuous, G_f has measure zero in R^2.

2. Find the range of
   (a) $g(x,y)=\frac{1}{\sqrt{xy}}$.
   (b) $g(x,y)=\frac{2xy}{x^2+y^2}$.

3. Sketch the graph of $\frac{x^2}{16}+\frac{y^2}{25}+\frac{z^2}{9}=1$.

4. Sketch the graph of $z = \frac{y^2}{b^2} - \frac{x^2}{a^2}$.

5. Sketch the level curves of the function $g(x, y) = \sqrt{9 - x^2 - y^2}$ for $k = 0, 1, 2, 3$.

6. Sketch the level curves of the functions:

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

   (b) $f(x, y) = \frac{1}{\sqrt{x^2+y^2}}$.

   (c) $h(x, y) = 4x^2 + y^2$.

7. Sketch the graph of the function

   (a) $f(x, y) = 3$.
   Hint. Graph $= \{(x, y, z) : z = 3\}$.

   (b) $f(x, y) = x$.

   (c) $f(x, y) = x^2 + 9y^2$.

   (d) $f(x, y) = \sqrt{16 - x^2 - 16y^2}$.

8. Draw a level curve of the following functions:

   (a) $f(x, y) = xy$.

   (b) $f(x, y) = x^2 + 9y^2$.

   (c) $f(x, y) = x/2$.

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

9. The magnitude of the gravitational force exerted on a unit mass at (x, y, z) by a point mass located at the origin is given by

$$F(x, y, z) = \frac{c}{x^2 + y^2 + z^2}$$

   where c is a positive constant. Describe the level surface of F.

10. Suppose a thin metal plate occupies the first quadrant of the $x-y$ plane and the temperature at (x, y) is given by $T(x, y) = xy$. Describe the isothermal curves i.e., the level curves of T.

3.3 Limits and Continuity

Limits

Let f be a function of two variables, as (x, y) approaches (x_0, y_0), $f(x, y)$ approaches L (or L is the limit of $f(x, y)$ if $f(x, y)$ is as close to L as we wish whenever (x, y) is close enough to (x_0, y_0) or $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$.

Definition 3.3.1. Let f be a function of two variables defined on a disk with center (x_0, y_0) except possibly at (x_0, y_0). Then we say that the limit of $f(x, y)$ as $(x, y) \to (x_0, y_0)$ is L and we write $\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$, if for every number $\varepsilon > 0, \exists$ a number $\delta > 0$ such that

$$0 < \sqrt{(x - x_0)^2 + y - y_0)^2} < \delta \Rightarrow |f(x, y) - L| < \varepsilon$$

or

$$0 < \|(x, y) - (x_0, y_0)\| < \delta \Rightarrow |f(x, y) - L| < \varepsilon$$

or

$$0 < \|P - P_0\| < \delta \Rightarrow |f(P) - L| < \varepsilon \texttt{ for } P = (x, y), P_0 = (x_0, y_0).$$

Notation.

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = \lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = L$$

.

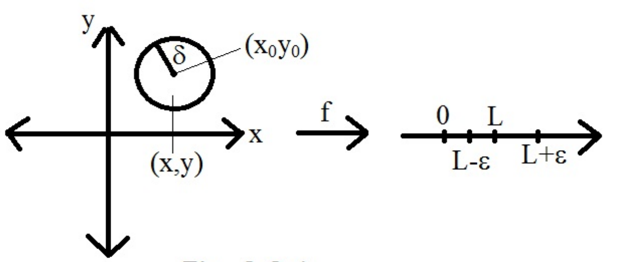

Fig. 3.3.1

Example 3.3.1. Show that $\lim_{(x,y)\to(x_0,y_0)} x = x_0$ and $\lim_{(x,y)\to(x_0,y_0)} y = y_0$.

Solution. To show $\lim_{(x,y)\to(x_0,y_0)} x = x_0$, given $\varepsilon > 0, \exists\delta > 0$ such that

$$0 < |(x, y) - (x_0, y_0)| < \delta \Rightarrow |x - x_0| < \varepsilon.$$

$\Rightarrow$

$$0 < \sqrt{(x - x_0)^2} < \delta \Rightarrow |x - x_0| < \varepsilon.$$

Now $\delta = \varepsilon$ then $0 < |x - x_0| < \delta = \varepsilon \Rightarrow |x - x_0| < \varepsilon$.

Similarly, we can prove $\lim_{(x,y)\to(x_0,y_0)} y = y_0$.

Note 3.3.1. For functions of single variable if $\lim_{x\to a^+} f(x) \neq \lim_{x\to a^-} f(x)$ then $\lim_{x\to a} f(x)$ does not exist. But for functions of two or more variables the situation is not simple, we let (x, y) approaches (x_0, y_0) from an infinite number of directions.

$*$ If the limit exists, then $f(x, y)$ must approaches the same limit no matter how $(x, y) \to (x_0, y_0)$.
$*$ If we find two different paths of approach along which $f(x, y)$ has different limit, then it follows that $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ does not exist.

If $f(x, y) \to L_1$ as $(x, y) \to (x_0, y_0)$ along a path c_1 and $f(x, y) \to L_2$ as $(x, y) \to (x_0, y_0)$ along a second path c_2 with $L_1 \neq L_2$ then $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ does not exist.

Example 3.3.2. Find $\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$ if it exists.

Solution. Let $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$, along the x axis $(y = 0)$ the function $f(x, 0) = \frac{x^2}{x^2} = 1$ for all $x \neq 0$.

$$f(x, y) \to 1 \texttt{ as } (x, y) \to (0, 0).$$

Along the y axis $(x = 0)$ and $f(0, y) = \frac{-y^2}{y^2} = -1, y \neq 0$.

$$f(x, y) \to -1 \texttt{ as } (x, y) \to (0, 0).$$

Hence f has two different limits along different lines. So the given limit does not exist.

$$\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2} = \nexists.$$

Example 3.3.3. If $f(x,y) = \frac{xy}{x^2+y^2}$ does $\lim_{(x,y)\to(0,0)} f(x,y)$ exists.

Solution. If $y = 0$ then $f(x,0) = 0$. Thus $f(x,y) = 0$ as $(x,y) \to (0,0)$. If $x = 0$ then $f(0,y) = 0$. Thus $f(x,y) \to 0$ as $(x,y) \to (0,0)$.

Along the line $y = x$ for $x \neq 0, f(x,x) = \frac{x^2}{2x^2} = 1/2$.
Thus $f(x,y) \to \frac{1}{2}$ as $(x,y) \to (0,0)$.

Along the line $y = -x, x \neq 0, f(x,-x) = \frac{-x^2}{2x^2} = -\frac{1}{2}, f(x,y) = -\frac{1}{2}$ as $(x,y) \to (0,0)$. Different limit along different path implies the limit does not exist.

Example 3.3.4. Find $\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2}$ if it exists.

Solution. One can show that the limit along any line through the origin is 0. To show this
Claim

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

Using the definition given $\varepsilon > 0, \exists\delta > 0$ such that if

$$0 < \sqrt{x^2+y^2} < \delta \Rightarrow \left|\frac{3x^2y}{x^2+y^2} - 0\right| < \varepsilon \Rightarrow \frac{3x^2|y|}{x^2+y^2} < \varepsilon.$$

Now $x^2 \leq x^2 + y^2 \Rightarrow \frac{x^2}{x^2+y^2} \leq 1 \Rightarrow \frac{3x^2|y|}{x^2+y^2} \leq 3|y|$
or

$$3\sqrt{y^2} \leq 3\sqrt{x^2+y^2} < \varepsilon \Rightarrow \sqrt{x^2+y^2} < \varepsilon/3.$$

If we choose $\delta = \varepsilon/3$ and $0 < \sqrt{x^2+y^2} < \delta = \varepsilon/3$, then

$$\left|\frac{3x^2y}{x^2+y^2} - 0\right| \leq 3\sqrt{x^2+y^2} < 3\delta = 3(\varepsilon/3) = \varepsilon.$$

Thus

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

Example 3.3.5. Show that $\lim_{(x,y)\to(-1,+1)} \frac{x^3+y^3}{x^2+y^2} = 0$.

Solution. Observe that $\lim_{(x,y)\to(-1,1)} x = -1$ and $\lim_{(x,y)\to(-1,1)} y = 1$. By this rule $\lim_{(x,y)\to(-1,1)} x^3 = -1, \lim_{(x,y)\to(-1,1)} y^3 = 1$, $\lim_{(x,y)\to(-1,1)} x^2 = 1$ and $\lim_{(x,y)\to(-1,1)} y^2 = 1$.

Thus

$$\lim_{(x,y)\to(-1,1)} \frac{x^3+y^3}{x^2+y^2} = \frac{-1+1}{1+1} = \frac{0}{2} = 0.$$

Continuity

Definition 3.3.2. Let f be a function of two variables defined on disk with center (a, b). Then f is said to be continuous at (a, b) if

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b).$$

Remark 3.3.1. If *dom* $f = D \subseteq R^2$ then the above definition of continuity of f is defined at the interior point of D.

Remark 3.3.2. For boundary point of D i.e., if a point (a, b) is a boundary point of D then the definition for continuity is given by $\varepsilon > 0, \exists \delta > 0$ such that

$$0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta \Rightarrow |f(x, y) - L| < \varepsilon, (x, y) \in D.$$

Example 3.3.6. If $f(x, y) = x$, prove that f is continuous on R^2.

Proof. Let $(a, b) \in R^2$ then we have to prove that

$$\lim_{(x,y)\to(a,b)} f(x, y) = \lim_{(x,y)\to(a,b)} x = a = f(a, b).$$

Given $\varepsilon > 0, \exists \delta > 0$ such that

$$\begin{aligned}
& 0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta \Rightarrow |x-a| < \varepsilon \\
\Rightarrow\quad & |x-a| \le \sqrt{(x-a)^2 + (y-b)^2} < \delta \Rightarrow |x-a| < \varepsilon \\
\Rightarrow\quad & |x-a| < \delta \Rightarrow |x-a| < \varepsilon. \texttt{ Now choose } \delta = \varepsilon \\
& \lim_{(x,y)\to(a,b)} f(x, y) = f(a, b) \texttt{ i.e., continuous at } (a, b) \in R^2.
\end{aligned}$$

Example 3.3.7. Where the function $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$ is continuous ?

Solution. The function f is discontinuous at $(0,0)$ because it is not defined there. Because f is a rational function, it is continuous on its domain which is the set $D = \{(x,y) : (x,y) \neq (0,0)\}$.

Example 3.3.8. Let

$$g(x,y) = \begin{cases} \frac{x^2-y^2}{x^2+y^2} & if \quad (x,y) \neq (0,0) \\ 0 & if \quad (x,y) = (0,0) \end{cases}.$$

Here g is defined at $(0,0)$ but g is discontinuous at $(0,0)$ because $\lim_{(x,y)\to(0,0)} g(x,y)$ does not exist.

Example 3.3.9. Let

$$f(x,y) = \begin{cases} \frac{3x^2y}{x^2+y^2} & if \quad (x,y) \neq (0,0) \\ 0 & if \quad (x,y) = (0,0) \end{cases}.$$

Show that f is continuous on R^2.

Solution.

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = 0 = f(0,0).$$

So f is continuous at $(0,0)$ and so it is continuous on R^2.

Example 3.3.10. Let R be the rectangular region consisting of all points (x,y) such that $0 \leq x \leq 1$ and $0 \leq y \leq 2$. Let

$$f(x,y) = \begin{cases} 4-x-y & \texttt{for} \quad (x,y) \in R \\ 0 & \texttt{for} \quad (x,y) \notin R \end{cases}.$$

Show that f is continuous on R but f is not continuous function.

Solution. Since the polynomial $4-x-y$ is continuous on R, f is also continuous on R.

$$f(x,y) = 4-x-y = 4-(x+y), f_{min} = 4-(1+2) = 1$$
$$\texttt{when } (x,y) = (1,2).$$

$f(x, y) = 1 \forall (x, y) \in R \Rightarrow f(x, y) \neq 0$ for $(x, y) \in R$. But $f(x, y) = 0 \forall (x, y) \notin R$. By definition ,$f$ has no limit point at any boundary point of R and thus f is not continuous.

Exercises 3.3

1. If $f(x, y) = \frac{xy^2}{x^2+y^4}$ does $\lim_{(x,y)\to(0,0)} f(x, y)$ exists ?

   Hint. $y = mx, x = y^2, \nexists$.

2. (a) If limit exists at a point then it is unique.

   (b) If $\lim_{p\to p_0} f(p) = a$ and $\lim_{p\to p_0} g(p) = b$ then

      i. $\lim_{p\to p_0}(f(p) \pm g(p)) = a \pm b$.

      ii. $\lim_{p\to p_0} f(p)g(p) = ab$.

      iii. $\lim_{p\to p_0} cf(p) = ca, c$ is a constant.

      iv. $\lim_{p\to p_0} \frac{f(v)}{g(p)} = \frac{a}{b}$ provided that $b \neq 0$.

3. If $g(x, y) = y$ prove that g is continuous on R^2.

4. If $f(x, y) = c$ (a constant function) then prove that f is continuous on R^2.

5. If functions f and g are continuous at a point $p = (a, b)$ in R^2, then

   (a) $f \pm g$ is continuous at $p = (a, b)$.

   (b) fg is continuous at $p = (a, b)$.

   (c) f/g is continuous at $p = (a, b), g(p) \neq 0$.

6. If f is continuous at (a, b) and g is a function of a single variable that is continuous at $f(a, b)$ then the composite function $h = gof$ defined by $h(x, y) = g(f(x, y))$ is continuous at (a, b).

7. Let

$$g(x, y) = \begin{cases} \frac{3x^2y}{x^2+y^2} & if \quad (x, y) \neq (0, 0) \\ 17 & if \quad (x, y) = (0, 0) \end{cases}$$

   show that g is not continuous at $(0, 0)$.

8. Find the limit,if it exists or show that the limit does not exists

(a) $\lim_{(x,y)\to(0,3)}(x^2y^2 - 2xy^4 + 3y)$.

(b) $\lim_{(x,y)\to(0,0)} \frac{x^2y^3+x^3y^2-5}{2xy}$.

(c) $\lim_{(x,y)\to(-2,1)} \frac{x^2+xy+y^2}{x^2-y^2}$.

(d) $\lim_{(x,y)\to(\pi,\pi)} x\sin\left(\frac{x+y}{4}\right)$.

(e) $\lim_{(x,y)\to(0,0)} \frac{x-y}{x^2+y^2}$.

(f) $\lim_{(x,y)\to(1,1)} \frac{x^2+xy-2y^2}{x^2-y^2}$.

9. Discuss the continuity of the given function.

(a) $g(x,y) = \frac{x+y}{1+x^2}$.

(b) $f(x,y) = \begin{cases} \frac{x^2y}{1+x} & if \quad x \neq -1 \\ y & if \quad x = -1 \end{cases}$

(c) $f(x,y) = \left[\begin{array}{lll} \frac{x^2y}{x^3+y^3} & if & (x,y) \neq (0,0) \\ 0 & if & (x,y) = (0,0) \end{array}\right.$

(d) $f(x,y) = \left[\begin{array}{lll} \frac{x^3+y^3}{x^2+y^2} & if & (x,y) \neq (0,0) \\ 0 & if & (x,y) = (0,0). \end{array}\right.$

3.4 Partial Derivatives

If f is a function of two variables x and y. Suppose we let x very while keeping y fixed, say $y = b$ (constant). Then f is a function of single variable. Denote $g(x) = f(x,y) = f(x,b)$. If g is differentiable at a then we call partial derivative of f with respect to x at (a,b) and denote it by $f_x(a,b)$. Thus $f_x(a,b) = g'(a)$ where $g(x) = f(x,b)$.

$$\begin{aligned} g'(a) &= \lim_{h\to 0}\frac{g(a+h)-g(a)}{h} = \lim_{x\to a}\frac{g(x)-g(a)}{x-a} \\ &= f_x(a,b) = \lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h} \\ &\qquad = \lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a}. \end{aligned}$$

Similarly, the partial derivative of f with respect to y at (a,b) denoted by $f_y(a,b)$ is obtained by keeping x fixed $(x = a)$ and

$\phi(y) = f(a, y)$ i.e.,

$$\phi'(b) = \lim_{h\to 0} \frac{\phi(b+h) - \phi(b)}{h} = \lim_{h\to 0} \frac{f(a, b+h) - f(a, b)}{h} = f_y(a, b).$$

Definition 3.4.1. If f is a function of two variables, its partial derivatives are the functions f_x and f_y defined by

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

(ii) $f_y(x, y) = \lim_{h\to 0} \frac{f(x,y+h)-f(x,y)}{h}$.

Notation 3.4.1. If $z = f(x, y)$, then

(i) $f_x(x, y) = f_x = \frac{\partial}{\partial x} f(x, y) = \frac{\partial z}{\partial x} = D_1 f = (D_x f)$

(ii) $f_y(x, y) = f_y = \frac{\partial}{\partial y} f(x, y) = \frac{\partial z}{\partial y} = D_2 f = (D_y f)$.

Remark 3.4.1. The above definition 3.4.1. is equivalent to (for a point (a, b))

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

(b) $f_y(a, b) = \lim_{y\to b} \frac{f(x,y+b)-f(a,b)}{y-b}$.

Example 3.4.1. If $f(x, y) = 4 - x^2 - 2y^2$, find $f_x(1, 1)$ and $f_y(1, 1)$ and interpret these numbers as slopes.

Solution. $f(x, y) = 4 - x^2 - 2y^2 \Rightarrow f_x(x, y) = -2x$ and $f_y(x, y) = -4y$. Hence $f_x(1, 1) = -2$ and $f_y(1, 1) = -4$.

The graph of $f(x, y) = 4 - x^2 - 2y^2$ and the vertical line $y = 1$ intersects it in the parabola $z = 2 - x^2, y = 1$. The slope of the tangent line to this parabola at $(1, 1, 1)$ in $f_x(1, 1) = -2$. Similarly, graph of f and $x = 1$ intersected at the parabola $z = 3-2y^2, x = 1$. The slope of the tangent line at $(1, 1, 1)$ is $f_y(1, 1) = -4$.

Theorem 3.4.1. Suppose f and g are functions of two variables. Also if $f_x(x, y), f_y(x, y), g_x(x, y)$ and $g_y(x, y)$ exist.

1. Sum and Difference rule

(i)

$$\begin{aligned}\frac{\partial}{\partial x}(f(x,y) \pm g(x,y)) &= \frac{\partial}{\partial x}f(x,y) \pm \frac{\partial}{\partial x}g(x,y) \\ &= f_x(x,y) \pm g_x(x,y) \\ &= f_x \pm g_x.\end{aligned}$$

(ii)

$$\begin{aligned}\frac{\partial}{\partial y}(f(x,y) \pm g(x,y)) &= \frac{\partial}{\partial y}f(x,y) \pm \frac{\partial}{\partial y}g(x,y) \\ &= f_y(x,y) \pm g_y(x,y) \\ &= f_y \pm g_y.\end{aligned}$$

2. Product rule

(i)

$$\begin{aligned}\frac{\partial}{\partial x}(f(x,y).g(x,y)) &= f_x(x,y)g(x,y) + f(x,y)g_x(x,y) \\ &= f_x g + f g_x.\end{aligned}$$

(ii)

$$\begin{aligned}\frac{\partial}{\partial y}(f(x,y).g(x,y)) &= f_y(x,y)g(x,y) + f(x,y)g_y(x,y) \\ &= f_y g + f g_y.\end{aligned}$$

3. Quotient Rule

(i)

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{f(x,y)}{g(x,y)}\right) &= \frac{f_x(x,y).g(x,y) - f(x,y)g_x(x,y)}{(g(x,y))^2} \\ &= \frac{f_x g - f g_x}{(g)^2}.\end{aligned}$$

(ii)

$$\begin{aligned}\frac{\partial}{\partial y}\left(\frac{f(x,y)}{g(x,y)}\right) &= \frac{f_y(x,y).g(x,y) - f(x,y)g_y(x,y)}{(g(x,y))^2} \\ &= \frac{f_y g - f g_y}{(g)^2}.\end{aligned}$$

Example 3.4.2. Let $f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2} & if \quad (x,y) \neq (0,0) \\ 0 & if \quad (x,y) = (0,0) \end{cases}$

then find f_x and f_y.

Solution 1. If $(x,y) = (0,0)$ then

(i)

$$\begin{aligned} f_x(0,0) &= \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} \\ &= \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0. \end{aligned}$$

(ii)

$$\begin{aligned} f_y(0,0) &= \lim_{h\to 0} \frac{f(0,0+h) - f(0,0)}{h} \\ &= \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = 0. \end{aligned}$$

Here $(0,0)$ is the point of discontinuity of $f(x,y)$ but $f_x(0,0)$ and $f_y(0,0)$ exist.

Solution 2. If $(x,y) \neq (0,0)$ then by using quotient rule

$$f_x(x,y) = \frac{2(y^2 - x^2y)}{(x^2+y^2)^2} \text{ and } f_y(x,y) = \frac{2(x^3 - y^2x)}{(x^2+y^2)^2}.$$

Remark 3.4.2. Partial derivatives can also be defined for functions of three or more variables, let $u = f(x,y,z)$ then

$$\begin{aligned} f_x(x,y,z) &= \lim_{h\to 0} \frac{(x+h,y,z) - f(x,y,z)}{h} \\ f_y(x,y,z) &= \lim_{h\to 0} \frac{f(x,y+h,z) - f(x,y,z)}{h} \\ f_z(x,y,z) &= \lim_{h\to 0} \frac{f(x,y,z+h) - f(x,y,z)}{h}. \end{aligned}$$

Exercises 3.4

1. Find the indicated partial derivatives.

(a) $z = \frac{x^3+y^3}{x^2+y^2}, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$.

(b) $z = x\sqrt{y} - \frac{y}{\sqrt{x}}, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$.

(c) $xy + yz = xz, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$.

(d) $xyz = \cos(x+y+z), \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$.

(e) $f(x,y,z) = xyz, f_y(0,1,2)$.

(f) $f(x,y,z) = \sqrt{x^2+y^2+z^2}$.

2. If $f(x,y) = \sin\left(\frac{y}{1+x}\right)$. Calculate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

3. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ if z is defined implicitly as a function of x and y by the equation $x^3 + y^3 + z^3 + 6xyz = 1$.

4. Find f_x, f_y and f_z if $f(x,y,z) = e^{xy} lnz$.

3.5 Higher Order Partial Derivatives

If f is a function of two variables, then its partial derivatives f_x and f_y are also functions of two variables, so we can consider their partial derivatives $(f_x)_x, (f_x)_y, (f_y)_x$ and $(f_y)_y$ which are called the second partial derivatives of f. If $z = f(x,y)$ then

$$\begin{aligned} (f_x)_x &= f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = D_1 D_1 f \\ (f_x)_y &= f_{xy} = D_2 D_1 f \\ (f_y)_y &= f_{yx} = D_1 D_2 f \\ (f_y)_y &= f_{yy} = D_2 D_2 f. \end{aligned}$$

Example 3.5.1. Let $f(x,y) = \begin{cases} \frac{x^3y-xy^3}{x^2+y^2} & if \quad (x,y) \neq (0,0) \\ 0 & if \quad (x,y) = (0,0) \end{cases}$.

Show that $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Solution. We have

$$\begin{aligned} f_x(x,y) &= \frac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2} \\ f_y(x,y) &= \frac{x^5-4x^3y^2-xy^4}{(x^2+y^2)^2} \\ f_x(0,0) &= f_y(0,0)=0 \\ f_x(0,y) &= -y, \quad y\neq 0 \\ f_y(x,0) &= x, \quad x\neq 0 \\ f_{xy}(0,0) &= \lim_{h\to 0}\frac{f_x(0,h)-f_x(0,0)}{h}=-1. \\ f_{yx}(0,0) &= \lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h}=1. \end{aligned}$$

Remark 3.5.1. Suppose f is defined on a disk D that contains the point (a,b). If the functions f_{xy} and f_{yx} are both continuous on D. Then $f_{xy}(a,b)=f_{yx}(a,b)$.

Remark 3.5.2. Partial derivatives of order 3 or higher can be defined as

$$(f_{x,y})_y = f_{xyy} = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial y\partial x}\right)=\frac{\partial^3 f}{\partial^2 y\partial x}$$

and using Remark 3.5.2. we have $f_{xyy}=f_{yxy}=f_{yyx}$ if these functions are continuous.

Example 3.5.2. Calculate f_{xxyz} if $f(x,y,z)=\sin(3x+yz)$.

Solution.

$$\begin{aligned} f(x,y,z) &= \sin(3x+yz) \\ f_x(x,y,z) &= 3\cos(3x+yz) \\ f_{xx}(x,y.z) &= -9\sin(3x+yz) \\ f_{xxy}(x,y,z) &= -9\cos(3x+yz).z \\ &= -9z\cos(3x+yz) \\ f_{xxyz}(x,y,z) &= 9yz\sin(3x+yz)-9\cos(3x+yz). \end{aligned}$$

Remark 3.5.2. Partial derivatives occur in partial differential equation that express certain physical equations. For instance, the p.d.e. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ is called Laplace's equation.

Remark 3.5.3. The wave equation $\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} = u_{tt} = a^2 u_{xx}$ describes the motion of a wave, which could be an ocean wave, a sound wave, a light wave, t =time, x =distance and a constant (depends on the density of the string and on the tension in the string).

Exercises 3.5

1. Find the second partial derivatives of $f(x,y) = x^3 + x^2y^3 - 2y^2$.

2. Show that the function $u(x,y) = e^x \sin y$ is a solution of Laplace's equation.

3. Verify that the function $u(x,t) = \sin(x - at)$ satisfies the wave equation.

4. Show that the following functions satisfies Laplace's equation

   (a) $z = x^4 - 6x^2y^2 + y^4$.

   (b) $z = \log(x^2 + y^2)$.

   (c) $u = \sin x \cos hy + \cos \sin hy$.

5. Show that each of the following function is solution of the wave equation.

   (a) $u = \sin kx \sin akt$.

   (b) $u = t/(a^2t^2 - xy)$.

   (c) $u = (x - at)^6 + (x + at)^6$.

3.6 Differentiability and Gradient

Definition 3.6.1. A function f is differentiable at $p_0 = (x_0, y_0)$ iff $\exists$ a first degree polynomial $g(x,y) = f(x_0,y_0) + (x - x_0)a + (y - y_0)b$ such that

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y) - g(x,y)}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = 0 \texttt{ where } \begin{array}{c} a = f_x(x_0,y_0) \\ b = f_y(x_0,y_0).* \end{array}$$

Remark 3.6.1. If $f(x)$ is a function of one variable with $F'(x_0) = a$, then

$$\begin{aligned} a = F'(x_0) &= \lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} \\ &\Rightarrow \lim_{x \to x_0} \left(\frac{F(x) - F(x_0)}{x - x_0} - a \right) = 0 \\ &\Rightarrow \lim_{x \to x_0} \left(\frac{F(x) - (F(x_0) + a(x - x_0))}{x - x_0} \right) = 0. \end{aligned}$$

Hence $F(x)$ is differentiable at x_0 if there exist a linear function $G(x) = F(x, 0) + a(x - x_0)$ for $a = F'(x_0) \ni \lim_{x \to x_0} \frac{F(x) - G(x)}{x - x_0} = 0$.

Theorem 3.6.1. Suppose that $g(P) = f(p_0) + (P - P_0).M, (p = (x, y), p_0 = (x_0, y_0), M = (a, b))$ is a polynomial satisfying $*$. Then

i) f is continuous at p_0.

ii) $D_1 f$ and $D_2 f$ exist at p_0.

iii) g has the slope i.e., $M = (f_x(p_0), f_y(p_0))$ is the slope.

Proof. (i). Let $f(p) = g(p) + \|p - p_0\| \frac{f(p) - g(p)}{\|p - p_0\|}$ then $\lim_{p \to p_0} f(p) = \lim_{p \to p_0} g(p) + \lim_{p \to p_0} \|p - p_0\| \frac{f(p) - g(p)}{\|p - p_0\|}$ (because f is differentiable at p_0)$\Rightarrow \lim_{p \to p_0} f(p) = \lim_{p \to p_0} g(p) = g(p_0) = f(p_0)$ .Thus $\lim_{p \to p_0} f(p) = f(p_0)$. Hence f is continuous at p_0.
(ii) and (iii) can be prove similarly.

Definition 3.6.2. The gradient of f denoted by ∇f at p_0, is the vector $\nabla f(p_0) = (D_1 f(p_0), D_2 f(p_0)) = (f_x(p_0), f_y(p_0))$. If both partial derivatives exist.

Example 3.6.1. Find ∇f if $f(x, y) = \sin x + e^{xy}$.

Solution.

$$\begin{aligned} f(x, y) &= \sin x + e^{xy} \\ \nabla f(x, y) &= (f_x(x, y), f_y(x, y)), f_x(x, y) = \cos x + y e^{xy} \\ &\quad \text{and } \; f_y(x, y) = x e^{xy}. \end{aligned}$$

Hence

$$\nabla f(x,y) = (\cos x + ye^{xy}, xe^{xy}) \text{ and } \nabla f(0,1) = (2,0).$$

Example 3.6.2. Let $f(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$ find $\nabla f(2\sqrt{2}, 2\sqrt{2}, -3)$.

Solution.

$$\begin{aligned} f(x,y,z) &= \frac{1}{\sqrt{x^2+y^2+z^2}}, f_x(z,y,z) = \frac{-x}{(x^2+y^2+z^2)^{3/2}} \\ f_y(x,y,z) &= \frac{-y}{(x^2+y^2+z^2)^{3/2}} f_z(x,y,z) = \frac{-z}{(x^2+y^2+z^2)^{3/2}}. \end{aligned}$$

Hence $\nabla f(x,y,z) = \frac{-xi-yj-zk}{(x^2+y^2+z^2)^{3/2}}$
$\Rightarrow \nabla f(2\sqrt{2}, 2\sqrt{2} - 3) = (-2\sqrt{2}i - 2\sqrt{2}j + 3k)\frac{1}{125}$.

The gradient vector plays a crucial role in the definition of the plane tangent to the graph of a function of several variables.

Definition 3.6.3. Let f be differentiable at a point (x_0, y_0, z_0) on a level surface S of f. If $\nabla f(x_0, y_0, z_0) \neq 0$, then the plane through (x_0, y_0, z_0) whose normal is $\nabla f(x_0, y_0, z_0)$ is the plane tangent to S at (x_0, y_0, z_0) and $\nabla f(x_0, y_0, z_0)$ is normal to S.

Remark 3.6.2. Equation of tangent plane through (x_0, y_0, z_0) with normal vector $\nabla f(x_0, y_0, z_0)$ is $(p - p_0).\nabla f(x_0, y_0, z_0) = 0$

$$\begin{aligned} \Longrightarrow (x - x_0)f_x(x_0, y_0, z_0) &+ (y - y_0)f_y(x_0, y_0, z_0) \\ &+ (z - z_0)f_z(x_0, y_0, z_0) = 0. \end{aligned}$$

Example 3.6.3. Find an equation of tangent plane to the sphere $x^2 + y^2 + z^2 = 4$ at the point $(-1, +1, \sqrt{2})$.

Solution. The sphere is the level curve $f(x,y,z) = 4$ where $f(x,y,z) = x^2+y^2+z^2 \Rightarrow f_x(x,y,z) = 2x, f_y(x,y,z) = 2y, f_z(x,y,z) = 2z \nabla f(x,y,z) = (2x, 2y, 2z)$ and $\nabla f(-1, 1, \sqrt{2}) = (-2, 2, 2\sqrt{2})$.

Hence equation of the plane tangent at $(-1, 1, \sqrt{2})$ is

$$-2(x+1) + 2(y-1) + 2\sqrt{2}(2 - \sqrt{2}) = 0 \Rightarrow -x + y + \sqrt{2}z = 0.$$

Remark 3.6.3. If f is a function of two variables that is differentiable at (x_0, y_0) then let $g(x, y, z) = f(x, y) - z = 0$. Hence the graph of f is the level surface $g(x, y, z) = 0$. We can find a plane tangent to the graph of f at $(x_0, y_0, f(x_0, y_0))$ and to be the plane tangent to the level curve $g(x, y, z) = 0$.

Hence $\nabla g(x_0, y_0, z_0) = f_x(x_0, y_0)i + f_y(x_0, y_0))j - k$.

$$\begin{aligned}
&f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0\\
&z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)\\
&z_0 = f(x_0, y_0).
\end{aligned}$$

Example 3.6.4. Find the tangent plane to the elliptic paraboloid $z = 2x^2 + y^2$ at the point $(1, 1, 3)$.

Solution. Let $f(x, y) = 2x^2 + y^2$. Then $f_x(x, y) = 4x$ and $f_y(x, y) = 2y \Rightarrow f_x(1, 1) = 4$ and $f_y(1, 1) = 2$.

Thus equation of the tangent plane at $(1, 1, 3)$ is

$$\begin{aligned}
z &= f(1, 1) + f_x(1, 1)(x - 1) + f_y(1, 1)(y - 1)\\
&= 3 + 4(x - 1) + 2(y - 1)\\
\Rightarrow z &= 3 + 4x - 4 + 2y - 2 \quad \Rightarrow 4x + 2y - z = 3.
\end{aligned}$$

The Chain Rule 1. Suppose that $z = f(x, y)$ is a differentiable function of $x\&y$, where $x = g(t)$ and $y = h(t)$ are both differentiable functions of t. Then z is a differentiable function of t and

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Exercise. If $z = x^2y + 3xy^4$, where $x = e^t, y = \sin t$, find $\frac{dz}{dt}$.

The Chain Rule 2. Suppose that $z = f(x, y)$ is a differentiable function of x and y, where $x = g(s, t), y = h(s, t)$ and the partial derivatives g_s, g_t, h_s and h_t exist. Then

$$\begin{aligned}
\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}\\
\frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.
\end{aligned}$$

Example 3.6.5. If $z = e^x \sin y$, where $x = st^2$, $\quad y = s^2t$, find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

Solution. Using the above formula

$$\begin{aligned}
\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\
&= e^x \sin y(t^2) + e^x \cos y(2st) \\
&= t^2 e^{st^2} \sin(s^2t) + e^{st^2} \cos(s^2t).2st \\
\frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} \\
&= e^x \sin y(2st) + e^y \cos y(s^2) \\
&= 2ste^{st^2} \sin(s^2t) + e^{st^2} \cos(s^2t)(s^2) \\
&= 2ste^{st^2} \sin(s^2t) + s^2 e^{st^2} \cos(s^2t).
\end{aligned}$$

Remark 3.6.4. In $z = f(x, y), x = g(s, t), y = h(s, t)$

- $s\&t$ are independent variables.
- $x\&y$ are intermediate variables.
- z is the dependent variable.

Remark 3.6.5. To remember the chain rule it is helpful to draw the tree diagram

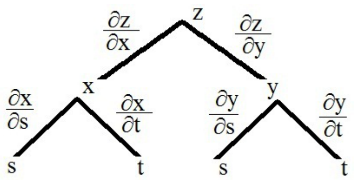

Fig. 3.6.1

$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$ similarly we can find $\frac{\partial z}{\partial t}$.

Example 3.6.6. Let $U = x \cos yz^2, x = \sin t, y = t^2$ and $z = e^t$. Find $\frac{\partial U}{dt}$.

Solution. By corresponding chain rule, we have

$$\frac{dU}{dt} = \frac{\partial U}{\partial x}\frac{dx}{dt} + \frac{\partial U}{\partial y}\frac{dy}{dt} + \frac{\partial U}{\partial z}\frac{dz}{dt}$$

$$= \cos(yz^2)\cos t - xz^2\sin(yz^2)(2t) - 2xyz\sin(yz^2)e^t$$
$$= \cos(t^2e^{2t})\cos t - 2te^t\sin t\sin(t^2e^{2t}) - 2t^2e^{2t}\sin t\sin(t^2e^{2t}).$$

Example 3.6.7. Let $U = \sqrt{x} + y^2z^2, x = 1 + s^2 + t^2, y = st$ and $z = 3s$. Find $\frac{\partial U}{\partial s}$ and $\frac{\partial U}{\partial t}$.

Solution. Consider the following chain rule

$$\begin{aligned}
\frac{\partial U}{\partial s} &= \frac{\partial U}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial U}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial U}{\partial z}\frac{\partial z}{\partial s} \\
&= \frac{s}{\sqrt{1+s^2+t^2}} + 135s^2t^2 \\
\frac{\partial U}{\partial t} &= \frac{\partial U}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial U}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial U}{\partial z}\frac{\partial z}{\partial t} \\
&= \frac{s}{\sqrt{1+s^2+t^2}} + 5 + s^5t.
\end{aligned}$$

Exercises 3.6

1. Find the equation of the tangent plane to f at the indicated point.
   (a) $f(x,y) = \sin x + e^{xy} \quad (x_0, y_0) = (0,1).$
   (b) $f(x,y) = x^2y^3 - 4y \quad (x_0, y_0) = (2,-1).$
2. Let $z = xlny, x = u^2 + \nu^2$ and $y = u^2 - \nu^2$. Find $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial \nu}$.
3. Write out the Chain rule for the case where $w = f(x,y,z,t)$ and $x = x(u,\nu), y = y(u,\nu), z = z(u,\nu)$ and $t = t(u,\nu)$.
4. If $U = x^4y + y^2z^3$, where $x = rse^t, y = rs^2e^{-t}$ and $z = t^2 s\sin t$. Find the value of $\frac{\partial u}{\partial s}$ when $r = 2, s = 1, t = 0$.
5. If $g(S,t) = f(S^2 - t^2, t^2 - S^2)$ and f is differentiable, show that g satisfies the equation $t\frac{\partial g}{\partial S} + S\frac{\partial g}{\partial t} = 0$.

   Hint. Let $x = S^2 - t^2$ and $y = t^2 - S^2$. Then $g(S.t) = f(x,y)$ and by the chain rule

$$\frac{\partial g}{\partial S} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial S} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial S} = 2Sf_x - 2Sf_y$$

$$\frac{\partial g}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} = -2tf_x - 2tf_y$$

$$\therefore \quad t\frac{\partial g}{\partial S} + S\frac{\partial g}{\partial t} = 0.$$

3.7 Directional Derivatives

Recall that if $z = f(x, y)$, then f_x and f_y are defined and represent the rates of change of z in the $x\&y$ directions i.e., in the directions of the unit vectors i and j.

Now we wish to find the rates of change of z at (x_0, y_0) in the direction of an arbitrary unit vector $u = (a, b)$. To find this we use the following definition.

Definition 3.7.1. Let f be a function of two variables defined on a disk D centered at (x_0, y_0) and let $u = (a, b)$ be a unit vector. Then the directional derivative of f at (x_0, y_0) in the direction of u, denoted $D_u f(x_0, y_0)$ is defined by

$$D_u f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if limit exists.

Remark 3.7.1. If $u = (1, 0) = i$ then $D_u f = D_i f = f_x$ and if $u = (0, 1)$ then $D_u f = D_j f = f_y$.

Theorem 3.7.1. If f is a differentiable function of x and y then f has a directional derivative in the direction of any unit vector $u = (a, b)$ and $D_u f(x, y) = f_x(x, y)a + f_y(x, y)b = \nabla f(x, y).u$.

Proof. Define $g(h) = f(x_0 + ha, y_0 + hb)$. This is single variable. Hence

$$\begin{aligned} g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h} &= \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h} \\ &= D_u f(x_0, y_0) \quad (*) \end{aligned}$$

on the other hand, we can write $g(h) = f(x, y)$ where $x = x_0 + ha, y = y_0 + hb$. So by chain rule

$$g'(h) = \frac{\partial f}{\partial x}\frac{dx}{du} + \frac{\partial f}{\partial y}\frac{dy}{du} = f_x.a + f_y.b.$$

Now if $h = 0$ then $g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b \quad (**)$
From $(*)$ and $(**)$, we have

$$\begin{aligned} D_u f(x_0, y_0) &= f_x(x_0, y_0)a + f_y(x_0, y_0)b \\ &= \nabla f(x_0, y_0).(a, b). \end{aligned}$$

Remark 3.7.2. If the unit vector $\cup$ makes an angle θ with the positive $x-$axis then we can write $\cup = (\cos\theta, \sin\theta)$ and $D_\cup f(x_0, y_0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta$.

Example 3.7.1. Let $f(x, y) = 6 - 3x^2 - y^2$ and let

$$\cup = \frac{2}{2}i + \frac{\sqrt{2}}{2}j. \texttt{ Find } D_\cup f(1, 2).$$

Solution. $f(x, y) = 6 - 3x^2 - y^2 \Rightarrow f_x(x, y) = -6x \;\; f_y(x, y) = -2y$

$$\begin{aligned} D_\cup f(1, 2) = \nabla f(1, 2), \cup &= (f_x(1, 2), f_y(1, 2)).\left(\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\right) \\ &= (-6, -4)\left(\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{2}\right) \\ &= -3\sqrt{2} + 2\sqrt{2} = -\sqrt{2}. \end{aligned}$$

Example 3.7.2. Find the directional derivative $D_\cup f(x, y)$ if $f(x, y) = x^3 - 3xy + 4y^2$ and $\cup$ is the unit vector given by the angle $\theta = \pi/6$. What is $D_\cup f(1, 2)$?

Solution. By using the above remark

$$\begin{aligned} D_\cup f(x, y) &= f_x(x, y)\cos\theta + f_y(x, y)\sin\theta \\ &= (3x^2 - 3y)\cos\pi/6 + (-3x + 8y)\sin\pi/6 \\ &= (3x^2 - 3y)\frac{\sqrt{3}}{2} + (-3x + 8y)\frac{1}{2} \\ &= \frac{1}{2}\left\{\left[3\sqrt{3}x^2 - 3x\right] + (8 - 3\sqrt{3})y\right\}. \end{aligned}$$

Hence $D_\cup f(1, 2) = \frac{1}{2}[3\sqrt{3} - 3 + (8 - 3\sqrt{3}), 2] = \frac{13 - 3\sqrt{3}}{12}$.

Example 3.7.3. Find the directional derivative of the function $f(x, y) = x^2y^3 - 1$ at the point $(2, -1)$ in the direction of the vector $V = (2, 5)$.

Solution. $f(x, y) = x^2y^3 - 4y \;\; (x_0, y_0) = (2, -1)$ and $V = (2, 5) \Rightarrow D_{\cup}f(x, y) = \nabla f(x, y).\cup$, where $\cup$ is the unit vector in the direction of V.

$$= (f_x(x, y), f_y(x, y)).\cup \;\; \texttt{where} \;\; \cup = \frac{V}{\|V\|} = \frac{1}{\sqrt{29}}(2, 5)$$

$$= (2xy^3, 3x^2y^2 - 4)\left(\frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}}\right)$$

$$\Rightarrow \;\; D_{\cup}f(2, -1) = (-4, 8).\left(\frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}}\right) = \frac{-8 + 40}{\sqrt{29}} = \frac{32}{\sqrt{29}}$$

$$\Rightarrow \;\; D_{\cup}f(2, -1) = \frac{32}{\sqrt{29}}.$$

Remark 3.7.3. The directional derivative in the direction of an arbitrary nonzero vector V is defined to be $D_{\cup}f(x_0, y_0)$ where $\cup = \frac{V}{\|V\|}$.

Definition 3.7.2. The directional derivative of f at (x_0, y_0, z_0) in the direction of a unit vector $\cup = (a, b, c)$ is

$$D_{\cup}f(x_0, y_0, z_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb, z_0 + hc) - f(x_0, y_0, z_0)}{h}$$

if this limit exists.

Remark 3.7.4. If $f(x, y, z)$ is differentiable and $\cup = (a, b, c)$ then the same method that was used to prove the above theorem can be used to show that

$$\begin{aligned} D_{\cup}f(x, y, z) &= f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z).C \\ &= (f_x, f_y, f_z).\cup = \nabla f(x, y, z) \cup . \end{aligned}$$

Example 3.7.4. If $f(x, y, z) = x \sin yz$ then find the directional derivative of f at $(1, 3, 0)$ in the direction of $V = (1, 2, -1)$.

Solution.

$$\begin{aligned} f(x, y, z) &= \nabla f(x, y, z) = (f_x(x, y, z).f_y(x, y, z).f_z(x, y, z)) \\ &= (\sin yz, xz \cos yz, xy \cos yz). \end{aligned}$$

At $(1,3,0)$ we have $A\nabla f(1,3,0) = (0,0,3)$ and the unit vector $\cup$ in the direction of V is

$$\cup = \frac{1}{\sqrt{6}}(1,2,-1).$$

Thus

$$\begin{aligned} D_{\cup}f(1,3,0) = \nabla f(1,3,0).\cup &= (0,0,3).\left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right) \\ &= -\frac{3}{\sqrt{6}} = -\sqrt{\frac{3}{2}}. \end{aligned}$$

Question 3.7.1. In which directions does f change fastest and what is the maximum rate of change ? The answers are given in the following theorem.

Theorem 3.7.2. Suppose f is differentiable function of two or three variables. The maximum value of the directional derivative $D_{\cup}f(x)$ is $\|\nabla f(x)\|$ and it occurs when $\cup$ has the same direction as the gradient vector $\nabla f(x)$.

Proof. $D_{\cup}f = \nabla f.\cup = \|\nabla f\|\| \cup \| \cos\theta = \|\nabla f\| \cos\theta, \theta$ is the angle between ∇f and $\cup$.

The maximum value of $\cos\theta = 1$ if $\theta = 0$.

Maximum value of $D_{\cup}f = \|\nabla f\|$ at $\theta = 0$.

Example 3.7.5. If $f(x,y) = xe^y$

a) Find the rate of change of f at the point $P(2,0)$ in the direction from P to $Q(\frac{1}{2}, 2)$.

b) In what direction does f have the maximum rate of change ? what is the maximum rate of change ?

Solution.

a) $D \cup f(2,0) = 1$ for $\cup = \frac{\overrightarrow{PQ}}{\|\overrightarrow{PQ}\|} = \left(\frac{-3}{5}, \frac{4}{5}\right)$.

b) f increases fastest in the direction of the gradient vector $\nabla f(2,0) = (1,2)$. The maximum rate of change is $\|\nabla f(2,0)\| = \|(1,2)\| = \sqrt{5}$.

Example 3.7.6. Suppose that the temperature at a point (x, y, z) in space is given by $T(x, y, z) = \frac{80}{1+x^2+2y^2+3z^2}$, where t is measured in C^0 and x, y, z in vectors. In which directions does the temperature increase fastest at the point $(1, 1, -2)$? What is the maximum rate of increase ?

Solution.

$$\begin{aligned}\nabla T &= \frac{\partial T}{\partial x}i + \frac{\partial T}{\partial y}j + \frac{\partial T}{\partial z}k \\ &= \frac{-160x}{(+x^2+2y^2+3z^2)^2}i + \frac{-320y}{(1+x^2+2y^2+3z^2)^2}j \\ &+ \frac{-480z}{(1+x^2+2y^2+3z^2)^2}k \\ &= \frac{160}{(1+x^2+2y^2+3z^2)^2}(-x, -2y, -3z).\end{aligned}$$

At $(1, 1, -2), \nabla T(1, 1, -2) = \frac{160}{256}(-1, -2, 6) = \frac{5}{8}(-1, -2, 6)$

Temperature increases fastest in the direction of the gradient vector. The maximum rate of increase in the length ∪ the gradient. i.e., $\|\nabla T(1, 1, -2)\| = \frac{5}{8}\|(-1, -2, 6)\| = \frac{5}{8}\sqrt{1+4+36} = \frac{5\sqrt{41}}{8}$.

Thus the maximum rate of increase of temperature is

$$\frac{5\sqrt{41}}{8} \simeq 4^o c/m.$$

Exercises 3.7

1. Define the function $f : R^2 \to R$ by

$$f(x, y) = \begin{cases} (x/|y|)\sqrt{x^2+y^2} & if \quad y \neq 0 \\ 0 & if \quad y = 0. \end{cases}$$

(i) Prove that the function $f : R^2 \to R$ is not continuous at the point $(0, 0)$.

(ii) Prove that the function $f : R^2 \to R$ has directional derivatives in all directions at the point $(0, 0)$.

(iii) Prove that if c is any number, then there is a vector p of the norm 1 such that

$$\frac{\partial f}{\partial p}(0,0) = c.$$

2. Consider the following assertions for a function $f : R^2 \to R$:

(i) The function $f : R^2 \to R$ is continuously differentiable.

(ii) The function $f : R^2 \to R$ has directional derivatives in all directions at each point in R^2.

(iii) The function $f : R^2 \to R$ has first order partial derivatives at each point in R^2. Explain the implications between these assertions.

3.8 Tangent Plane Approximation

Recall that an equation of the tangent plane to the surface $z = f(x, y)$ at the point $P(x_0, y_0, z_0)$ is

$$\begin{aligned} z - z_0 &= f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ \Rightarrow z &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \end{aligned}$$

To discuss tangent plane approximation, let us define differentiation.

For a function of one variable $y = f(x)$, we defined the increment of y as $\triangle y = f(x+\triangle x) - f(x)$ and differential of y as $dy = f'(x)dx$

$\triangle y = f(a + \triangle x) - f(a) =$ change in the height of the curve $y = f(x)$.

$dy =$ the change in the height of tangent line.

Now let, $t = \frac{\triangle y - dy}{\triangle x} = \frac{f(x+\triangle x)-f(x)-f'(x)dx}{\triangle x} = \frac{f(x+\triangle x)-f(x)}{\triangle x} - f'(x) \Rightarrow f'(x) - f'(x) \to 0$ as $\triangle x \to 0$.

Now for a function of two variables, $z = f(x, y)$

$$\triangle z = f(x + \triangle x, y + \triangle y) - f(x, y).$$

Then the differential dz, also called the total differential, is defined by

$$dz = f_x(x,y)dx + f_y(x,y)dy = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy.$$

If we take $dx = \triangle x = x - a, dy = \triangle y = y - b$ then

$$dz = f_x(a,b)(x-a) + f_y(a,b)(y-b) \qquad (*).$$

But from equation of tangent plane at $(a, b, f(a,b))$ is

$$z - f(a,b) = f_x(a,b)(x-a) + f_y(a,b)(y-b) \qquad (**).$$

From $(*)$ and $(**)$, we have dz = the change in height of the tangent plane, where $\triangle z$ = the change in height of the surface $y = f(x,y)$ when (x,y) changes from (a,b) to $(a + \triangle x, b + \triangle y)$.

Definition 3.8.1. A function f of two variables is differentiable at (a,b) if there exist a disk D centered at (a,b) and functions ε_1 and ε_2 of two variables s.t.

$$\begin{aligned} f(x,y) - f(a,b) &= f_x(a,b)(x-a) + f_y(a,b)(y-b) + \varepsilon_1(x,y) \\ &\quad (x-a) \\ &= +\varepsilon_2(x,y)(y-b) \texttt{ for } (x,y) \in D, \end{aligned}$$

where $\lim_{(x,y)\to(a,b)} \varepsilon_1(x,y) = 0$ and $\lim_{(x,y)\to(a,b)} \varepsilon_2(x,y) = 0$.

Define $\triangle z - dz = \varepsilon_1 \triangle x + \varepsilon_2 \triangle y$ where $\varepsilon_1 \& \varepsilon_2$ are functions of $\triangle x$ and $\triangle y$ that approach 0 as $\triangle x \& \triangle y$ approach 0.

$$\Rightarrow \triangle z - dz \simeq 0 \texttt{ and so } \triangle z \simeq dz.$$

Hence change in z is approximately equal to the differential dz when $\triangle x$ and $\triangle y$ are small.

$$\begin{aligned} \Rightarrow \quad & \triangle z = f(a + \triangle x, b + \triangle y) - f(a,b) \simeq dz \\ \Rightarrow \quad & f(a + \triangle x, b + \triangle y) \simeq f(a,b) + dz = f(a,b) + \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \\ & = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b). \end{aligned}$$

Therefore in the approximation, we used the tangent plane at $(a.b, f(a,b))$ as an approximation to the surface $z = f(x,y)$ when (x,y) is close to (a,b).

Example 3.8.1. Let $z = f(x, y) = x^2 + 3xy - y^2$

a) find the differential dz.

b) If x changes from 2 to 2.05 and y changes from 3 to 2.9 compare the values of $\triangle z$ and dz.

Solution.

a) $f(x, y) = x^2 + 3xy - y^2$ then

$$\begin{aligned} dz &= \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy = (2x + 3y)dx + (3x - 2y)dy \\ \Rightarrow \quad dz &= (2x + 3y)dx + (3x - 2y)dy. \end{aligned}$$

b) Putting $x = 2, \triangle x = dx = .5, y = 3$ and $\triangle y = dy = -0.04$, we get

$$dz = [2(2) = 3(3)](0.5) + [3(2) - 2(3)](-0.04) = 0.65.$$

The increment of z is

$$\begin{aligned} \triangle z &= f(2.05, 2.96) - f(2, 3) \\ &= [(2.05)^2 + 3(2.05)(2.96) - (2.96)^2] \\ &\quad -[2^2 + 3(2)(3) - 3^2] \\ &= 0.6449. \end{aligned}$$

observe that $\triangle z = dz$ but dz is easier to compute.

Example 3.8.2. Use differentials to find an approximate value for

$$\sqrt{9(1.95)^2 + (8.1)^2}.$$

Solution. Consider the function $z = f(x, y) = \sqrt{9x^2 + y^2}$ and observe that we can easily calculate $f(2.8) = 10$ and $f(1.95, 8.1) = \sqrt{9(1.95)^2 + (8.1)^2}$.

So take $a = 2, b = 8, dx = \delta x = -0.05$ and $dy = \triangle y = 0.1$.

$$f_x(x, y) = \frac{9x}{\sqrt{9x^2 + y^2}} f_y(x, y) = \frac{y}{\sqrt{9x^2 + y^2}}$$

we have

$$\begin{aligned} f(1.95, 8.1) &= \sqrt{9(1.95)^2 + (8.1)^2} \simeq f(2,8) + dz \\ &= f(2,9) + f_x(2,8)dx + f_y(2,9)dy \\ &= 10 + \frac{18}{10}(-0.05) + \frac{8}{10}(0.1) = 9.99. \end{aligned}$$

Exercises 3.8

1. Find the differential of the function

   (a) $z = x^2y^3$.

   (b) $u = e^x \cos xy$.

   (c) $z = ye^{xy}$.

2. If $z = 5x^2 + y^2$ and (x, y) changes from $(1, 2)$ to $(1.05, 2.1)$ compare the values of $\triangle z$ and dz.

3.9 Maxima and Minima (Extreme Values)

Partial derivatives are important to determine maximum and minimum values of functions of two variables.

Definition. A function of two variables has a local maximum at (a, b) if $f(x, y) \leq f(a, b)$ for all points (x, y) in some disk with center (a, b). The number $f(a, b)$ is called a local maximum value. If $f(x, y) \geq f(a, b)$ for all (x, y) in such a disk, $f(a, b)$ is local minimum value.

Remark 3.9.1. In the above definition, if the inequality hold for all points (x, y) in the domain of f, then f has an absolute maximum (or absolute minimum at (a, b)).

Theorem 3.9.1. If f has a local extremum (i.e., local maximum or minimum) at (a, b) and the first order partial derivatives of f exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Proof. For any constant H, define the function

$$F(t) = f(A + tH) \ \texttt{where} \ A = (a, b)$$

$\Rightarrow F'(t) = H.\nabla f(A+tH)$ and F has a local extreme at $t = 0$ hence $0 = F'(0) = H\nabla f(A)$, since H is arbitrary vector,
put $H = \nabla f(A) \Rightarrow \nabla f(A).\nabla f(A) = 0 \Rightarrow |\nabla f(A)|^2 = 0$

$$\Rightarrow \nabla f(A) = 0 \Rightarrow (f_x(A), f_y(A)) = 0 \Rightarrow f_x(a,b) = 0$$

and $f_y(a,b) = 0$ for $A = (a,b)$.

Remark 3.9.2. A point (a,b) such that $f_x(a,b) = 0$ and $f_y(a,b) = 0$ or one of these partial derivatives does not exist is called a critical point of f. If f has local extremum at (a,b) then (a,b) is a critical point of f. But not all critical points give rise to extreme.

Example 3.9.1. Let $f(x,y) = x^2 + y^2 - 2x - 5y + 14$, then

$$f_x(x,y) = 2x - 2, f_y(x,y) = 2y - 6.$$

Hence

$$f_x(x,y) = 0 \text{ and } f_y(x,y) = 0 \Rightarrow \quad \begin{array}{l} 2x - 2 = 0 \Rightarrow x = 1 \\ 2y - 6 = 0 \Rightarrow y = 3. \end{array}$$

$\Rightarrow (1,3)$ is the only critical point.
And

$$\begin{aligned} f(x,y) &= x^2 - 2x + 1 + y^2 - 6y + 9 + 4 \\ &= (x-1)^2 + (y-3)^2 + 4. \end{aligned}$$

Since $(x-1)^2 \geq 0$ and $(y-3)^2 \geq 0$, we have $f(x,y) \geq 4 \forall x, y$.

$\therefore f(1,3) = 4$ is a local minimum.
[In fact it is the absolute minimum] Range $[4, \infty)$.

Example 3.9.2. Find the extreme values of $f(x,y) = y^2 - x^2$.

Solution. Since $f_x = -2x, f_y = 2y$, the only critical point is $(0,0)$. On the $x-$axis, we have $y = 0$ so $f(x,y) = -x^2 < 0$(if $x \neq 0$) on the $y-$axis, we have $x = 0$. So $f(x,y) = y^2 > 0$ (if $y \neq 0$). Thus every disk center $(0,0)$ contains points where f takes positive values as well as points where f takes negative values.

$\therefore f(0,0) = 0$ can not be an extreme value. So f has no extreme value. Such point $(0,0)$ is called a Saddle point of f. Range $(-\infty, \infty)$.

The second derivatives of f usually helps to distinguish between maximum, minimum and saddle points.

Theorem 3.9.2. (Second Derivative Test). Suppose the second partial derivatives of f are continuous in a disk with center (a, b), and suppose that $f_x(a, b) = 0$ and $f_y(a, b) = 0$ (i.e., $A = (a, b)$ is a critical point of f). Let

$$D = D(a, b) = f_{xx}(a, b).f_{yy}(a, b) - (f_{xy}(a, b))^2$$

a) If $D > 0$ and $f_{xx}(a, b) > 0$, then $f(a, b)$ is a local minimum.

b) If $D > 0$ and $f_{xx}(a, b) < 0$ then $f(a, b)$ is a local maximum.

c) If $D < 0$, then $f(a, b)$ is not a local extremum value on (a, b) is called a saddle point of f.

Remark 3.9.3. If $D = 0$ then the test gives no information: f could have a local maximum or local minimum at (a, b) or (a, b) could be a saddle point of f.

Remark 3.9.4. To remember the formula for D write it as a determinant

$$D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - (f_{xy})^2.$$

Example 3.9.3. Find the local extreme of $f(x, y) = x^4 + y^4 - 4xy + 1$.

Solution. To find the critical points,

$$\begin{aligned} f_x &= 4x^3 - 4y = 0 \Rightarrow x^3 - y = 0 \Rightarrow y = x^3 \\ f_y &= 4y^3 - 4x = 0 \Rightarrow y^3 - x = 0 \Rightarrow y^3 = x \Rightarrow (x^3)^3 - x = 0 \\ &\quad \rightarrow x(x^8 - 1) = 0 \\ &\Rightarrow x(x^4 - 1)(x^4 + 1) = 0 \\ &\Rightarrow x(x^2 - 1)(x^3 - 1)(x^4 + 1) = 0 \\ &\Rightarrow x(x - 1)(x + 1)(x^2 + 1)(x^4 + 1) = 0. \end{aligned}$$

The real roots are $x = 0, 1, -1$.
Hence critical points $(0, 0)(-1, -1)(1, 1)$.

The second partial derivatives are

$$f_{xx} = 12x^2 \quad f_{xy} = -4, \quad f_{yy} = 12y^2$$
$$D(x, y) = f_{xx}f_{yy} - (f_{xy})^2 = 144x^2y^2 - 16.$$

i) Since $D(0, 0) = -16 < 0$ by Second Derivative Test $(0, 0)$ is a saddle point.

ii) Since $D(-1, -1) = 144 - 16 = 128 > 0$ and $f_{xx}(-1, -1) = 12 > 0$, so $f(-1, -1) = -1$ is a local minimum.

iii) Since $D(1, 1) = 144 - 16 = 128 > 0$ and $f_{xx}(1, 1) = 12 > 0$ by the above theorem $f(1, 1) = -1$ is also a local minimum.

Example 3.9.4. Find the shortest distance from the point $(1, 0, -2)$ to the plane $x + 2y + z = 4$.

Solution. Let (x, y, z) be on the plane and the distance d to the point $(1, 0, -2)$ is $d = \sqrt{(x-1)^2 + y^2 + (z+2)^2}$.

But (x, y, z) lies on the plane $x + 2y + z = 4 \Rightarrow z = 4 - x - 2y$ and so $d = \sqrt{(x-1)^2 + y^2 + (6-x-2y)^2}$

$$\Rightarrow \quad d^2 = (x-1)^2 + y^2 + (6-x-2y)^2 = f(x, y)$$
$$\Rightarrow \quad f_x = 2(x-1) + 2(6-x-2y) = 4x + 4y - 14 = 0$$
$$f_y = 2y - 4(6-x-2y) = 4x + 10y - 24 = 0$$

then the only critical point is $(11/6, 5/3)$ and

$$f_{xx} = 4, f_{xy} = 4, f_{yy} = 10, f_{yx} = 4.$$

Hence $D(x, y) = f_{xx}f_{yy} - (f_{xy})^2 = 40 - 16 = 24 > 0$ and $f_{xx} > 0 \Rightarrow$ by Second Derivative Test f has a local minimum at $(11/6, 5/3)$. Hence

$$\begin{aligned} d &= \sqrt{(x-1)^2 + y^2 + (6-x-2y)^2} \\ &= \sqrt{(5/16)^2 + (5/3)^2 + (5/6)^2} = \frac{5\sqrt{6}}{6} \end{aligned}$$

$\therefore$ The shortest distance from $(1, 0, -2)$ to the plane $x + 2y + z = 4$ is $\frac{5\sqrt{6}}{6}$.

Example 3.9.5. A rectangular box without cover is to be made from $12m^2$ of card board. Find the maximum volume of such a box.

Solution. Consider the volume $v = xyz$.

Area of 4 side faces and bottom of the box is

$$\begin{aligned} & 2xz + 2yz + xy = 12 \\ & \Rightarrow = \frac{12 - xy}{2x + 2y} = \frac{12 - xy}{2(x+y)} \\ & \Rightarrow V = xy\left(\frac{12 - xy}{2(x+y)}\right) = \frac{12xy - x^2y^2}{2(x+y)}, \end{aligned}$$

$$\begin{aligned} V_x &= y^2\frac{(12 - 2xy - x^2)}{2(x+y)^2} \\ V_y &= x^2\frac{(12 - 2xy - y^2)}{2(x+y)^2}. \end{aligned}$$

V is minimum $\Rightarrow V_x = 0$ and $V_y = 0$ and $x = 0$ or $y = 0 \Rightarrow V = 0$

$$\begin{aligned} \Rightarrow \quad & 12 - 2xy - x^2 = 0, 12 - 2xy - y^2 = 0 \Rightarrow x^2 = y^2 \text{ and} \\ & x = y.(x, y > 0) \\ & \Rightarrow 12 - 2x^2 - x^2 = 0 \Rightarrow 12 - 3x^2 = 0 \Rightarrow x = 2 \Rightarrow y = 2 \end{aligned}$$

and $z = 1$. By S.D.T. we can show that V is maximum at $x = 2$ maximum volume $V = 2.2.1 = 4m^3$.

Example 3.9.6. Find the local extreme value of $f(x, y) = x^3 + y^3 - 3xy$.

Solution. $(0, 0)$ saddle point, $f(1, 1)$ local minimum $= -1$.

Exercises 3.9

1. Find the maximum of $\{x + y + z : |x| + |y| + |z| \leq 1\}$ by inspection.
2. Find the maximum of $\{x^2 + y^2 + z^2 : 2x^2 + y^2 + 3z^2 \leq 1\}$.
3. For numbers a, b and c, find the minimum of

$$\{ax + by + cz : x^2 + y^2 + z^2 \leq 1\}.$$

4. Find the point on the plane $ax+by+cz+d=0$ that is closest to the point $(0,0,0)$.

5. For positive numbers a, b and c find a point on the ellipsoid

$$S = \left\{(x,y,z) \texttt{ in } R^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\right\}$$

that is closest to the point $(0,0,0)$.

6. Show that $x+y+z \geq 3$ for all (x,y,z) in R^3 such that $x > 0, y > 0, z > 0$, and $xyz = 1$.

7. Use Exercise 6 to verify the following Geometric Mean/ Arithmetic Mean Inequality : If a_1, a_2 and a_3 are positive numbers, then

$$(a_1 a_2 a_3)^{1/3} \leq \frac{a_1 + a_2 + a_3}{3}.$$

Generalize this inequality from $n = 3$ to general positive integer n.

3.10 Absolute maximum and Minimum Values

Recall for a function f of one variable, if f is continuous on a closed interval $[a,b]$, then f has an absolute minimum value and an absolute maximum value. There is a similar situation for function of two variables just as a closed interval contains its end points. A closed set in R^2 is one that contains all its boundary points.

Example 3.10.1. The disk $D = \{(x,y)/x^2+y^2 \leq 1\}$.

Theorem 3.10.1. If f is continuous on a closed, bounded set D in R^2, then f attains an absolute maximum value $f(x_1,y_1)$ and an absolute minimum value $f(x_2,y_2)$ at some points (x_1,y_1) and (x_2,y_2) in D.

To find the absolute maximum and minimum value of a continuous function f on a closed, bounded set D, we use the following steps:

1. Find the values of f at the critical points of f in D.

2. Find the extreme values of f on the boundary of D.

3. The largest of the values from step 1 and step 2 is the absolute maximum value; the smallest of these values is the absolute minimum value.

Example 3.10.2. Find the absolute maximum and minimum value of the function $f(x, y) = x^2 - 2xy + 2y$ on the rectangle

$$D = \{(x, y)/0 \leq x < 3, 0 \leq y \leq 2\}.$$

Solution. Since f is a polynomial, it is continuous on the closed bounded rectangle. By the above theorem both absolute maximum and absolute minimum values exist. Hence

i) By step 1 we can determine the critical points; thus

$f_x = 2x - 2y = 0 \Rightarrow (x, y) = (1, 1)$ is the only critical point
$f_y = -2x + 2 = 0$
and $f(1, 1) = 1$.

ii) By step 2 we took the values of f on the boundary of D i.e., along $L_1, L_2, L_3 \& L_4$.

$=$ on $L_1, y = 0$ and $f(x, 0) = x^2 \quad 0 \leq x \leq 3f$ is $\uparrow$.

So minimum value is $f(0, 0) = 0$ and maximum value is $f(3, 0) = 9$.

- On $L_2, x = 3$ and $f(3, y) = 9 - 6y + 2y = 9 - 4y \quad 0 \leq y \leq 2$. f is decreasing function of y so its maximum value is $f(3, 0) = 9$ and its minimum value is $f(3, 2) = 1$.

- On $L_3, y = 2$ and $f(x, 2) = x^2 - 4x + 4 \quad 0 \leq x \leq 3$, but $f(x, 2) = (x - 2)^2$ for $0 \leq x \leq 3$.

minimum value is $f(3,2) = 0$ and maximum value is $f(0,2) = 4$.

- Finally on L_4, we have $x = 0$ and $f(0,y) = 2y$ for $0 \leq y \leq$ maximum value $f(0,2) = 4$ and minimum value $f(0,0) = 0$. Thus on the boundary, the minimum value of f is 0 and the maximum value is 9.

iii) By step 3 comparing these values with $f(1,1) = 1$ (at the critical point). We conclude that the absolute maximum value of f and D is $f(3,0) = 9$ and the absolute minimum value is $f(0,0) = f(3,2) = 0$.

Example 3.10.3. Find the absolute maximum and minimum value of

$$f(x,y) = 2 + 2x + 2y - x^2 - y^2$$

on the triangular plate in the first quadrant bounded by the line $x = 0, y = 0$ and $y = 9 - x$.

Solution. $f(x,y) = 2+2x+2y-x^2-y^2$ and consider the following figure

Since f is continuous and differentiable on R^2, by the above theorem f is continuous and differentiable on D and it has absolute maximum and absolute minimum on D.

i) To determine critical point, we set

$$\begin{cases} f_x = 0 \\ f_y = 0 \end{cases} = 1 \begin{cases} 2 - 2x = 0 \\ 2 - 2y = 0 \end{cases} \Rightarrow x = 1, y = 1 \quad (1,1)$$

is critical point . and $f(1,1) = 4$.

ii) We determine the extreme values on the boundary of D.

- On $\overline{OA}, y = 0$ then $f(x,y) = f(x,0) = 2 + 2x - x^2$ where $0 \leq x \leq$. Its extreme values may occur at the end points:

$$x = 0 \Rightarrow f(0,0) = 2$$
$$x = 9 \Rightarrow f(0,0) = 20 - 81 = -61$$

and at the interior points where $f'(x,0) = 2-2x = 0 \Rightarrow x = 1$

$$\Rightarrow f(x,0) = f(1,0) = 3.$$

- On $\overline{OB}, x = 0$ and $f(x, y) = f(0, y) = 2 + 2y - y^2$ where $0 \leq y \leq$. By symmetry of x and y in the above calculation, the candidates on this segment are

$$f(0,0) = 2, f(0,9) = -61, f(0,1) = 3.$$

- On $\overline{AB}, y = 9 - x$ and we have

$$\begin{aligned} f(x, 9-x) &= f(x, y) \\ &= 2 + 2x + 2(9-x) - x^2 - (9-x)^2 \\ &= -61 + 18x - 2x^2 \\ &\Rightarrow f^{'}(x, 9-x) \\ &= 18 - 4x = 0 \Rightarrow x = \frac{18}{4} = \frac{9}{2} \Rightarrow y \\ &= 9 - \frac{9}{2} = \frac{9}{2} \end{aligned}$$

and $f\left(\frac{9}{2}, \frac{9}{2}\right) = \frac{-41}{2}$.
$\therefore$ candidates are $4, 2, -61, 3, \frac{-41}{2}$.
Hence $f(1,1) = 4$ is absolute maximum and $f(9,0) = f(0,9) = -61$ is absolute minimum value.

Exercises 3.10

1. Find the absolute maximum and minimum values on the set D where

   (a) $f(x, y) = 5 - 3x + 4y$. D is the closed triangular region with vertices $(0,0), (4,0)$ and $(4,5)$.

   (b) $f(x, y) = 1 + xy - x - y$. D is the region bounded by the parabola $y = x^2$ and the line $y = 4$.

3.11 Lagrange Multipliers

Let f be a function of two variables, to maximize or minimize $f(x, y)$ subject to $g(x, y) = k$ is to find the largest value of c such that the level curve $f(x, y) = c$ intersects $g(x, y) = k$. This happens when these curves just touch each other. i.e. when they have a common tangent value.

This means that the normal lines at the point (x_0, y_0) where their touch are identical. Therefore, the gradient vectors are parallel i.e. for some scaler λ. Where λ is called Lagrange multiplier and the procedure bases on equation

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$$

is called the method of Lagrange multipliers and is as follows.

To find the maximum and minimum values of f subject to the constraint $g(x, y, z) = k$ (assuming that the extreme values exist).

a) Find all values of x, y, z and λ such that

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$$

and $g(x, y, z) = k$.

b) Evaluate f at all the points (x, y, z) that arise from step (a). The largest of these values is the maximum value of f, the smallest is the minimum value of f.

Remark 3.11.1. For a function of three variables we have

$$\nabla f = \lambda \nabla g \quad \Rightarrow \quad \begin{aligned} f_x &= \lambda g_x \\ f_y &= \lambda g_y \\ f_z &= \lambda g_z \end{aligned}$$

and $g(x, y, z) = k$. This is a system of four equations in the four unknowns x, y, z and λ.

For a function of two variables we have

$$\begin{aligned} &\nabla f(x, y) = \lambda \nabla g(x, y) \ \texttt{and} \ g(x, y) = k \\ \Rightarrow \quad &f_x = \lambda g_x, f_y = \lambda g_y \ \texttt{and} \ g(x, y) = k. \end{aligned}$$

Examples 3.11.1. A rectangular box without a lid is to be from $12m^2$ of cardboard. Find the maximum volume of such a box.

Solution. Let x, y and z be the length, width and height respectively of the box in meters. Then we wish to maximize

$$v = f(x, y, z) = xyz$$

subject to the constraint $g(x, y, z) = 2xz + 2yz + xy = 12$.

By using the method of Lagrange multipliers we have

$$\nabla v = \lambda \nabla g \ \texttt{and} \ g(x, y, z) = 12.$$

$\Rightarrow v_x = \lambda g_x, v_y = \lambda gy, v_z \lambda g_z, \quad 2xz + 2yz + xy = 12$

$$\begin{aligned} \Rightarrow \quad & y_z = \lambda(2z + y) \ldots\ldots\ldots (1) \\ & x_z = \lambda(2z + x) \ldots\ldots\ldots (2) \\ & x_y = \lambda(2x + 2y) \ldots\ldots\ldots (3) \\ & 2xz + 2yz + xy = 12 \ldots\ldots\ldots (4). \end{aligned}$$

If we multiply (1) by x, (2) by y and (3) by z we have

$$\begin{aligned} xyz &= \lambda(2xz + xy) \ldots\ldots\ldots (5) \\ xyz &= \lambda(2yz + xy) \ldots\ldots\ldots (6) \\ xyz &= \lambda(2xz + 2yz) \ldots\ldots\ldots (7) \end{aligned}$$

observe that $\lambda \neq 0$ because if $\lambda = 0 \Rightarrow yz = xz = xy = 0$ and by (4) $\quad 0 = 12$ which is contradiction.

Hence from (5) and (6) we have

$$\begin{aligned} & 2xz + xy = 2yz + xy \Rightarrow 2xz = 2yz \\ \Rightarrow \quad & xz = yz. \ \texttt{But} \ z \neq 0 \ \texttt{(Since} \ z = 0 \ \texttt{gives} \ v = 0). \end{aligned}$$

So $x = y$. From (6) and (7) we have

$$2yz + xy = 2xz + 2yz \Rightarrow 2xz = xy$$

and $x \neq 0$ implies $y = 2z$.

Hence $x = y = 2z$. From (4) we have

$$4z^2 + 4z^2 + 4z^2 = 12 \Rightarrow 12z^2 = 12 \Rightarrow z^2 = 1 \Rightarrow z = \pm 1.$$

Since x, y and z are all positive, we have $x = 2, y = 2$ and

$$v_{max} = 2.2.1. = 4m^3.$$

Example 3.11.2. Find the extreme values of the function $f(x, y) = x^2 + 2y^2$ on the circle $x^2 + y^2 = 1$.

Solution. We are asked for the extreme values of f subject to the constraint for $g(x, y) = x^2 + y^2 = 1$. By using Lagrange multipliers we have $\nabla f = \lambda \nabla g, g(x, y) = 1$. Which can be written as

$$
\begin{aligned}
& f_x = \lambda g_x, \quad f_y = \lambda g_y, \quad g(x, y) = 1 \\
\Rightarrow \quad & 2x = 2x\lambda \ldots\ldots\ldots (1) \\
& 4y = 2y\lambda \ldots\ldots\ldots (2) \\
& x^2 + y^2 = 1 \ldots\ldots\ldots (3).
\end{aligned}
$$

From (1) we have $x = 0$ or $\lambda = 1$. If $x = 0$ then (3) gives $y = \pm 1$. If $\lambda = 1$ then from (2), $y = 0$ so from (3), $x^2 = 1 \Rightarrow x = \pm 1$.

$\therefore f$ has possible extreme values at the points $(0, 1), (0, -1) (1, 0)$ and $(-1, 0)$. Evaluating f at these four points, we get

$$f(0, 1) = 2, \quad f(0, -1) = 2, \quad f(1, 0) = 1, \quad f(-1, 0) = 1.$$

Therefore, the maximum value of f on the circle $x^2 + y^2 = 1$ is $f(0, \pm 1) = 2$ and the minimum value is $f(\pm 1, 0) = 1$.

Exercises 3.11

1. Find the points on the sphere $x^2 + y^2 + z^2 = 4$ that are closed to and farthest from the point $(3, 1, -1)$.

2. Use Lagrange multipliers to find the maximum and minimum values of f subject to the given constraints

   (a) $f(x, y) = x^2 y : x^2 + y^2 = 1$.

   (b) $f(x, y) = \frac{1}{x} + \frac{1}{y} : \frac{1}{x^2} + \frac{1}{y^2} = 1$.

   (c) $f(x, y, z) = x^2 + 2y^2 + 3z^2 : x + y + z = 1$.

Chapter 4

Functions, Limit and Continuity in R^n

In this chapter we have studied the limits and continuity of real and vector valued functions in R^n, compactness and continuity,connected and path connected sets and connectedness and continuity in R^n.

4.1 Vector Valued Functions

Vector Function. A function $f : A \to R^m, A \subset R^n$ is called a vector function of a vector variable.

Real Valued Function. $f(x) = (f^1(x), \ldots, f^m(x)) \forall x \in A$. Here f^i's are real valued function with the same domain as f and called the component functions of f.

Elementary Operations. If $f \& g$ are two functions with the same domain $A \subset R^n$ and range in R^m and c is any real number then,

$$\begin{aligned} (f+g)x &= f(x) + g(x) \quad \forall x \in A. \\ (c.f)(x) &= c(f(x)) \quad \forall x \in A. \end{aligned}$$

If $f : A \to R^m, A \subset R^n$ and $g : B \to R^k, B \subset R^m$, and such that range $f \subset B$ then

$$(gof)(x) = g(f(x)) \forall x \in A.$$

4.2 Limit and Continuity of Vectors and Real Valued Functions

Definition 4.2.1. (Limit of a function) Let $f : A \to R^m, A \subset R^n$ and c is a cluster point of A. f is said to have a limit l in R^m

at the point c in R^n if for every $\varepsilon > 0, \exists\delta > 0$ such that $\forall x \in A$

$$0 < |x - c| < \delta \Rightarrow |f(x) - l| < \varepsilon, \quad lim_{x\to c}f(x) = l.$$

Example 4.2.1. $f : R^3 \to R^2$ is defined by $f(x, y, z) = (x+yz, x-yz)$.

$$\lim_{(x,y,z)\to(1,1,1)} f(x, y, z) = (2, 0).$$

Solution.

$$\begin{aligned}
|f(x, y, z) - (2, 0)| &= |(x + yz, x - yz) - (2, 0)| \\
&= |x + yz - 2, x - yz| \\
&\leq |x + yz - 2| + |x - yz|. \\
&= |x - 1 + yz - 1| + |x - 1 + 1 - yz| \\
&\leq |x - 1| + |yz - 1| + |x - 1| + |yz - 1| \\
&= 2(|x - 1| + |yz - 1|) \\
&= 2(|x - 1| + |yz - y + y - 1|) \\
&\leq 2(|x - 1| + |y||z - 1| + |y - 1|).
\end{aligned}$$

so

$$|(x, y, z) - (1, 1, 1)| < \delta \Rightarrow 2(|x - 1| + |y||z - 1| + |y - 1|) < \varepsilon.$$

Now assume $0 < \delta \leq 1$. Then $|y - 1| < 1 \Rightarrow |y| < 2$.
But then

$$\begin{aligned}
|(x, y, z) - (1, 1, 1)| &< \delta \Rightarrow 2(|x - 1| + |y||z - 1| + |y - 1|) \\
&< 2|x - 1| + 4|z - 1| + 2|y - 1|.
\end{aligned}$$

Now choosing $\delta = \min[1, (\varepsilon/8)]$

$$\Rightarrow 2|x - 1| + 4|z - 1| + 2|y - 1| < 2\frac{\varepsilon}{8} + 4\frac{\varepsilon}{8} + 2\frac{\varepsilon}{8} = \varepsilon.$$

Hence

$$|(x, y, z) - (1, 1, 1)| < \delta \Rightarrow |f(x, y, z) - (2, 0)| < \varepsilon.$$

Note. If f has a limit at c along every straight line through $c \not\Rightarrow f$ has limit at c.

Example 4.2.2. $f : R^2 \to R$

$$f(x,y) = \begin{matrix} \frac{xy}{x^2-y^2} & y \neq \pm x \\ 0 & y = \pm x \end{matrix} .$$

Let $c = (0,0)$, and straight line $y = kx, k \neq \pm 1$, then

$$\lim_{x \to 0} f(x, kx) = \frac{kx^2}{x^2 - k^2 x} = \frac{k}{1-k}$$

$\lim_{(x,y)\to(0,0)} f(x,y) = 0$ along $y = \pm x$.

Along y coordinate, limit $\to 0$. So all limits are different so limit does not exist.

Theorem 4.2.1. Let $f : A \to R^m$, where $A \subset R^n$ and c is cluster point of A. $\lim_{x\to c} f(x) = l$ iff $\lim_{x\to c} f^i(x) = l^i \quad i = 1, \dots, m$.

Proof. Suppose $\lim_{x\to c} f(x) = l, \varepsilon > 0$. Choose $\delta > 0$ such that $\forall x \in A : 0 < |x - c| < \delta \Rightarrow |f(x) - l| < \varepsilon$.

But

$$\begin{aligned} & \forall x, |f^i(x) - l^i| \leq |f(x) - l| < \varepsilon, i = 1, \dots, m \\ \Rightarrow\ & \forall x, |f^i(x) - l^i| < \varepsilon, i = 1, \dots, m, 0 < |x - c| < \delta. \\ \Rightarrow\ & \lim_{x\to c} f^i(x) = l^i \forall i = 1, \dots, m. \end{aligned}$$

Conversely suppose that $\lim_{x\to c} f^i(x) = l^i, i = 1, \dots, m$. Let $\varepsilon > 0$. For each $i =, \dots, m$, choose $\delta_i > 0$ such that $\forall x \in A$:

$$0 < |x - c| < \delta_i \Rightarrow |f^i(x) - l^i| < \frac{\varepsilon}{m}, \ \texttt{let}\ \delta = \min(\delta_, \dots, \delta_m).$$

Then $\forall x \in A, 0 < |x - c| < \delta \Rightarrow |f(x) - l| \leq |f^1(x) - l| + \cdots + f^m(x) - l| < \varepsilon$.

$$\Rightarrow \lim_{x\to c} f(x) = l.$$

Properties of Limits

Lemma 4.2.1. Suppose $f : A \to R, A \subset R^n$ and c is a cluster point of A. If $\lim_{x\to c} f(x)$ exists, then $\exists M > 0, \delta > 0$ such that $\forall x \in A$

$$0 < |x - c| < \delta \Rightarrow |f(x)| < M.$$

Proof. Let $\lim_{x\to c} f(x) = l$. By definition of limit $\exists \delta > 0$, such that $\forall x \in A, 0 < |x - c| < \delta \Rightarrow |f(x) - l| < 1$.

Hence

$$\begin{aligned} \forall x \in A, 0 < |x - c| < \delta \ &\Rightarrow \ |f(x)| < 1 + |l| \quad M = 1 + |l| \\ &= \ |f(x)| < M. \end{aligned}$$

Lemma 4.2.2. Suppose $f : A \to R, A \subset R^n$ and c is a cluster point of A. If $\lim_{x\to c} f(x)$ exists and is different from zero then $\exists m > 0, \delta > 0$ such that $\forall x \in A, 0 < |x - c| < \delta \Rightarrow |f(x)| > m$.

Proof. Let $\lim_{x\to c} f(x) = l \neq 0$. By definition of limit $\exists \delta > 0$ such that $\forall x \in A, 0 < |x - c| < \delta \Rightarrow |f(x) - l| < \frac{|l|}{2}. \Rightarrow |l| < \frac{|l|}{2} + |f(x)|$, i.e.

$$\forall x \in A, 0 < |x - c| < \delta \ \Rightarrow \ \frac{|l|}{2} < |f(x)| \quad m = \frac{|l|}{2}.$$

Theorem 4.2.2. Let $f \& g$ be two real valued function with common domain $A \subset R^n$, and c a cluster point of A. If $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ exists, then

(i) $\lim_{x\to c}(f + g)(x) = \lim_{x\to c} f(x) + \lim_{x\to c} g(x)$

(ii) $\lim_{x\to c}(f.g)(x) = \lim_{x\to c} f(x). \lim_{x\to c} g(x)$

(iii) $\lim_{x\to c}(f/g)(x) = \lim_{x\to c} f(x)/\lim_{x\to c} g(x)$ provided $g\&$ its limit at c not vanish.

Proof.

(i) Let $\varepsilon > 0; \lim_{x\to c} f(x) = l_1, \lim_{x\to c} g(x) = l_2$. Then by definition of limit, $\exists \delta_1 > 0, \delta_2 > 0$, such that $\forall x \in A$

$$0 < |x - c| < \delta_{\Rightarrow}|f(x) - l_1| < \frac{\varepsilon}{2} \text{ and}$$
$$0 < |x - c| < \delta_2 \Rightarrow |g(x) - l_2| < \frac{\varepsilon}{2}.$$

Choose $\delta = \min(\delta_1, \delta_2)$. Then $\forall x \in A$

$$\begin{aligned} 0 < |x - c| < \delta \Rightarrow |f(x) &+ g(x) - l_1 - l_2| \leq |f(x) - l_1| \\ &+ |g(x) - l_2| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

$$\begin{aligned} \Rightarrow \lim_{x\to c}(f(x) + g(x)) = l_1 + l_2 \Rightarrow \lim_{x\to c}(f+g)(x) &= \lim_{x\to c} f(x) \\ &+ \lim_{x\to c} g(x). \end{aligned}$$

(ii) Let $\varepsilon > 0, \lim_{x\to c} f(x) = l_1$ and $\lim_{x\to c} g(x) = l_2$. Then by definition of limit and Lemma 4.2.1. $\exists \delta_1 > 0, \delta_2 > 0, \delta_3 > 0$ and $M > 0$ such that $\forall x \in A$,

$$\begin{aligned} 0 < |x - c| < \delta_1 &\Rightarrow |f(x)| < M \\ 0 < |x - c| < \delta_2 &\Rightarrow |g(x) - l_2| < \frac{\varepsilon}{2M} \\ 0 < |x - c| < \delta_3 &\Rightarrow |f(x) - l_1| < \frac{\varepsilon}{2(|l_2|)}. \end{aligned}$$

Now choose $\delta = \min(\delta_1, \delta_2, \delta_3)$. Then $\forall x \in A$

$$\begin{aligned} & 0 < |x - c| < \delta \Rightarrow |f(x)g(x) - l_1 l_2| \\ = \ & |f(x)g(x) - l_2 f(x) + l_2 f(x) - l_1 l_2| \\ \leq \ & |f(x)||g(x) - l_2| + |l_2||f(x) - l_1| \\ < \ & M.\frac{\varepsilon}{2M} + \frac{\varepsilon.l_2}{2(|l_2|)} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(iii) As $\lim_{x\to c} g(x) = l_2$, by definition
$0 < |x - c| < \delta_2 \Rightarrow |g(x) - l_2| < \varepsilon$. Since $l_2 \neq 0$, choice of $\varepsilon = \frac{|l_2|}{2}. \Rightarrow \left|\frac{l_2}{2}\right| < |g(x)| < \frac{3|l_2|}{2}$
and

$$\left|\frac{1}{g(x)} - \frac{1}{l_2}\right| = \left|\frac{g(x) - l_2}{g(x)l_2}\right| = \frac{|g(x) - l_2|}{|g(x)||l_2|} \leq \frac{2|g(x) - l_2|}{|l_2|^2}.$$

Since $\varepsilon > 0$ is arbitrary so $\left(\frac{|l_2|^2}{2}\right)\varepsilon$ is also positive and $\exists$ a positive number δ_4 such that

$$|g(x) - l_2| < \frac{|l_2|^2}{2}\varepsilon, 0 < |x - c| < \delta_4.$$

For $\delta = \lim(\delta_2, \delta_4)$, we get

$$\left|\frac{1}{g(x)} - \frac{1}{l_2}\right| < \varepsilon, 0 < |x - c| < \delta$$
$$\Rightarrow \lim_{x \to c} \frac{1}{g(x)} = \frac{1}{l_2}.$$

Applying the product law of limit we get

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} f(x). \lim_{x \to c} \frac{1}{g(x)} = l_1.\frac{1}{l_2} = \frac{l_1}{l_2}.$$

Proposition 4.2.1. If $f : A \to R^m, A \subset R^n$ and c is a cluster point of A such that $\lim_{x \to c} f(x)$ exists then $\lim_{x \to c} |f(x)|$ exists and furthermore $\lim_{x \to c} |f(x)| = |\lim_{x \to c} f(x)|$.

Definition 4.2.2. (Continuity) A function $f : A \to R^m, A \subset R^n$ is said to be continuous at a point $a \in A$ if a is not a cluster point of A or $\lim_{x \to a} f(x) = f(a)$.

Theorem 4.2.3. Let $f : A \to R^m.A \subset R^n$ and $a \in A$. f is continuous at a iff $f^i, i = 1, \ldots, m$ is continuous at a.

Proof. Suppose f is continuous at a, then $\lim_{x \to a} f(x) = f(a)$. But $\lim_{x \to a} f^i(x) = f^i(a)$ for $i = 1, \ldots, m \Rightarrow$ each f^i is continuous at a.

Conversely, suppose each component function f^i is continuous at a. Then $\lim_{x \to a} f^i(x) = f^i(a)$ for $i = 1, \ldots, m$. Again by Theorem $\lim_{x \to a} f(x) = f(a)$.

Now we have the following theorem which can be prove easily.

Theorem 4.2.4. Let $f, g : A \to R^m, A \subset R^n$ and $\varphi : A \to R$. If f, g and φ are continuous at $a \in A$, then $f + g, (f.g)$ and $\varphi.f$ are all continuous at $a \in A$.

Proposition 4.2.2. If $T : R^n \to R^m$ is a linear transformation, then T is continuous.

Proof. Let $\varepsilon > 0$, and $a \in R^n$. So $\exists M > 0$, such that $|Th| \le M|h| \forall h \in R^n$. Let $\delta = \frac{\varepsilon}{M+1}$. Hence $\forall x \in R^n$,

$$0 < |x - a| < \delta \Rightarrow |T_x - T_a| = |T(x-a)| \le M|x-a| < \varepsilon$$
$$\Rightarrow \lim_{x \to a} T_x = T_a.$$

Identity Function. $i : R^n \to R^n$ such that $i(x) = x \forall x \in R^n$. It is continuous function.
Projection Function. $\pi_i : R^n \to R, i = 1, \ldots, n$ such that $\pi_i(x) = x^i \forall x \in R^n$. Since for all $x \in R^n, i(x) = (\pi_1(x), \ldots, \pi_n(x))$. π_i is continuous due to continuity of i.
Constant Function. $c : R^n \to R^m$ such that $c(x) = c \forall x \in R^n$. It is also continuous.
Sum Function. $S(x,y) = x + y, \quad S, p : R^2 \to R.$
Product Function. $p(x,y) = xy.$

Proposition 4.2.3. S and p are continuous functions.

Proof. To prove S is continuous. Let $\varepsilon > 0$, and $(x_0, y_0) \in R^2$. We see that $|x - x_0| \le |(x,y) - (x_0,y_0)|$ and $|y - y_0| \le |(x,y) - (x_0,y_0)|$ and $|S(x,y) - S(x_0,y_0)| = |x + y - x_0 - y_0| \le |x - x_0| + |y - y_0|$.

Let $\delta = \frac{\varepsilon}{2}$. Now

$$0 < |(x,y) - (x_0,y_0)| < \delta \Rightarrow |x - x_0| < \delta, |y - y_0| < \delta$$
$$\Rightarrow |x - x_0| < \frac{\varepsilon}{2} |y - y_0| < \frac{\varepsilon}{2} \Rightarrow |S(x,y) - S(x_0,y_0)| < \varepsilon$$
$$\Rightarrow \lim_{(x,y) \to (x_0,y_0)} S(x,y) = S(x_0,y_0).$$

To prove p is continuous. Let $\varepsilon > 0$ and $(x_0, y_0) \in R^2$. Again $|x - x_0| \le |(x,y) - (x_0,y_0)|, |y - y_0| \le |(x,y) - (x_0,y_0)|$ and

$$|p(x,y) - p(x_0,y_0)| = |xy - x_0y_0| = |xy - xy_0 + xy_0 - x_0y_0| \le |x||y - y_0| + |y_0||x - x_0|.$$

Let $0 < \delta \le 1$. Then

$$|(x,y) - (x_0,y_0)| < \delta \Rightarrow |x - x_0| < 1 \Rightarrow |x| \le 1 + |x_0|.$$

Hence $\forall (x,y)$ such that

$$0 < |(x,y) - (x_0,y_0)| < \delta, \text{ and}$$
$$|p(x,y) - p(x_0,y_0)| \le (1 + |x_0|)(|y - y_0|) + |y_0||x - x_0|.$$

Now $\delta = \min\left(1, \frac{\varepsilon}{1+|x_0|+|y_0|}\right)$
then

$$\begin{aligned} 0 < |(x,y)-(x_0,y_0)| < \delta \quad &\Rightarrow \quad |p(x,y)-p(x_0,y_0)| \\ &\leq \quad (1+|x_0|)\frac{\varepsilon}{1+|x_0|+|y_0|} \\ &+ \quad |y_0|\frac{\varepsilon}{1+|x_0|+|y_0|} = \varepsilon \end{aligned}$$

$\Rightarrow \lim_{(x,y)\to(x_0,y_0)} p(x,y) = p(x_0,y_0)$.

Now we have

$$P(x^1,\ldots,x^n) = \sum_{0\leq i_1+\cdots+i_n\leq k} a_{i_1,\ldots,i_n}(x^1)^{i_1},\ldots,(x^n)^{i_n}, k\in N,$$

of n variables and degree k is a continuous function.

Lemma 4.2.3. Let $f : A \to R^m$ and $g : B \to R^k, A \subset R^n$ and $B \subset R^m$ such that $B \supseteq$ range f. If $\lim_{x\to a} f(x) = b \in B$ and g is continuous at b then $\lim_{x\in a}(gof)(x) = g(b)$.

Proof. Let $\varepsilon > 0$. Since g is continuous at $b \exists \delta_1 > 0$ such that $\forall y \in B, 0 < |y-b| < \delta_1 \Rightarrow |g(y)-g(b)| < \varepsilon$.

Since $\lim_{x\to a} f(x) = b, \exists \delta > 0$ such that $\forall x \in A$

$$0 < |x-a| < \delta \Rightarrow |f(x)-b| < \delta_1.$$

Hence $\forall x \in A; 0 < |x-a| < \delta \Rightarrow |g(f(x))-g(b)| < \varepsilon$.

$$\Rightarrow \lim_{x\to a}(gof)(x) = g(b).$$

Theorem 4.2.5. Let $f : A \to R^m, g : B \to R^k, A \subset R^m$ and such that range $f \subset B$. Suppose f is continuous at $a \in A$, and g is continuous at $f(a) \in B$. Then gof is continuous at a.

Example 4.2.3. Let $f : R^4 \to R$ defined by $f(x,y,z,w) = \sin[\cos(x+y+z)+w^2]$. Show that f is continuous.

Solution.

$$f = \sin 0s(\cos 0s(s(\pi_1,\pi_2),\pi_3), p(\pi_4,\pi_4)),$$

continuity of f follows from the continuity of $\sin, \cos, S, p$ and π_1, π_2, π_3 and π_4.

Theorem 4.2.6. Let $f : A \to R^m, A \subset R^n$. f is continuous on A iff for every open set $U \subset R^m$ there is an open set $V \subset R^n$ such that $f^{-1}(U) = V \cap A$.

Corollary 4.2.1. Let $f : A \to R^m, A \subset R^n$. f is continuous on A iff for every closed set $F \subset R^m$ there is a closed set $H \subset R^n$ such that $f^{-1}(F) = H \cap A$.

Definition 4.2.3. Let $f : A \to R, A \subset R^n$ be a bounded function, $a \in R^n$. The oscillation $O(f,a)$ of f at a is defined by $O(f,a) = \lim_{\delta\to 0}[M(a,f,\delta) - m(a,f,\delta)]$, where

$$\begin{aligned} M(a,f,\delta) &= \sup\{f(x) : x \in A \texttt{ and } |x-a| < \delta\} \\ m(a,f,\delta) &= \inf\{f(x) : x \in A \texttt{ and } |x-a| < \delta\}. \end{aligned}$$

Theorem 4.2.7. Let $f : A \to R, A \subset R^n$ be a bounded function, $a \in R^n$. f is continuous at a iff the oscillation of f at a is zero.

Proof. Suppose f is continuous at a. Let $\varepsilon > 0$. Then $\exists\delta > 0$ such that $\forall x \in A : |x-a| < \delta \Rightarrow |f(x) - f(a)| < \frac{\varepsilon}{2}$.

$\Rightarrow \forall x \in A$,

$$f(a) - \frac{\varepsilon}{2} < f(x) < f(a) + \frac{\varepsilon}{2}$$

$\Rightarrow M(a,f,\delta) \le f(a) + \frac{\varepsilon}{2}$ and $m(a.f,\delta) \ge f(a) - \frac{\varepsilon}{2}$.

Hence

$$M(a,f,\delta) - m(a.f,\delta) \le \frac{\varepsilon}{2}.$$

Consequently for all $\delta' < \delta, M(a,f,\delta') - m(a,f,\delta') \le \varepsilon$

Hence

$$\begin{aligned} &\lim_{\delta\to 0}[M(a,f,\delta) - m(a.f,\delta)] \le \varepsilon \\ \Rightarrow\ &\lim_{\delta\to 0}[M(a,f,\delta) - m(a.f,\delta)] = 0. \end{aligned}$$

Conversely, suppose $O(f,a) = 0$. Let $\varepsilon > 0, \exists\delta > 0$ such that

$[M(a, f, \delta') - m(a.f, \delta')] < \varepsilon \forall \delta' \leq \delta$. But then $\forall x_1, x_2 \in B(a, \delta) \cap A : |f(x) - f(x_2)| < \varepsilon \Rightarrow \forall x \in A : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon \Rightarrow f$ is continuous.

Theorem 4.2.8. Let $A \subset R^n$ be closed. If $f : A \to R$ is a bounded function and $\varepsilon > 0$, then $\{x \in A : O(f, x) \geq \varepsilon\}$ is closed.

Proof. Let $B = \{x \in A : O(f, x) \geq \varepsilon\}$. Suppose $x \notin B$. Then either $x \notin A$ or $x \in A$ and $O(f, x) < \varepsilon$. If $x \notin A$ then since A is closed $\exists$ an open ball $B(x) \subset A^c$. Hence $B(x) \subset B^c$. If $x \in A$ and $O(f, x) < \varepsilon$ then $\exists \delta > 0$ such that $M(x, f, \delta) - m(x, f, \delta) < \varepsilon$. Now consider the open ball $B(x, \delta)$. Let $y \in B(x, \delta) \cap A$. Set $\delta_1 = \delta - |x - y|$. But then $M(y, f, \delta_1) - m(y, f, \delta_1) < \varepsilon$. Hence $O(f, y) < \varepsilon$. Hence $y \notin B^c$. If $y \notin A$ then clearly $y \in B^c$. Hence in any case $B(x, \delta) \subset B^c \Rightarrow B$ is closed set.

Corollary 4.2.2. Let $A \subset R^n$ be a closed set in R^n, and $f : A \to R$ be a bounded function. Then the set of all points of discontinuity of f is the union of a countable family of closed sets.

Proof. Let $B_n = \{x \in A : O(f, x) \geq \frac{1}{n}\} n \in N$. If f is discontinuous at a point $c \in A$ then $O(f, c) > 0$. Hence $c \in B_n$ for some n, and conversely if $c \in B_n$ for some n then f is not continuous at c. Hence the set of all points of discontinuity of f is $U_{x \in N} B_n$. But each B_n is closed.

Exercises 4.2

1. Let m and n be positive integers. Show that the limit

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

exists if and only if $m + n > 2$.

2. Give an example of a function $f : R^n \to R$ such that $lim_{x\to 0}$ $f(x) = 0$ but that $lim_{x\to 0} f(x)/\|x\| \neq 0$.

3. For a function $f : R^n \to R$ and a positive integer m, show that

$$\lim_{x\to 0} f(x)/\|x\|^{m+1} = 0 \text{ implies that } \lim_{x\to 0} f(x)/\|x\|^m = 0,$$

but that the converse does not hold.

4. Define the function $f : R^2 \to R$ by

$$f(x, y) = \cos(x + y) + x^2y^2 \texttt{ for } (x, y) \in R^2.$$

Prove that $f : R^2 \to R$ is continuous.

5. Define $O = \{(x, y, z) \text{ in } R^3 : (x, y, z) \neq (0, 0, 0)\}$ and define the function $f : O \to R$ by

$$f(x, y, z) = \frac{x}{x^2 + y^2 + z^2} \texttt{ for } (x, y, z) \texttt{ in } O.$$

Prove that the function $f : O \to R$ is continuous.

6. Suppose that the function $f : R^n \to R$ is continuous and that $f(u) > 0$ if the point u in R^n has at least one rational component. Prove that $f(u) \geq 0$ for all points u in R^n.

7. For a point ν in R^n and define the function $f : R^n \to R$ by

$$f(u) = (u, \nu) \texttt{ for } u \texttt{ in } R^n.$$

Prove that the function $f : R^n \to R$ is continuous.

4.3 Compactness and Continuity

Theorem 4.3.1. If $f : A \to R^m, A \subset R^n$ is continuous and A is compact then $f(A)$ is compact.

Proof. Suppose $f(A)$ is not compact. Then $\exists$ an open cover O of $f(A) \ni$ no finite sub collection of O covers $f(A)$. Let $O' = \{V \subset R^n : V$ is open in R^n and $\ni u \in O$ such that $V \cap A : f^{-1}(u)\}, O'$ is an open cover of A. Since A is compact $\exists$ a finite sub collection $\{V_1, \ldots, V_k\}$ that covers A. But then the sub collection $\{u_1, \ldots, u_k\}$ of O where $f^{-1}(u_i) = V_i \cap A, i = 1, \ldots, k$ covers $f(A)$. Contradiction. Hence $f(A)$ is compact.

Corollary 4.3.1. Extremum Value Theorem. Let $A \subset R^n$, be compact. If $f : A \to R$ is continuous then f takes on a maximum and minimum value on A.

Proof. $f(A)$ is compact. Hence $f(A)$ is closed and bounded in R. Hence $f(A)$ has a supremum and an infimum that are contained in $f(A)$. But then the supremum and infimum are respectively the maximum and the minimum of f on A.

Corollary 4.3.2. Let $A \subset R^n$ be compact. If $f : A \to R^m$ is continuous then $|f|$ takes on a maximum and minimum value on A.

Proof. f is continuous $\Rightarrow |f|$ is also continuous, and hence using above Corollary 4.3.1. to $|f|$, we get result.

Theorem 4.3.2. Let $A \subset R^n$ be compact. If $f : A \to R^m$ is continuous and one to one then the inverse function is continuous.

Proof. Since $f : A \to R^m$ is one to one the inverse of f denoted by $g : f(A) \to A$ exists. Since A is compact $\Rightarrow A$ is closed. Let F be any closed set in R^n. Clearly

$$g^{-1}(F) = g^{-1}(F \cap A) = f(F \cap A).$$

$F \cap A$ is closed and bounded $\Rightarrow$ it is compact. Hence $f(F \cap A)$ is compact $\Rightarrow$ it is closed. But $f(F \cap A) \subset f(A)$. Hence $g^{-1}(F) = f(A) \cap f(F \cap A)$

$\Rightarrow g$ is continuous.

Definition 4.3.1. Given a function $f : A \to R^m, f$ is said to be uniformly continuous on A iff $\forall \varepsilon > 0, \exists \delta > 0$ such that

$$\forall x, y \in A : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Note 4.3.1. If f is uniformly continuous on A, then f is continuous at every point of A i.e., f is continuous on A.

Note 4.3.2. But f is continuous on A need not imply that f is uniformly continuous on A.

Example 4.3.1. $f : (0, 1) \to R, \quad f(x) = 1/x$

f is continuous on $(0,1)$. But f is not uniformly continuous on $(0,1)$, we have to show that $\exists\varepsilon > 0, \forall\delta > 0$ such that $x_0, \& y_0 \in (0,1)$

$$|x_0 - y_0| < \delta \text{ but } |f(x_0) - f(y_0)| \geq \varepsilon.$$

Taking

$$\begin{aligned} x_0 = \delta, y_0 = \frac{\delta}{\varepsilon+1} \text{ then } |x_0 - y_0| &= \left|\delta - \frac{\delta}{\varepsilon+1}\right| = \frac{\varepsilon\delta}{\varepsilon+1} \\ &< \delta, \quad (0 < \delta < 1) \end{aligned}$$

but

$$|f(x_0) - f(y_0)| = \left|\frac{1}{x_0} - \frac{1}{y_0}\right| = \left|\frac{1}{\delta} - \frac{\varepsilon+1}{\delta}\right| = \frac{\varepsilon}{\delta} > \varepsilon.$$

Theorem 4.3.3. Let $A \subset R^n$ be compact. If $f : A \to R^m$ is continuous then f is uniformly continuous.

Proof. $f : A \to R^m$ is continuous $\Rightarrow$ for any $a \in A$ and $\forall\varepsilon > 0, \exists\delta > 0$ such that

$$|x - a| < \delta_a \Rightarrow |f(x) - f(a)| < \varepsilon/2.$$

The collection

$$\Im = \left\{B_{\frac{\delta a}{2}}(a) : a \in A\right\} \text{ is an open cover of } A.$$

Since A is compact $\exists$ a finite sub collection

$$\left\{B_{\frac{\delta a_1}{2}}(a_1), B_{\frac{\delta a_2}{2}}(a_2), \ldots, B_{\frac{\delta a_n}{2}}(a_n)\right\} \text{ of } \Im \text{ covering } A.$$

Now let $\delta = \min\left\{\frac{\delta a_1}{2}, \frac{\delta a_2}{2}, \ldots, \frac{\delta a_n}{2}\right\}$.

Let $x, y \in A$ be any points such that $|x - y| < \delta$, then we have to show that $|f(x) - f(y)| < \varepsilon$.
Moreover $y \in A \Rightarrow y \in B_{\frac{\delta a_j}{2}}$ for some $j \in \{1, \ldots, n\}$.

$$\Rightarrow |y - a_j| < \frac{\delta a_j}{2} < \delta a_j \Rightarrow |f(y) - f(a_j)| < \varepsilon/2.$$

[In fact for any $z \in A : |z - a_i| < \delta a_i \Rightarrow |f(x) - f(a_i)| < \varepsilon$]

$$\begin{aligned} |x - a_j| &\le |x - y| + |y - a_j| \\ &\le \delta + \frac{\delta a_j}{2} \le \delta a_j \Rightarrow |f(x) - f(a_j)| < \varepsilon. \end{aligned}$$

Thus

$$|x - y| < \delta \Rightarrow |x - y| < \delta a_j \forall j.$$

$$\begin{aligned} &|x - y| \le |x - a_j| + |y - a_j| < \delta a_j \\ &\Rightarrow |f(x) - f(y)| \le |f(x) - f(a_j)| + |f(y) - f(a_j)| < \varepsilon. \end{aligned}$$

Exercises 4.3

1. Can a symmetric neighborhood in R^n be compact ?
2. Let A be a subset of R^n and the function $f : A \to R^m$ be continuous. If A is bounded, is $f(A)$ bounded ?
3. Suppose that the function $f : R^n \to R$ is continuous, and $f(u) \ge \|u\|$ for every point u in R^n. Prove that $f^{-1}([0, 1])$ is compact.
4. Let A and B be compact subsets of R. Define $K = \{(x, y) \in R^2 : x \in A, y \in B\}$. Prove that K is compact.
5. Let u be a point in R^n, and let r be a positive number. Prove that the set $\{\nu \in R^n : d(u, \nu) \le r\}$ is compact.
6. Which of the following subsets of R is compact.
   (a) $\{x \in R : x^2 > x\}$.
   (b) $\{x \in R : e^x - x^2 \le 2\}$.

4.4 Connected and Path-connected Sets

Definition 4.4.1. Let $A \subseteq R^n$ and U, V open subsets of R^n. $U \& V$ are said to be a separation of A if

(i) $U \cap A \ne \phi$.

(ii) $V \cap A \neq \phi$.

(iii) $(U \cap A) \cap (V \cap A) = \phi$.

iv) $(U \cap A) \cup (V \cap A) = A$.

Definition 4.4.2. A subset A of R^n is said to be a disconnected set if it has a separation U and V. A is called connected if it is not disconnected.

Example 4.4.1.

(i) $A = R \backslash \{1\}$ is a disconnected set.

Let $U = (-\infty, 1), V = (1, \infty), U \& V$ are separation of A.

(ii) $A = \{(x, y) : y = 1/x \forall x > 0\} \subset \{(x, y) : y = 0, x > 0\}$.

Let $C = \{(x, y) : y = 1/x, x > 0\}$ $B = \{(x, y) : y = 0, x > 0\}$.

A is not disconnected. Any open set containing B contains some point of C.

Theorem 4.4.1. A set $C \subseteq R$ is connected iff it is an interval.

Proof. Let C be a connected set. Suppose C is not an interval $\Rightarrow \exists x \& y \in C, x < y$ such that $[x, y] \nsubseteq C \Rightarrow \exists z \in (x, y)$ such that $z \notin C$. Now consider the open sets $U = (-\infty, z)$ and $V = (z, \infty), U \& V$ are separation of $C \Rightarrow C$ is not connected $\Rightarrow C$ must be an interval.

Conversely, suppose now C is an interval and assume that C is not connected. Let $U \& V$ be separation of C. Let $w \in U \cap C$ and $z \in V \cap C$ and let $w < z$. Note that $[w, z] \subseteq C$. Consider now the sets $\begin{array}{l} S_1 = \{x \in [w, z]/[w, x) \subseteq U\} \\ S_2 = \{x \in [w, z]/[x, z) \subseteq V\} \end{array}$ and.

Both $S_1 \& S_2$ are non empty and bounded sets.

Let $\alpha = lub S_1$ and $\beta = glb S_2$

$\alpha \leq z$ and $\beta \geq w$

i) If $\alpha \in U$, then since U is open $\exists x_0 \in U$ such that $\alpha < x_0$ and $[z, x_0) \leq U \Leftrightarrow \alpha = \sup S_1$.

ii) If $\alpha \notin U$, then since $\alpha \in C, \alpha \in V$ and $\beta = glbS_2 \Rightarrow \beta < \alpha$.

[If $\beta = \alpha$ then $\beta \in V \Rightarrow \exists x_1 \in V$ such that $x_1 < \beta$ and $(x_1, w] \leq V \Leftrightarrow$ to $\beta = glbS_2$]

$\beta < \alpha \in V, V$ is open $\Rightarrow \exists x_1 \in V$ such that $\beta < x_1 < x$.

(a) If $x_1 \in U$, then $x_1 \in intU$ and $x_1 \in intV$ $\Rightarrow (intU) \cap (intV) \neq \phi$.

(b) If $x_1 \notin U$, then x_1 is an upper bound and $x_1 < \alpha \Leftrightarrow$ since $\alpha = lubU$. Thus $\alpha \in V$ is not possible.

Hence $\alpha \notin U$ and $\alpha \notin V$ but $\alpha \in C$ and $U \& V$ are separation of C.

$$\Rightarrow \alpha \in [U \cap C] \cup [V \cap C] = C$$
$$\Rightarrow \alpha \in U \& \alpha \in V,$$

which is a contradiction. Hence C must be connected.

Exercises 4.4

1. Let Q be the set of rational numbers. Show that Q is not connected.

2. Show that the set $S = \{(x, y) \in R^2$: either x or y is rational $\}$ is path-wise connected.

3. Let A and B be path-wise connected subsets of R whose intersection $A \cap B$ is nonempty. Prove that the union $A \cup B$ is also path-wise connected.

4. Let a and b be positive real numbers. Show that the ellipse

$$\{(x, y) \in R^2 : x^2/a^2 + y^2/b^2 = 1\}$$

is path-wise connected.

5. Let A and B be convex subsets of R^n. Prove that the intersection $A \cap B$ is also convex. Is it true that the intersection of two path-wise connected subsets of R^n is also path-wise connected ?

6. Given a point u in R^n and a point ν in R^m, we define the point (u, ν) to be the point in R^{n+m} whose first n components coincide with the components of u and whose last m components coincide with those of ν. Suppose A is subset of R^n and that $F : A \to R^m$ is continuous. The graph G of this mapping is defined by

$$G = \{(u, \nu) \in R^{n+m} : u \in A, \nu = F(u)\}.$$

Show that if A is path-wise connected, then G is also path-wise connected.

4.5 Connectedness and Continuity

Theorem 4.5.1. Let $A \subseteq R^n$ be a connected set. If $f : A \to R^m$ is continuous then $f(A)$ is connected.

Proof. Suppose $f(A)$ is disconnected. Then $\exists U, V \subseteq R^m$ open sets which are separation of $f(A)$.

f is continuous ,$U \& V$ are open sets $\Rightarrow \exists U_0, V_0 \subseteq R^n$ such that $f^{-1}(U) = U_0 \cap A$ and $f^{-1}(V) = V_0 \cap A$

Claim. U_0 and V_0 are separation of A.

(i) $f^{-1}(U) \cap f^{-1}(V) = \phi$ (otherwise let $x \in f^{-1}(U) \& x \in f^{-1}(V) \Rightarrow$ $f(x) \in U \cap f(A)$ and $f(x) \in V \cap f(A)$

$\Rightarrow [U \cap f(A)] \cap [V \cap f(A)] \neq \phi$.

Thus $f^{-1}(U) \cap f^{-1}(V) = \phi \Leftrightarrow (U_0 \cap A) \cap (V_0 \cap A) = \phi$.

(ii) $(U_0 \cap A) \cup (V_0 \cap A) = A$

let

$$\begin{aligned} & x \in A \Rightarrow f(x) \in f(A) = (U \cap A) \cup (V \cap A) \\ \Rightarrow\ & f(x) \in (U \cap A) \texttt{ or } f(x) \in (V \cap A) \\ \Rightarrow\ & f(x) \in U \texttt{ or } f(x) \in V \\ \Rightarrow\ & x \in f^{-1}(U) \texttt{ or } x \in f^{-1}(V) \Rightarrow x \in (U_0 \cap A) \cup (V_0 \cap A). \end{aligned}$$

Since $(U_0 \cap A) \cup (V_0 \cap A) \subseteq A \Rightarrow (U_0 \cap A) \cup (V_0 \cap A) = A$.

(iii) $U \cap f(A) \neq \phi \Rightarrow f^{-1}(U) = U_0 \cap A \neq \phi$

$V \cap f(A) \neq \phi \Rightarrow f^{-1}(V) = V_0 \cap A \neq \phi.$

From (i) (ii) & (iii) A is disconnected $f(A)$ must be connected.

Definition 4.5.1. A set $C \subseteq R^n$ is said to be a convex set if for all $x, y \in C$ and $t \in [0,1], tx + (1-t)y \in C$.

Note. $z = tx + (1-t)y$

$= y + t(x - y)$

$\Rightarrow$ every point on the line segment determined by the point $x \& y$ in C.

Fig. 4.5.1

Theorem 4.5.2. Any convex set in R^n is connected.

Proof. Let C be a convex set in R^n. Suppose C is not connected, then $\exists$ open sets U and V in R^n such that U and V are separation of C. Now let $x \in U \cap C \neq \phi$ and $y \in V \cap C \neq \phi$. Consider the function $f : [0,1] \to C$ defined by $f(t) = tx + (1-t)y$, f is continuous on $[0,1]$. Hence $\exists$ open sets O_1 and $O_2 \subseteq R$ such that

$$f^{-1}(U) = O_1 \cap [0,1] \text{ and } f^{-1}(V) = O_2 \cap [0,1]$$
$$\exists f(0) = y \Rightarrow O_1 \cap [0,1] \neq \phi, f(1) = x \Rightarrow O_2 \cap [0,1] \neq \phi$$

furthermore $f^{-1}(U) \cap f^{-1}(V) = \phi$ and $f^{-1}(U) \cup f^{-1}(V) = [0,1]$

$\Rightarrow O_1$ and O_2 are separation of $[0,1]$ (since $[0,1]$ is an interval). Thus C must be connected.

Corollary 4.5.1. R^n is connected.

Proof. R^n is convex. (For any $x, y \in R^n$ and $t \in [0,1]$

$$tx + (1-t)y \in R^n.$$

Example 4.5.1. In R^n, the only subsets that are open and closed are ϕ and R^n.

Proof. Let $S \subseteq R^n, S \neq \phi$ and $S \neq R^n$ and S is both open and closed. Let $V = S$ and $U = S^c$. V and U are open sets, $U \neq \phi$ and $V \neq \phi$.

$$(U \cap R^n) \cap (V \cap R^n) = \phi \ \texttt{and} \ (U \cap R^n) \cup (V \cap R^n) = R^n$$

$\Rightarrow U$ and V are separations of $R^n \Rightarrow$. Thus R^n and ϕ are the only subsets.

Example 4.5.2. Let A be a connected set and $A \subseteq B \subseteq \overline{A}$. Then B is also connected.

Proof. Let A is connected, we have to prove that B is connected. Assume that B is not connected then $\exists$ open sets $U, V \subseteq R^n$ such that $U \cap B \neq \phi, V \cap B \neq \phi$.

$$(U \cap B) \cap (V \cap B) = \phi, (U \cap B) \cup (V \cap B) = B.$$

i) $\Rightarrow U \cap A \neq \phi$ and $V \cap A \neq \phi$ (U, V are open sets and $x \in B \Rightarrow x \in A$ or x is a limit point of A).

ii) $(U \cap A) \cap (V \cap A) \subseteq (U \cap B) \cap (V \cap B) = \phi$.

iii) $(U \cap A) \cap (V \cap A) \subseteq A$. Conversely let $x \in A$

$\Rightarrow x \in B = (U \cap B) \cup (V \cap B) \Rightarrow x \in U \cap B$ or $x \in V \cap B$

$\Rightarrow x \in U \cap A$ or $x \in V \cap A \Rightarrow A \subseteq (U \cap A) \cup (V \cap A)$

$\Rightarrow (U \cap A) \cap (V \cap A) = A$.

$\Rightarrow A$ is not connected $\Leftrightarrow B$ must be connected. It follows that $\overline{A}$ is connected.

Definition 4.5.2. A set $A \subseteq R^n$ is said to be a path connected set if for every $x, y \in A$, there is a continuous function $f : [a, b] \to A$ such that $f(a) = x$ and $f(b) = y$, where $[a, b]$ is a closed interval.

i.e., A is path connected if any two points $x, y \in A$ can be connected by a continuous curve which lie in A.

Theorem 4.5.3. Any path connected set is a connected set.

Proof. Let A be a path connected set. Suppose A is not connected. Then $\exists$ open sets $U\&V \subseteq R^n$ that separate A. Let $x \in U \cap A$ and $y \in V \cap A$. Since A is path connected there is a continuous function $f : [a, b] \to A$ such that $f(a) = x, f(b) = y$, then $\exists$ open sets $O_1, O_2 \subseteq R$ such that $f^{-1}(U) = O_1 \cap [a, b]$ and $f^{-1}(V) = O_2 \cap [a, b], O_1\&O_2$ are separation of $[a, b] \Rightarrow A$ is connected.

Exercises 4.5

1. Prove that if A is open and connected, then it is path connected.

2. Show that

$$\lim_{(x,y)\to(0,0)} \frac{xy^3}{x^2 + y^6} \texttt{ does not exist} .$$

3. Which subsets of R are both compact and connected ?

4. Let K be a compact subset of R^n. Prove that K is not connected if and only if there are nonempty, disjoint subsets A and B of K, with $A \cap B = K$ and a positive number ε such that $d(u, \nu) > \varepsilon$ for all u in A and all ν in B. Is the assumption of compactness necessarily for the existence of such an ε ?

5. Suppose that A is a subset of R^n that fails to be connected, and let u and ν be open subsets of R^n that separate A. Suppose that B is a subset of A that is connected. Prove that either $B \leq u$ or $B \subseteq \nu$.

Chapter 5

Differentiation in R^n

In this chapter we consider functions mapping R^n into R^m, and define the derivative of such functions.Chain rule, partial derivatives, directional derivatives and Mean Value Theorem also have been discussed.Two major results of this chapter are the inverse function theorem, which give conditions under which a differentiable function from R^n to R^n has a differentiable inverse, and the implicit function theorem, which provides the theoretical understanding for the technique of implicit differentiation as studied in calculus.Here we have simply generalized facts that are already familiar in calculus.

5.1 Introduction

Definition 5.1.1. Let A be an open set in R^n and $f : A \to R^m$. f is said to be differentiable at $a \in A$ if and only if there is a linear transformation $\lambda : R^n \to R^m$ such that

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

If f is differentiable at every point of A, then f is said to be a differentiable function.

Remark 5.1.1. The above definition is equivalent to say that f is differentiable at $a \in A \subseteq R^n$ if $\exists$ a linear transformation $\lambda : R^n \to R^m$ such that $\forall \varepsilon > 0$ there exists $\delta > 0$ such that $\forall x \in A : |x - a| < \delta \Rightarrow |f(x) - f(a) - \lambda(x - a)| \leq \varepsilon |x - a|$.

Example 5.1.1. Let $f : R^n \to R^m$ be given by $f(x) = c, c \in R^m$ is constant then f is differentiable.

Solution. $\lambda(h) = \overline{0}$, the zero transformation, then

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = \lim_{h \to 0} \frac{|c - c - \overline{0}|}{|h|} = 0.$$

$\Rightarrow f$ is differentiable.

Example 5.1.2. Let $f : R^n \to R^m$ be a linear transformation, then f is differentiable.

Solution. Let $\lambda(h) = f(h)$, then

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

$\Rightarrow f$ is differentiable.

Theorem 5.1.1. Let $A \subseteq R^n$ be open. If f is differentiable at $a \in A$, then the linear transformation $\lambda : R^n \to R^m$ is unique.

Proof. Let f be differentiable at $a \in A$ and $\lambda : R^n \to R^m$ and $\mu : R^n \to R^m$ be linear transformations such that

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = 0 = \lim_{h \to 0} \frac{|f(a+h) - f(a) - \mu(h)|}{|h|}. \tag{5.1.1}$$

Note that $\forall x \in R^n$ if we show
$|\mu(x) - \lambda(x)| = 0$, then we are done.

Consider

$$\begin{aligned} \frac{|\mu(h) - \lambda(h)|}{|h|} &= \frac{|\mu(h) - f(a+h) + f(a) + f(a+h) - f(a) - \lambda(h)|}{|h|} \\ &\leq \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} \\ &\quad + \frac{|f(a+h) - f(a) - \mu(h)|}{|h|} \end{aligned}$$

using (5.1.1), we get

$$\lim_{h \to 0} \frac{|\mu(h) - \lambda(h)|}{|h|} = 0.$$

Let $x \neq 0$ be in R^n, then for any $t \in R$

$$\frac{|\mu(tx) - \lambda(tx)|}{|tx|} = \frac{|\mu(x) - \lambda(x)|}{|x|}.$$

Furthermore

$$\lim_{t \to 0} \frac{\mu(tx) - \lambda(tx)}{|tx|} = 0 \quad \Rightarrow \quad \frac{|\mu(x) - \lambda(x)|}{|x|} = 0$$
$$\Rightarrow \quad \mu = \lambda.$$

Definition 5.1.2. Let $A \subseteq R^n$ be open and the function $f : A \to R^m$ differentiable. Then the unique linear map $\lambda : R^n \to R^m$ in the definition of differentiability of f, is called the derivative of f at a and is denoted by $Df(a)$.

Example 5.1.3. If $f : A \to R, A \subseteq R, f$ is differentiable at $a \in A$ means

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \texttt{ exists,}$$

and this limit is denoted by $f^{'}(a)$.

If we let $\lambda(h) = f^{'}(a)h$, then

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - f^{'}(a)h}{h} = 0.$$

Example 5.1.4. Let $f : R^2 \to R$ be defined by $f(x, y) = \sin x$. Then $Df(a, b) = \lambda$ satisfies $\lambda(x, y) = \cos a.x$.

Solution. Note that

$$\lim_{(h,k) \to 0} \frac{f(a+h, b+k) - f(a, b) - \lambda(h, k)}{|(h, k)|}$$
$$= \lim_{(h,k) \to 0} \frac{|\sin(a+h) - \sin a - \cos a.h|}{|(h, k)|}$$

we have that

$$\lim_{h \to 0} \frac{|\sin(a+h) - \sin a - \cos a.h|}{|h|} = 0$$

since $|(h,k)| \geq |h|$, it is also true that

$$\lim_{h\to 0} \frac{\sin(a+h) - \sin a - \cos a.h}{|(h,k)|} = 0.$$

Theorem 5.1.2. If f is differentiable at a then it is continuous at a.

Proof. We have to prove that $\lim_{h\to 0} |f(a+h) - f(a)| = 0$. f is differentiable at $a \Rightarrow$.

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

Note that

$$\begin{aligned} 0 \leq |f(x) - f(a)| &\leq |f(x) - f(a) - \lambda(x-a)| + |\lambda(x-a)| \\ &\leq |f(x) - f(a) - \lambda(x-a)| + M|x-a|, \\ &\quad \texttt{for some } M > 0. \end{aligned}$$

$$\lim_{x\to a} \frac{|x-a||f(x) - f(a) - \lambda(x-a)|}{|x-a|} + M|x-a| = 0.$$

$\Rightarrow \lim_{x\to a} |f(x) - f(a)| = 0$. Therefore f is continuous at a.

Exercises 5.1

1. Let $g : R \to R$ be function of class C^2. Show that

$$\lim_{h\to 0} \frac{g(a+h) - 2g(a) + g(a-h)}{h^2} = g''(a).$$

2. Show that the function $f(x,y) = |xy|$ is differentiable at 0, but is not of class C^1 in any neighborhood of 0.

3. Suppose that A is an $m \times n$ matrix. Define the mapping $F : R^n \to R^m$ by

$$F(x) = Ax \texttt{ for every } x \texttt{ in } R^n.$$

Prove that $DF(x) = A$ for all x in R^n.

5.2 Chain Rule

If $f : A \to R^m$ is differentiable at $a \in A \subseteq R^n$ and $g : B \to R^P$ is differentiable at $f(a) \in B \subseteq R^m$, then $gof : A \to R^P$ is differentiable at a, and

$$D(gof)(a) = Dg(fa)0Df(a).$$

If $m = n = 1$, we obtain the old Chain Rule.

Proof. Let $b = f(a)$, let $\lambda = Df(a)$ and $\mu = Dg(f(a))$. If we define

1. $\phi(x) = f(x) - f(a) - \lambda(x - a) : \lambda : R^n \to R^m$.
2. $\varphi(y) = g(y) - g(b) - \mu(y - b) : \mu : R^m \to R^P$.
3. $S(x) = gof(x) - gof(a) - \mu 0 \lambda(x - a) : \mu 0 \lambda : R^n \to R^P$ then
4. $\lim_{x\to a} \frac{|\phi(x)|}{|x-a|} = 0$ since f is differentiable at a.
5. $\lim_{y\to b} \frac{|\varphi(y)|}{|y-b|} = 0$ g is differentiable at b.

and we need to show

$$\lim_{x\to a} \frac{|S(x)|}{|x-a|} = 0.$$

Note.

$$\begin{aligned} S(x) &= g(f(x)) - g(f(a)) - \mu(\lambda(x - a)) \\ &= g(f(x)) - g(f(a)) - \mu(f(x) - f(a) - \phi(x)) \\ &= g(f(x)) - g(f(a)) - \mu(f(x) - f(a)) + \mu(\phi(x)) \\ &= g(f(x)) - g(b) - \mu(f(x) - b) + \mu(\phi(x)) \\ S(x) &= \varphi(f(x)) + \mu(\phi(x)) \\ \Rightarrow |S(x)| &\leq |\varphi(f(x))| + |\mu(\phi(x))|. \end{aligned}$$

Thus, if we show

$$\lim_{x\to a} \frac{|\varphi(f(x))|}{|x-a|} = 0 \text{ and } \lim_{x\to a} \frac{|\mu(\phi(x))|}{|x-a|} = 0.$$

Then we are done.

$|\mu(\phi(x))| \leq M|\phi(x)|$, for some M.

$$\lim_{x \to a} \frac{M|\phi(x)|}{|x-a|} = 0 \quad \texttt{by } (4). \Rightarrow \lim_{x \to a} \frac{|\mu(\phi(x))|}{|x-a|} = 0.$$

Now we have

$$\lim_{f(x) \to b} \frac{|\varphi(f(x))|}{|f(x)-b|} = 0 \texttt{ by } (5).$$

$\Rightarrow \forall \varepsilon > 0$ there exists $\delta_1 > 0$ such that

$$|f(x) - b| < \delta_1 \Rightarrow |\varphi(f(x))| \leq \varepsilon |f(x) - b|$$

and $|f(x) - b| < \delta_1$ and f is continuous at $a \Rightarrow$ there exists $\delta > 0, |x - a| < \delta \Rightarrow |f(x) - b| < \delta_1$.

Now $|x - a| < \delta \Rightarrow |f(x) - b| < \delta_1 \Rightarrow$

$$\begin{aligned} |\varphi(f(x)) &\leq \varepsilon|f(x) - b| = \varepsilon|\phi(x) + \lambda(x-a)| \\ |\varphi(f(x))| &\leq \varepsilon|\phi(x)| + \varepsilon|\lambda(x-a)| \\ &\leq \varepsilon|\phi(x)| + \varepsilon M|x-a| \texttt{ for some } M > 0. \end{aligned}$$

Thus now $\forall \varepsilon > 0$ there exists $\delta > 0$ such that

$$\begin{aligned} |x-a| &< \delta \Rightarrow |\varphi(f(x))| \leq \varepsilon(|\phi(x)| - M|x-a|) \\ &\Rightarrow \lim_{x \to a} \frac{|\varphi(f(x))}{|x-a|} = 0. \end{aligned}$$

Hence since $\frac{|S(x)|}{|x-a|} \leq \frac{|\varphi(f(x))|}{|x-a|} + \frac{|\mu(\phi(x))|}{|x-a|}$

$$\Rightarrow \lim_{x \to a} \frac{S(x)|}{|x-a|} = 0.$$

Example 5.2.1. Let $A \subseteq R^n$. If $f : A \to R^m$ is a constant function i.e., for some $c \in R^m$, $f(x) = c$ for some $x \in A$. Then $Df(a) = 0$.

Solution.

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - 0|}{|h|} = 0 \Rightarrow Df(a) = 0.$$

Example 5.2.2. If $f : R^n \to R^m$ is a linear transformation, then $Df(a) = f$.

Solution.

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

$\Rightarrow Df(a)=f.$

Example 5.2.3. Let $A\subseteq R^n$ be open, $f:A\to R^m$ is differentiable at $a\in A$ if and only if each f^i is differentiable and

$$Df(a)=(Df^1(a),\ldots,Df^m(a)).$$

Solution. Suppose f is differentiable at a. Then since π^i (the projection function) is differentiable with $\lambda=\pi^i$. i.e., $D\pi^i(a)=\pi^i$ and f is differentiable at a, the composition $\pi^i of$ is differentiable at a (Chain Rule) and

$$\begin{aligned} D(\pi^i of)(a) &= Df(a)D(\pi^i f(a)) = D\pi^i(f(a)oDf(a)) = \pi^i.Df(a) \\ &\Rightarrow Df^i(a)=Df(a) \texttt{ for each } i. \end{aligned}$$

Conversely, suppose each f^i be differentiable at a.

Let $\lambda=(Df^1(a),Df^2(a),\ldots,Df^m(a)),\lambda^i=Df^i(a)$.
Consider $\lim_{h\to 0}\frac{|f(a+h)-f(a)-\lambda(h)|}{|h|}$

$$\begin{aligned} & |f(a+h)-f(a)-\lambda(h)| \\ = & |f^1(a+h)-f^1(a)-\lambda^1(h),\ldots,f^m(a+h)-f^m(a)-\lambda^m(h)| \\ \leq & \sum_{i=1}^{m}|f^i(a+h)-f^i(a)-\lambda^i(h)| \end{aligned}$$

but

$$\lim_{h\to 0}\left[\frac{1}{|h|}\sum_{i=1}^{n}|f^i(a+h)-f^i(a)-\lambda^i(h)|\right]=0$$

and

$$\begin{aligned} 0 \quad \leq \quad & \frac{|f(a+h)-f(a)-\lambda(h)|}{|h|}\leq\frac{1}{|h|}\sum_{i=1}^{n}|f^i(a+h)-f^i(a)-\lambda^i(h)| \\ \Rightarrow \quad & \lim_{h\to 0}\frac{|f(a+h)-f(a)-\lambda h|}{|h|}=0. \end{aligned}$$

i.e., $f'(a) = Df(a) = (Df^1(a).Df^2(a), \ldots, D^m(a)) = \lambda$.

Example 5.2.4. $S : R^2 \to R$ defined by $S(x, y) = x + y$ is differentiable with $DS(a, b) = S$. i.e., $DS(a, b)(x, y) = S(x, y)$.

Solution. S is linear hence by $(2), S' = S \Rightarrow DS(a, b) = S$.

Example 5.2.5. $P : R^2 \to R$ be defined by $P(x, y) = xy$, then P is differentiable and $DP(a, b)(x, y) = bx + ay$.

Then

$$P'(a, b) = (b, a).$$

Solution. $P(x, y) = xy$. Let $\lambda(x, y) = bx + ay, \lambda = DP(a, b)$.

Then

$$\frac{|P(a+h, b+k) - P(a, b) - \lambda(h, k)|}{|(h, k)|}$$

$$= \frac{|(a+h)(b+k) - ab - bh - ak|}{|(h, k)|} = \frac{|hk|}{\sqrt{h^2+k^2}} \leq \frac{h^2+k^2}{\sqrt{h^2+k^2}},$$

$$\therefore \quad \lim_{(h,k)\to(0,0)} \frac{|P(a+h, b+k) - P(a, b) - \lambda(h, k)|}{|(h, k)|} = 0 \Rightarrow \lambda(x, y) = bx + ay.$$

Thus it follows that the matrix representation of the linear transformation $f'(a) = (b, a)$.

Remark 5.2.1. Let $f : A \to R^m$ where A is an open set in R^n. If f is differentiable at $a \in A$. Then the matrix representation of $Df(a)$ with respect to the standard basis in R^n and R^m is called the Jacobian matrix of f at a and is denoted by $f'(a)$.

Corollary 5.2.1. If $f : A \to R^n$ is differentiable at $a \in A \subseteq R^n$, A open, then $f'(a)$ is the $m \times n$ matrix whose i^{th} row is $(f^i)^1(a)$.

Proof. We know that if each f^i is differentiable at a then

$$Df(a) = (Df^1(a), \ldots, Df^m(a)).$$

Suppose $(f^i)^1(a) = (a_{i1}, \ldots, a_{in})$ for $i = 1, \ldots, m$. Then $(f^i)^1(e_j) = a_{ij}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Therefore

$$\begin{aligned} Df(a)(e_j) &= Df^1(a)e_j, \ldots, Df^m(a)e_j, j = 1, \ldots, n \\ &= (f^1)^1(a)(e_j), \ldots, (f^m)^1(a)(e_j), j = 1, \ldots, n \\ &= (a_{ij}, \ldots, a_{mj}), j = 1, \ldots, n \\ &= \sum_{i=1}^{m} a_{ij}e_i, j = 1, \ldots, n. \end{aligned}$$

$$\Rightarrow f^{'}(a) = \begin{pmatrix} a_{11} & \ldots a_{1n} \\ \vdots & \\ a_{m1} & \ldots a_{mn} \end{pmatrix} \texttt{ or } f^{'}(a) = \begin{pmatrix} (f^1)^1(a) \\ \vdots \\ (f^m)^1(a) \end{pmatrix}.$$

Example 5.2.6. $P : R^2 \to R, \quad P(x, y) = xy$, then

$$\begin{aligned} &DP(a, b) = \lambda \texttt{ and } \lambda(x, y) = bx + ay \\ &DP(a) = Df^{'}(a) \\ &DP(a, b)(x, y) = (b, a)\begin{pmatrix} x \\ y \end{pmatrix} \\ &P^{'}(a, b) = (b, a). \end{aligned}$$

Example 5.2.7. $f(x, y) = \sin x, \quad \lambda(x, y) = \cos a.x$, then

$$\begin{aligned} &Df(a, b) = \lambda \\ &Df(a, b)(x, y) = \cos a.x + 0.y \\ &f^{'}(a, b) = (\cos a, 0). \end{aligned}$$

Example 5.2.8. $f(x, y, z) = (2x + z, y + 3z), \quad f : R^3 \to R^2$
f is linear, $\lambda = f$, then

$$\begin{aligned} Df(a) &= (Df^1(a), Df^2(a)) \\ Df^1(a) &= f^1, \quad Df^2(a) = f^2 \\ Df^1(a)(x, y, z) &= 2x + z \\ Df^2(a)(x, y, z) &= y + 3z \\ Df(a)(x, y, z) &= \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= (2x + z, y + 3z) \end{aligned}$$

$$\begin{aligned} f^1(a) &= \begin{pmatrix} (f^1)^1(a) \\ (f^2)^1(a) \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & 3 \end{pmatrix}. \end{aligned}$$

Corollary 5.2.2. Let $f, g : R^n \to R$ be differentiable at $a \in R^n$, then $f + g, f.g$ and f/g for $g(a) \neq 0$ are all differentiable at a. Furthermore

i) $D(f+g)(a) = Df(a) + Dg(a)$.

ii) $D(f.g)(a) = g(a)Df(a) + f(a)Dg(a)$.

iii) $D(f/g)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{[g(a)]^2}$.

Proof.

i) $(f+g)(x) = So(f,g)(x)$

$$\begin{aligned} D(f+g)(a) &= D(So(f,g)(a)) \\ &= DS(f(a),g(a))oD(f(a),g(a)) \\ &= So(Df(a),Dg(a)) \\ &= Df(a) + Dg(a). \end{aligned}$$

ii) $(f.g)(a) = po(f,g)(a)$

$$\begin{aligned} D((f.g)(a)) &= D[po(f,g)(a)] \\ &= Dp(f(a),g(a))o(Df(a),Dg(a)) \\ &= g(a)Df(a) + f(a)Dg(a). \end{aligned}$$

iii) First show that $D\left(\frac{1}{g}\right)(a) = -\frac{Dg(a)}{[g(a)]^2}$, then apply (ii).

To show $D\left(\frac{1}{g}\right)(a) = -\frac{Dg(a)}{[g(a)]^2}$, by direct computation we have

$$\begin{aligned}
& \left|\frac{1}{g(a+h)} - \frac{1}{g(a)} + \frac{Dg(a)}{[g(a)]^2}(h)\right| \\
= & \left|\frac{1}{g(a+h)} - \frac{1}{g(a)} + \frac{Dg(a)(h)}{[g(a)]^2}\right| \\
= & \left|\frac{g(a) - g(a+h)}{g(a)g(a+h)} + \frac{Dg(a)(h)}{[g(a)]^2}\right| \\
= & \left|\frac{g(a) - g(a+h) - Dg(a)(h) + Dg(a)(h)}{g(a)g(a+h)} + \frac{Dg(a)(h)}{[g(a)]^2}\right| \\
= & \left|\frac{-[g(a+h) - g(a) - Dg(a)(h)]}{g(a)g(a+h)}\right. \\
& \left. - \frac{Dg(a)(h)}{g(a)g(a+h)} + \frac{Dg(a)(h)}{[g(a)]^2}\right| \\
= & \left|\frac{-[g(a+h) - g(a) - Dg(a)(h)]}{g(a)g(a+h)}\right. \\
& \left. + \frac{Dg(a)(h)[[g(a)] - g(a+h)]}{[g(a)]^2 g(a+h)}\right| \\
\leq & \frac{|g(a+h) - g(a) - Dg(a)(h)|}{|g(a)g(a+h)|} + \frac{M|h||g(a) - g(a+h)|}{[g(a)]^2|g(a+h)|}
\end{aligned}$$

$$\begin{aligned}
& \lim_{h\to 0} \frac{|g(a+h) - g(a) - Dg(a)(h)|}{|h|} . \frac{1}{|g(a)g(a+h)|} \\
+ & \frac{M|g(a) - g(a+h)|}{[g(a)]^2|g(a+h)|} = 0 \\
\Rightarrow & \lim_{h\to 0} \frac{1}{|h|}\left|\left(\frac{1}{g}\right)(a+h) - \frac{1}{g}(a) + \left(\frac{Dg(a)}{[g(a)]^2}\right)(h)\right| = 0 \\
\Rightarrow & D\left(\frac{1}{g}\right)(h) = \frac{-Dg(a)}{[g(a)]^2}.
\end{aligned}$$

$$\begin{aligned}
D\left(\frac{f}{g}\right)(a) &= \left(\frac{1}{g}\right)(a)Df(a) + f(a)D\left(\frac{1}{g}(a)\right) \qquad by (ii) \\
&= \frac{1}{g(a)}Df(a) - \frac{f(a)Dg(a)}{[g(a)]^2} \\
&= \frac{g(a)Df(a) - f(a)Dg(a)}{[g(a)]^2}.
\end{aligned}$$

Remark 5.2.2. $f : A \to R^m, A \subseteq R^n, g : B \to R^k, B \subseteq R^m f$ is differentiable at a and g is differentiable at $f(a)$ then by Chain Rule

$$D(g0f)(a) = Dg(f(a))oDf(a).$$

In terms of the Jacobian matrix this equation can be written as

$$(gof)^{'}(a) = g^{'}(f(a)).f^{'}(a).$$

Remark 5.2.3. $\pi^i : R^n \to R$ is projection $D\pi^i(a) = \pi^i$ and $(\pi^i)^{'} = (0, \ldots, 1, 0 \ldots 0)$.

Example 5.2.9. $f(x, y) = \sin(xy^2)$, find $f^{'}(a, b)$ which is a 1×2 matrix.

Solution. $f = \sin o(\pi^1(\pi^2)^2)$.

$$\begin{aligned} f^{'}(a,b) &= [\sin' \pi^1(\pi^2)^2(a,b)].[(\pi^1)'(\pi^2)^2 + \pi^1[((\pi^2)^2)^{'}](a,b)] \\ &= \cos ab^2[(\pi^1)'(a.b)(\pi^2)^2(a,b) + \pi^{(}a,b)((\pi^2)^2)^{'}(a,b)] \\ &= \cos ab^2[(1,0)b^2 + a[2\pi^2(\pi^2)^{'}(a,b)]] \\ &= \cos ab^2[(b^2,0) + a(2b(0,1))] \\ &= \cos ab^2(b^2, 2ab) = (b^2\cos(ab^2), 2ab\cos(ab^2)). \end{aligned}$$

Exercises 5.2

1. Find derivatives of the following:
   (a) $f(x, y, z) = (x^y, z)$.
   (b) $f(x, y, z) = \sin(x \sin y)$.
   (c) $f(x, y, z) = x^{y+z}$.
   (d) $f(x, y) = (\sin(xy), \sin(x \sin y), x^y)$.

2. Let $f : R^3 \to R$ and $g : R^2 \to R$ be differentiable. Let $F : R^2 \to R$ be defined by the equation

$$F(x, y) = f(x, y, g(x, y)).$$

   (a) Find DF in terms of the partials of f and g.

(b) If $F(x, y) = 0$ for all (x, y), find $D_1 g$ and $D_2 g$ in terms of the partials of f.

3. Let $f : R^2 \to R^3$ and $g : R^3 \to R^2$ be given by the equations

$$\begin{aligned} f(x) &= (e^{2x_1+x_2}, 3x_2 - \cos x_1, x_1^2 + x_2 + 2), \\ g(y) &= (3y_1 + 2y_2 + y_3^2, y_1^2 - y_3 + 1). \end{aligned}$$

(a) If $F(x) = g(f(x))$, find $DF(0)$.

(b) If $G(y) = f(g(y))$, find $DG(0)$.

4. Suppose that the function $h : R^3 \to R$ is continuously differentiable. Define the function $\eta : R^3 \to R$ by

$$\eta(u, \nu, w) = (3u + 2\nu)h(u^2, \nu^2, u\nu w) \ \texttt{for} \ (u, \nu, w) \ \texttt{in} \ R^3.$$

Find $D_1\eta(u, \nu, w), D_2\eta(u, \nu, w)$ and $D_3\eta(u, \nu, w)$.

5. Suppose that the functions $g : R \to R$ and $h : R \to R$ have continuous second-order partial derivatives. Define the function $u : R^2 \to R$ by

$$u(s, t) = g(s - t) + h(s + t) \ \texttt{for} \ (s, t) \ \texttt{in} \ R^2.$$

Prove that $\frac{\partial^2 u}{\partial t^2}(s, t) - \frac{\partial^2 u}{\partial s^2}(s, t) = 0$ for all (s, t) in R^2.

6. Suppose that the functions $f : R \to R$ and $g : R \to R$ have continuous second-order partial derivatives. Also that there is a number λ such that

$$f''(x) = \lambda f(x) \ \texttt{and} \ g''(x) = \lambda g(x) \ \texttt{for all} \ x \ \texttt{in} \ R.$$

Define the function $u : R^2 \to R$ by

$$u(x, y) = f(x)g(y) \ \texttt{for} \ (x, y) \ \texttt{in} \ R^2.$$

Prove that $\frac{\partial^2 u}{\partial x^2}(x, y) - \frac{\partial^2 u}{\partial y^2}(x, y) = 0$ for every (x, y) in R^2.

7. Suppose that the function $g : R^n \to R$ is continuously differentiable. For points x and p in R^n, if $\varphi(t) = g(x + tp)$ for t in R, then

$$\varphi'(t) = < Dg(x + tp), p > \ \texttt{for every} \ t \ \texttt{in} \ R.$$

Show that this formula is a special case of the Chain Rule.

8. Suppose that the function $u : R^2 \to R$ is harmonic. Let a, b, c and d be real numbers such that

$$a^2 + b^2 = 1, c^2 + d^2 = 1 \texttt{ and } ac + bd = 0.$$

Define the function $\nu : R^2 \to R$ by

$$\nu(x, y) = u(ax + by, cx + dy) \texttt{ for } (x, y) \texttt{ in } R^2.$$

Prove that the function $\nu : R^2 \to R$ is also harmonic.

5.3 Partial Derivative

Definition 5.3.1. Let $A \subset R^n$ be open. If $f : A \to R$ and $a \in A$, then the limit

$$\lim_{h \to 0} \frac{f(a^1, \ldots, a^i + h, \ldots, a^n) - f(a^1, \ldots, a^n)}{h}$$

if it exists, is denoted by $D_i f(a)$, and called the i^{th} partial derivative of f at a.

$$D_i f(a) = \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h}.$$

Suppose $f : R^n \to R^m$ is a given function.

$$f(x) = (y_1, y_2, \ldots, y_m) \in R^m.$$

Components of $f : f : R^n \to R \quad i = 1, 2, \ldots, m,$

$$f(x) = (f_1(x), f_2(x), \ldots, f_m(x)).$$

Conversely if m functions $\phi_j : R^n \to R$ are defined then $f : (\phi_1, \phi_2, \ldots, \phi_m)$ is a function from $R^n \to R^m$.

Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis of R^n. We define for $f = (f_1, f_2, \ldots, f_m)$ the function $D_j f_i$ as:

Let $E \subseteq R^n$ be an open set and $x \in E$.

$$(D_j f_i)(x) = \lim_{h \to 0} \frac{f_i(x + he_j) - f_i(x)}{h} \texttt{ provided the limit exists.}$$

or

$$(D_j f_i)x = \frac{\partial f_i}{\partial x_j}.$$

Suppose $f : R^n \to R^m$ is differentiable at $x \in E$,

$$\lim_{t\to 0} \frac{|f(x+t) - f(x) - f'(x)t|}{|t|} = 0.$$

Let $t = he_j$, then we have

$$\lim_{h\to 0} \frac{f(x+he_j) - f(x)}{h} = f'(x)(e_j).$$

Let us put the components of f in numerator of left hand side

$$\lim_{h\to 0} \frac{f_1(x+he_j), f_2(x+he_j), \ldots, f_m(x+he_j) - f_1(x), f_2(x), \ldots, f_m(x)}{h} = f'(x)(e_j)$$

i.e.,

$$\lim_{h\to 0} \frac{f_1(x+he_j) - f_1(x), f_2(x+he_j) - f_2(x), \ldots, f_m(x+he_j) - f_m(x)}{h}.$$

Hence the individual limits

$$\lim_{h\to 0} \frac{f_i(x+he_j) - f_i(x)}{h}, i = 1, 2, \ldots, m.$$

etc. exists.

Thus if f is differentiable at $x \in E$ then each component is also differentiable.

But converse is not true:

$$f(0,0) = 0. \quad f(x,y) = \frac{x^3}{x^2+y^2},$$

function being ratio of polynomial is continuous on R^2 at $(0,0)$

$$D_1 f : \lim_{x\to 0} \frac{\frac{x^3}{x^2+y^2} - 0}{x} = \lim_{x\to 0} \frac{x^3}{x^2 - x} = 1.$$

$$D_2 f : \lim_{y\to 0} \frac{\frac{0}{y^2}}{y} = 0.$$

at $y = mx$

$$\frac{x^3}{\frac{x^2(1+m^2)}{x}} = \frac{1}{1+m^2} \Rightarrow$$

limit is not unique.

Example 5.3.1. $f(x,y,z) = z\sin(xy^2)$. Then

$$\begin{aligned}
D_1 f(x,y,z) &= z\cos(xy^2).y^2 = zy^2\cos(xy^2)\\
D_2 f(x,y,z) &= 2zyx\cos(xy^2)\\
D_3 f(x,y,z) &= \sin(xy^2).
\end{aligned}$$

Remark 5.3.1. If $D_j f(x)$ exists $\forall x \in A \subseteq R^n$ we obtain a fix $D_j f : A \to R$. Then i^{th} partial derivative of this function at $x \in A$, is denoted by $D_{i,j} f(x)$. $D_{i,j} f(x)$ is called the second order partial derivatives of f.

Remark 5.3.2. If $D_{i,j} f$ and $D_{j,i} f$ are continuous in an open set contain a, then

$$D_{i,j} f(a) = D_{j,i} f(a).$$

Lemma 5.3.1. Let $A \subset R^n$ be open and $f : A \to R$ be differentiable at $a \in A$, then the Jacobian matrix of f at a is

$$f'(a) = (D_1 f(a), D_2 f(a), \ldots, D_n f(a)).$$

Proof. Consider the function $h : R \to A \subset R^n$ defined by

$$h(x) = (a^1, \ldots, a^{j-1}, x, a^{j+1}, \ldots, a^n).$$

h is differentiable at a^i (infect at any point t)

$$\begin{aligned}
& \lim_{k\to 0} \frac{|h(t+k) - h(t) - \lambda(k)|}{k}\\
= \;& \lim_{k\to 0} \frac{|(0,\ldots,0,k,0,\ldots,0) - \lambda(k)|}{k} = 0\\
\Rightarrow \;& \lambda(k) = (0,\ldots,k,\ldots,0) \qquad \lambda : R \to R^n
\end{aligned}$$

is a linear transformation.

Thus h is differentiable with

$$h^{'}(a_j.) = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} j^{th} \texttt{ place} \ .$$

Now consider the function foh, since h is differentiable at a_j and f is differentiable at $a = h(a_j)$, foh is differentiable at $a_j^{'}$ and by Chain Rule

$$(foh)^{'}(a_j^{'}) = f^{'}(h(a_j))h^{'}(a_j)$$

$$\Rightarrow D_jf(a) = f^{'}(a). \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \end{pmatrix} j^{th} \texttt{ place.}$$

i.e., $D_jf(a)$ is the j^{th} component of $f^{'}(a)$.

$$\begin{aligned} \Rightarrow f^{'}(a) &= (D_1f(a), \ldots, D_nf(a)). \\ (foh)^{'}(a_j) &= \lim_{k\to 0} \frac{(foh)(a_j + k) - (foh)(a_j)}{k} \\ &= \lim_{k\to 0} \frac{f(a_1, \ldots, a_j + k, \ldots, a_n) - f(a_1, \ldots, a_n)}{k} \\ &= D_jf(a). \end{aligned}$$

Theorem 5.3.1. Let $A \subseteq R^n$ be open set and $f : A \to R^m$ be differentiable at $a \in A$. Then $D_jf^i(a)$ exists for each $i = 1, \ldots, m$, and $j = 1, \ldots, n$. Furthermore the Jacobian of f at a $f^{'}(a)$ is given by

$$f^{'}(a) = \begin{pmatrix} D_1f^1(a) & D_2f^1(a) & \ldots & D_nf^1(a) \\ \vdots & & & \\ D_1f^m(a) & D_2f^m(a) & \ldots & D_nf^m(a) \end{pmatrix}.$$

Proof. Let $f = (f^1, f^2, \ldots, f^m)$, differentiable at a, then for each $i = 1, \ldots, m$, f^i is differentiable at a.

Furthermore

$$f^1(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \\ \vdots \\ (f^m)'(a) \end{pmatrix} f^i : A \to R \texttt{ for each } i.$$

Thus by Lemma 5.3.1 $(f^i)'(a) = (D_1 f^i(a), D_2 f^i(a), \ldots, D_n f^i(a))$

$$\Rightarrow f^i(a) = (D_i f^j(a))_{m\times n}.$$

Note. Existence of partial derivatives need not imply differentiability.

Example 5.3.2. Let

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

$D_1 f(0,0) = 0$, $D_2 f(0,0) = 0$, but f is not even continuous at $(0,0)$.

Example 5.3.3. Find the Jacobian matrix for each of the following.

a) $f(x,y,z) = (x^y, z)$.

b) $f(x,y,z) = \sin(x \sin y)$.

c) $f(x,y,z) = x^{y+z}$.

d) $f(x,y) = (x\sin(x,y), \sin(x \sin y), x^y)$.

Solution.

a)

$$\begin{aligned} f(x,y,z) &= (x^y, z) \\ f &= (f^1, f^2) \\ f'(a) &= \begin{pmatrix} D_1 f^1(a), D_2 f^1(a) D_3 f^1(a) \\ D_1 f^2(a), D_2 f^2(a) D_3 f^2(a) \end{pmatrix} \\ &= \begin{pmatrix} yx^{y-1} & x^y \log x & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

b) It can be find similarly.

c) It can also be find similarly.

d)

$$\begin{aligned} f(x,y) &= (x\sin(xy), \sin(x\sin y), x^y), \\ f &= (f_1, f_2, f_3) \quad f_i = f_i(x,y), \\ f^1(a) &= x\sin(xy) \\ f^2(a) &= \sin(x\sin y) \\ f^3(a) &= x^y. \\ f^{'}(a) &= \begin{pmatrix} D_1f^1(a) & D_2f^1(a) \\ D_1f^2(a) & D_2f^2(a) \\ D_1f^3(a) & D_2f^3(a) \end{pmatrix} \\ &= \begin{pmatrix} \sin xy + xy\sin(xy) & x^2\cos xy \\ \cos(x\sin y)(\sin xy) & \cos(x\sin y)xcosy \\ yx^{y-1} & x^y lnx \end{pmatrix}. \end{aligned}$$

Theorem 5.3.2. Let $A \subseteq R^n$ be open and $f : A \to R^m$. If all partial derivatives $D_jf^i(x)$ exists in some open set containing a and if each function D_jf^i is continuous at a, then f is differentiable at a. (Such function f is called continuously differentiable at a).

Proof. First consider the case $m = 1$. Then $f : A \to R, A \subseteq R^n$ we have to show that

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = 0$$

for some linear transformation $\lambda : R^n \to R$. Note: $a = (a^1, \ldots, a^n)$ $h = (h^1, \ldots, h^n)$.

Now $f(a+h) - f(a) = f(a+h^1e_1) - f(a) + f(a+h^1e_1+h^2e_2) - f(a+h^1e_1) + f(a+h^1e_1+\cdots+h^ie_i) - f(a+h^1e_1+\cdots+h^{i-1}e_{i-1}) + f(a+h) - f(a+h^1e_1+\cdots+h^{n-1}e_{n-1})$.

But for each $i, D_if(x)$ exists in some open set containing a then by M.V.T. there exists t_1 between a^1 and a^1+h^1 such that $D_1f(t_1)h_1^1 = f(a+h^1e_1) - f(a)$. Thus in general for each $i = 1, \ldots, n$, there exists t_i between a^i and a^i+h^i such that

$$D_if(t_i)h^i = f(a+h^1e_1+\cdots+h^ie_i) - f(a+h^1e_1+\cdots+h^{i-1}e_{i-1}).$$

Thus now

$$f(a+h) - f(a) = \sum_{i=1}^{n} D_i f(t_i) h^i.$$

$$\begin{aligned}
& \lim_{h\to 0} \frac{|f(a+h) - f(a) - \sum_{i=1}^{n} D_i f(a) h^i|}{|h|} \\
= & \lim_{h\to 0} \frac{|\sum_{i=1}^{n} D_i f(t_i) h^i - \sum_{i=1}^{n} D_i f(a) h^i|}{|h|} \\
= & \lim_{h\to 0} \frac{|\sum_{i=1}^{n} D_i f(t_i) - D_i f(a) h^i|}{|h|} \\
= & \lim_{h\to 0} \frac{|\sum_{i=1}^{n} (D_i f(t_i) - D_i f(a))(h^i)|}{|h|} \\
\leq & \lim_{h\to 0} \frac{\sum_{i=1}^{n} |D_i f(t_i) - D_i f(a)||h^i|}{|h|} \\
\leq & \lim_{h\to 0} \sum_{i=1}^{n} |D_i f(t_i) - D_i f(a)|.
\end{aligned}$$

Since $D_i f$ is continuous at $a \Rightarrow \lim_{h\to 0} D_i f(t_i) = D_i f(a)$. Therefore $\lim_{h\to 0} \frac{|f(a+h)-f(a)-\sum_{i=1}^{n} D_i f(a) h^i}{|h|} = 0 \Rightarrow f$ is differentiable at a.

Remark 5.3.2. Suppose now $f : A \to R^m$ is in above theorem and $m > 1$, let $f = (f^1, f^2, \ldots, f^m)$, then for each $i = 1, \ldots, n$, $D_i f^i(a)$ exists in some open interval containing a and $D_i f^i$ is continuous at $a \Rightarrow f^i$ for each $i = 1, \ldots, m$ is differentiable at a.

Then it follows by theorem that f is differentiable.

Theorem 5.3.3. Let $g_1, \ldots, g_m : R^n \to R$ be continuously differentiable at a and let $f : R^m \to R$ be continuously differentiable at $(g_1(a), \ldots, g_m(a))$. Define $F : R^m \to R$ by

$$\begin{aligned}
F(x) &= f(g_1(x), \ldots, g_m(x)). \text{ Then} \\
D_i F(a) &= \sum_{j=1}^{m} D_j f(g_1(a), \ldots, g_m(a)).D_i g_j(a).
\end{aligned}$$

Proof. $f : R^m \to R$ is continuously differentiable at $(g_1(a), \ldots, g_m(a))$ $\Rightarrow f$ is differentiable at $(g_1(a), \ldots, g_m(a))$. In each $i = 1, \ldots, m$, $g_i :$

$R^n \to R$ is continuously differentiable at $a \Rightarrow g_i$ is differentiable at a. Thus by Chain Rule

$F = fog$ where $g : R^n \to R^m, g(a) = (g_1(a), \ldots, g_m(a))$ is differentiable at a

$$F^{'}(a) = (D_1F(a), D_2F(a), \ldots, D_nF(a)$$

and

$$\begin{aligned} F^{'}(a) &= f^{'}(g(a)).g^{'}(a) \\ &= (D_1f(g(a)), D_2f(g(a)), \ldots, D_nf(g(a)) \\ &\quad . \begin{pmatrix} D_1g_1(a) & \ldots & D_ng_1(a) \\ D_1g_2(a) & \ldots & D_ng_2(a) \\ \vdots & & \\ D_1g_m(a) & \ldots & D_ng_m(a) \end{pmatrix} \\ &= \left(\sum_{i=1}^{m} D_if(g(a))D_1g_i(a), \sum_{i=1}^{m} D_if(g(a))D_2g_i(a), \ldots, \right. \\ &\quad \left. \sum_{i=1}^{m} D_if(g(a))D_ng_i(a)\right). \end{aligned}$$

But $F^{'}(a) = (D_1F(a), \ldots, D_nF(a))$.
So

$$D_jF(a) = \sum_{i=1}^{m} D_if(g(a))D_jg_i(a).$$

Remark 5.3.3. Above theorem called weak Chain Rule because it is weaker than Chain Rule. gof could be differentiable without g_i or f being continuously differentiable.

Example 5.3.4. Let $F : R^2 \to R$ be given by

$$F(x, y) = f(g(x, y), h(x), k(y)).$$

Determine $D_1F(x, y)$ and $D_2F(x, y)$.

Solution. In order to apply Theorem 5.3.3, we need slight modification of $h, k : R \to R$. Define $\overline{h}, \overline{k} : R^2 \to R$ by $\overline{h}(x, y) = h(x)$

and $\overline{k}(x,y) = k(y)$. Then

$$\begin{aligned} D_1\overline{h}(x,y) &= h^{'}(x) \text{ and } D_2\overline{h}(x,y) = 0 \\ D_1\overline{k}(x,y) &= 0 \text{ and } D_2\overline{k}(x,y) = k^{'}(y). \end{aligned}$$

Thus now $F(x,y) = f(g(x,y), \overline{h}(x,y), \overline{k}(x,y))$ and letting $b = (g(x,y), h(x), k(y))$.

$$\begin{aligned} D_1F(x,y) &= D_1f(b).D_1g_1(x,y) + D_2f(b).D_1g_2(x,y) \\ &\quad +D_3f(b)D_1g_3(x,y) \\ &= D_1f(b)D_1g_1(x,y) + D_2f(b).D_1\overline{h}^{'}(x,y) + 0 \\ D_2F(x,y) &= D_1f(b)D_2g_1(x,y) + D_3f(b).D_2g_3(x,y) \\ &= D_1f(b)D_2g_(x,y) + D_3f(b)D_2\overline{k}^{'}(x,y). \end{aligned}$$

Example 5.3.5. $F(x,y,z) = 2x + xy^2 + zy, \quad x = v + u, y = \sin v, z = u + 2v$. Find $D_1F(x,y,z)$, $D_2F(x,y,z)$,and $D_3F(x,y,z)$.

Solution. Let $f(x,y,z) = x + y + z$,

$$\begin{aligned} g_1(x,y,z) &= 2x, g_2(x,y,z) = xy^2, g_3(x,y,z) = zy. \\ D_1F(x,y,z) &= \sum_{j=1}^{3} D_jf(g_1,g_2,g_3)D_1g_j(x,y,z) \\ &= D_1f(g_1,g_2,g_3)D_1g_1(x,y,z) + D_2f(g_1,g_2,g_3)D_1g_2 \\ &\quad +D_3f(g_1,g_2,g_3)D_1g_3 \\ &= 1.2 + 1.y^2 + 1.0 = 2 + y^2. \\ D_2F(x,y,z) &= 2xy + z, D_3F(x,y,z) = y. \end{aligned}$$

Example 5.3.6. Let $F(x,y) = \sin(x^2y + x) + \cos(x + y^2)$,

$$\begin{aligned} g : R^2 \to R^2, \quad & g(x,y) = (g_1(x,y), g_2(x,y)), \\ & g_1(x,y) = x^2y + x, g_2(x,y) = x + y^2. \end{aligned}$$

$$\begin{aligned} & f : R^2 \to R, f(x,y) = \sin x + \cos y, \\ & f(g_1(x,y), g_2(x,y)) = \sin g_1(x,y) + \cos g_2(x,y), \end{aligned}$$

$$F = fog,$$

$$
\begin{aligned}
D_1F(x,y) &= \sum_{i=1}^{2} D_i f(g_1,g_2).D_1 g_i(x,y)\\
&= D_1 f(g_1,g_2)D_1g_1(x,y) + D_2 f(g_1,g_2)D_1g_2(x,y)\\
&= \cos g_1(2xy+1) - \sin g_2(1)\\
&= \cos(x^2y+x)(2xy+1) - \sin(x+y^2).\\
D_2F(x,y) &= D_1 f(g_1,g_2)D_2g_1(x,y) + D_1(f(g_1,g_2)D_2g_2(x,y))\\
&= \cos g_1(x^2) - \sin g_2.2y\\
&= x^2\cos(x^2y+x) - 2y\sin(x+y^2).
\end{aligned}
$$

Exercises 5.3

1. For the function $f : R^2 \to R$ defined by

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

show that neither the function $\partial f/\partial x : R^2 \to R$ nor the function $\partial f/\partial y : R^2 \to R$ is continuous at the point $(0,0)$.

2. Suppose that the function $g : R^2 \to R$ has the property that

$$
|g(x,y)| \leq x^2 + y^2 \texttt{ for all } (x,y) \texttt{ in } R^2.
$$

Prove that $g : R^2 \to R$ has partial derivatives with respect to both x and y at the point $(0,0)$.

3. Given a pair of functions $\phi : R^2 \to R$ and $\varphi : R^2 \to R$, it is often useful to know whether there exists some continuously differentiable function $f : R^2 \to R$ such that

$$
\frac{\partial f}{\partial x}(x,y) = \phi(x,y) \texttt{ and } \frac{\partial f}{\partial y}(x,y) = \varphi(x,y) \texttt{ for}(x,y) \in R^2.
$$

Such a function $f : R^2 \to R$ is called a potential function for the pair of functions (ϕ, φ).

(i) Show that if a potential function exists for the pair (ϕ,φ), then this potential is uniquely determined up to an additive constant that is, the difference of any two potentials is constant.

(ii) Show that if there is a potential function for the pair of continuously differentiable functions $\phi : R^2 \to R$ and $\varphi : R^2 \to R$, then

$$\frac{\partial \varphi}{\partial x}(x, y) = \frac{\partial \phi}{\partial y}(x, y) \texttt{ for all } (x, y) \texttt{ in } R^2.$$

4. Suppose that the function $f : R^2 \to R$ has continuous second order partial derivatives and let (x_0, y_0) be a point in R^2. Prove that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that if $0 < |h| < \delta$ and $0 < |k| < \delta$, then

$$\left| \frac{f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0)}{hk} - \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right| < \varepsilon.$$

5. Define $f : R^2 \to R$ by setting $f(0) = 0$, and $f(x, y) = xy(x^2 - y^2)/(x^2 + y^2)$ if $(x, y) \neq 0$.

   (a) Show that $D_1 f$ and $D_2 f$ exist at 0.

   (b) Calculate $D_1 f$ and $D_2 f$ at $(x, y) \neq 0$.

   (c) Show f is of class C^1 in R^2.

   (d) Show that $D_2 D_1 f$ and $D_1 D_2 f$ exist at 0, but are not equal there.

6. Show that if $A \subset R^m$ and $f : A \to R$, and if the partials $D_j f$ exist and are bounded in a neighborhood of a, then f is continuous at a.

5.4 Directional Derivatives

Definition 5.4.1. Let $S \subseteq R^n$ and $f : S \to R$. Let $a \in S$(S being open) and let y be an arbitrary point in R^n, the derivative of f at a with respect to y is denoted by the symbol $f'(a; y)$ and is defined by the equation

$$f'(a; y) = \lim_{h \to 0} \frac{f(a + hy) - f(a)}{h}$$

when the limit on the right exists.

Example 5.4.1. If $f : S \to R$ is linear, then $f(a+hy) = f(a)+hf(y)$. Then $f(a;y)$ always exists and $f^{'}(a;y) = f(y)$.

Definition 5.4.2. Let $A \subseteq R^n$ be open and $f : A \to R$. Let $\times \in R^n$ be any vector, then directional derivative of f at $a \in A$ in the direction of $\times$ is denoted by $D_{\times} f(a)$ and defined as

$$D_{\times} f(a) = \lim_{t \to 0} \frac{f(a+t\times) - f(a)}{t}$$

provided this limit exist.

Remark 5.4.1. $De_i f(a) = D_i f(a), \quad \times = e_i$

$$De_i f(a) = \lim_{t \to 0} \frac{f(a+te_i) - f(a)}{t} = D_i f(a).$$

Theorem 5.4.1. Let $f : A \to R, A \subseteq R^n$ open. If f is differentiable at $a \in \times$ then $D_{\times} f(a)$ exists in any direction and $D_{\times} f(a) = Df(a)(\times)$.

Proof. Let $\lambda = Df(a)$, then for $\times \neq 0$.

$$\begin{aligned}
& \lim_{t \to 0} \frac{|f(a+t\times) - f(a) - \lambda(t\times)|}{|t\times|} = 0 \\
\Rightarrow \quad & \lim_{t \to 0} \left| \frac{f(a+t\times) - f(a) - \lambda(t\times)}{t\times} \right| = 0 \\
\Leftrightarrow \quad & \lim_{t \to 0} \frac{1}{|\times|} \left| \frac{f(a+t\times) - f(a)}{t} - \lambda(\times) \right| = 0 \\
\Rightarrow \quad & \lim_{t \to 0} \left| \frac{f(a+t\times) - f(a)}{t} - \lambda(\times) \right| = 0 \\
\Rightarrow \quad & \lim_{t \to 0} \frac{f(a+t\times) - f(a)}{t} - \lambda(\times) \\
\Rightarrow \quad & D_{\times} f(a) = Df(a)(\times).
\end{aligned}$$

Note. The converse of the above theorem is not true.

Example 5.4.2. Let $A = \{(x,y) \in R^2 : x > 0 \text{ and } 0 < y < x^2\}$.

Define $f : R^2 \to R$ by

$$f(x) = \begin{cases} 0 & x \notin A \\ 1 & x \in A. \end{cases}$$

f is not continuous at $(0,0)$ and hence f is not differentiable at $(0,0)$.

$D_\nu f(0,0)$ exists $\forall \nu$.

$$\begin{aligned} D_\nu f(0,0) &= \lim_{h \to 0} \frac{f(0,0) + h(\nu_1, \nu_2) - f(0,0)}{h} \\ &= \lim_{h \to 0} \frac{f(h\nu_1, h\nu_2)}{h}. \end{aligned}$$

Now let $\nu = (\nu_1, \nu_2) \in R^2 . \nu \neq \overline{0}$, then there exists $h \in R$ such that $h\nu \notin A$ and for all $t \in R$ and $|t| < |h|, t\nu \notin A$. In this case then

$$D_\nu f(0,0) = \lim_{h \to 0} \frac{f(h\nu_1, h\nu_2)}{h} = 0.$$

i.e., the directional derivative along any vector ν exists.

Corollary 5.4.1. Let A be an open subset of R^n and $a \in A$. Let $x, y \in R^n$

i) If $D_\times f(a)$ exists then for any $r \in R, D_{r\times} f(a)$ exists and $D_{r\times} f(a) = r D_\times f(a)$.

ii) If f is differentiable at a, then

$$D_{x+y} f(a) = D_x f(a) + D_y f(a).$$

Proof.

i) $D_\times f(a)$ exists and for $r \in R$ we have

$$\begin{aligned} D_{r\times} f(a) &= \lim_{h \to 0} \frac{f(a + hr\times) - f(a)}{h} \\ &= r \lim_{h \to 0} \frac{f(a + hr\times) - f(a)}{hr} \\ &= r \lim_{h \to 0} \frac{f(a + t\times) - f(a)}{t}, \qquad t = rh \quad t \to 0 \quad \text{as } rh \to 0 \\ &= r D_\times f(a). \end{aligned}$$

ii) f is differentiable at $a \Rightarrow D_{\times} f(a) = Df(a)(\times)$. Then

$$\begin{aligned} D_{x+y} f(a) &= Df(a)(x+y) \\ &= Df(a)(x) + Df(a)(y) \\ &= D_x f(a) + D_y f(a). \end{aligned}$$

Exercises 5.4

1. Consider the following assertions for a function $f : R^2 \to R$:

   (i) The function $f : R^2 \to R$ is continuously differentiable.

   (ii) The function $f : R^2 \to R$ has directional derivatives in all directions at each point in R^2.

   (iii) The function $f : R^2 \to R$ has first order partial derivatives at each point in R^2.
   Explain the implications between these assertions.

2. Define the function $f : R^2 \to R$ by

$$f(x,y) = \begin{cases} (x/|y|)\sqrt{x^2+y^2} & if \quad y \neq 0 \\ 0 & if \quad y = 0. \end{cases}$$

   (i) Prove that the function $f : R^2 \to R$ is not continuous at the point $(0,0)$.

   (ii) Prove that the function $f : R^2 \to R$ has directional derivatives in all directions at the point $(0,0)$.

   (iii) Prove that if c is any number, then there is a vector P of norm 1 such that

$$\frac{\partial f}{\partial P}(0,0) = c.$$

3. Suppose that the functions $f : R^n \to R$ and $g : R \to R$ are continuously differentiable. Find a formula for $D(gof)(x)$ in terms of $Df(x)$ and $g'(f(x))$.

5.5 Mean Value Theorem

Mean Value Theorem. Let $A \subseteq R^n$ be open, $f : A \to R$ be differentiable. If A contains the point a and b together with the line segment forming a and b, then there exists c on the line segment join a and b (i.e., on $\overline{ab}$) such that

$$f(b) - f(a) = f^{'}(c).(b - a)$$

or

$$f(b) - f(a) = Df(c)(b - a).$$

Proof. Consider the function $g : [0, 1] \to A$ defined by $g(t) = (1 - t)a + tb$, g is differentiable on $(0, 1)$ (since it is linear). Now let $F = fog : F : [0, 1] \to R$. Since f and g are differentiable function $\Rightarrow$ F is differentiable with $F^{'}(t) = f^{'}(g(t)).g^{'}(t)$. Also F is continuous on $[0, 1]$ and differentiable on $(0, 1)$ $\Rightarrow$ there exists $c \in (0, 1)$ such that

$$F^{'}(c) = \frac{F(a) - F(0)}{1 - 0} = \frac{f(g(1)) - f(g(0))}{1}$$

or

$$\begin{aligned} & f(b) - f(a) = f^{'}(g(c)).g^{'}(c) \\ \Rightarrow \quad & f(b) - f(a) = f^{'}(d).(b - a), \end{aligned}$$

$d = g(c)$ is a point on the line on $\overline{ab}$.

Lemma 5.5.1. Let $A \subseteq R^n$ be open and $f : A \to R^m$ be differentiable. Let $y \in R^m$ and $G : A \to R$ is defined by $G(x) =< f(x), y >$. Then G is differentiable on A; Moreover, for each $x \in A$

$$DG(x)u =< Df(x)u, y > .$$

Proof. For each $x \in A$ define $l_x = R^n \to R$ by $l_x(x) =< Df(x), y >$, here l_x is a linear map. Now consider

$$|G(x + h) - G(x) - l_x(h)|$$

$$\begin{aligned} |G(x + h) - G(x) - l_x(h)| &= | < f(x + h), y > - < f(x), y > \\ & \quad - < Df(x)h, y > | \\ &= | < f(x + h) - f(x) - Df(x)h, y > | \\ &\leq |f(x + h) - f(x) - Df(x)h||y|. \end{aligned}$$

[By Cauchy Schwartz inequality].

$$\begin{aligned}
\Rightarrow \quad & \lim_{h\to 0} \frac{1}{|h|} |G(x+h) - G(x) - l_x(h)| \\
& \leq \lim_{h\to 0} \frac{1}{|h|} |f(x+h) - f(x) - Df(x)h||y| \\
& \leq |y| \lim_{h\to 0} \frac{1}{|h|} |f(x+h) - f(x) - Df(x)h| = 0 \\
& [\ \texttt{since } f \texttt{ is differentiable}\ ] \\
\leq \quad & \lim_{h\to 0} \frac{1}{|h|} |G(x+h) - G(x) - l_x(h)| = 0 \Rightarrow G \\
& \texttt{is differentiable on } A.
\end{aligned}$$

Theorem 5.5.1. Let $A \subseteq R^n$ be open, and $f : A \to R^m$ be differentiable. Suppose A contains a and b together with the line segment joining a and b. Then there exists $c \in \overline{ab}$ such that

$$|f(b) - f(a)| \leq |Df(c)(b-a)|.$$

Proof. If $f(b) - f(a) = \overline{0}$, then done. Suppose $f(b) \neq f(a)$. Let

$$y = \frac{f(b) - f(a)}{|f(b) - f(a)|}.$$

Next define $G : A \to R$ by $G(x) =< f(x), y >$

$$\begin{aligned}
\Rightarrow G(b) - G(a) &= < f(b) - f(a), y > \\
&= \frac{1}{|f(b) - f(a)|} < f(b) - f(a), f(b) - f(a) >, \\
G(b) - G(a) &= |f(b) - f(a)|.
\end{aligned}$$

On the other hand by the Lemma 5.5.1 $G : A \to R$ is differentiable. Thus there exists $c \in \overline{ab}$ such that $G(b) - G(a) = DG(c)(b-a)$ (M.V.T.)

$$\begin{aligned}
& \therefore |f(b) - f(a)| = DG(c)(b-a) \\
& |f(a) - f(a)| = DG(c)(b-a) =< Df(c)(b-a), y > \\
& \quad \texttt{(Lemma 5.5.1)} \\
\Rightarrow \quad & |f(b) - f(a)| \leq |Df(c)(b-a)||y| \texttt{(By Cauchy Schwartz)} \\
& \Rightarrow |f(b) - f(a)| \leq |Df(c)(b-a)|.
\end{aligned}$$

Corollary 5.5.1. Let $A \subseteq R^n$ be open and connected and $f : A \to R^m$ differentiable with $Df(x) = 0$ for all $x \in A$, then f is constant.

Proof. Let $F, S \in A$. Since A is open and connected, there exists $x_0, x_1, \ldots, x_k$ in A with $x_0 = a$ and $x_k = b$ and $\overline{x_{i-1}x_i} \in A$. Hence by the above theorem

$$|f(x_i) - f(x_{i-1})| \leq |Df(c_i)(x_i - x_{i-1})| = 0 \texttt{ for some } c_i \texttt{ on } \overline{x_{i-1}x_i}$$

but

$$\begin{aligned} |f(b) - f(a)| &= \left|\sum_{i=1}^{k} f(x_{i-1}) - f(x_i)\right| \leq \sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| = 0 \\ \Rightarrow f(b) &= f(a). \end{aligned}$$

Since a and b are arbitrary hence f is constant on A.

Exercises 5.5

1. Define the function $f : R^3 \to R$ by

   $$f(x, y, z) = xyz + x^2 + y^2 \texttt{ for } (x, y, z) \texttt{ in } R^3.$$

   The Mean Value Theorem implies that there is a number θ with $0 < \theta < 1$ for which

   $$f(1, 1, 1) - f(0, 0, 0) = \frac{\partial f}{\partial x}(\theta, \theta, \theta) + \frac{\partial f}{\partial y}(\theta, \theta, \theta) + \frac{\partial f}{\partial z}(\theta, \theta, \theta).$$

   Find the value of θ.

2. Suppose that the function $f : R^n \to R$ has first order partial derivatives and that the point x in R^n is a local minimizer for $f : R^n \to R$, meaning that there is a positive number r such that

   $$f(x + n) \geq f(x) \texttt{ if } d(x, x + h) < r.$$

   Prove that $Df(x) = 0$.

3. Suppose that the function $f : R^n \to R$ is continuously differentiable. Define $K = \{x \in R^n : \|x\| \leq 1\}$.

(i) Prove that there is a point x in K at which the function $f : K \to R$ attains a smallest value.

(ii) Now suppose also that if p is any point in R^n of norm 1, then $< Df(p), p >> 0$. Show that the minimizer x in (i) has norm less than 1.

5.6 Surjective Function Theorem and Open Mapping Theorem

Definition 5.6.1. Let $A \subset R^n$ be an open set and $f : A \to R^m$. f is said to be of class c^1 if the partial derivative of f exists on A and are continuous.

Theorem 5.6.1. Let $A \subset R^n$ and $f : A \to R^m$ be a function of class c^1. Then for every compact set $K \subset A$ and for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\forall x, y \in K$:

$$|x - y| < \delta \Rightarrow |Df(x)u - Df(y)u| \leq \varepsilon|u|, \forall u \in R^n.$$

Proof. Let $S : \{u \in R^n : |u| = 1\}$. Define $Df : K \times S \to R^m$ by $Df(x, u) = Df(x)u$. But note that

$$\begin{aligned} Df(x,u) &= Df(x).u = \begin{pmatrix} D_1 f^1(x) & \dots & D_n f^1(x) \\ \vdots & & \\ D_1 f^m(x) & \dots & D_n f^m(x) \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^n \end{pmatrix} \\ &= \left(\sum_{i=1}^{n} D_i f^1(x) u^i, \dots, \sum_{i=1}^{n} D_i f^m(x) u^i \right). \end{aligned}$$

Since f is of class $C^1, D_i f^j(x) u^j$ are all continuous, since $K \times S$ is compact. Thus given $\varepsilon > 0$,there exists $\delta > 0$ such that $|(x, u) - (y, u)| < \delta$

$$\begin{aligned} &\Rightarrow \quad |Df(x)u - Df(y)u| < \varepsilon \\ &\Rightarrow \quad |x - y| < \delta \Rightarrow |Df(x)u - Df(y)u| < \varepsilon. \quad |u| = 1. \end{aligned}$$

Hence for $x, y \in K$ and $u \in R^n$,

$$\begin{aligned} |x - y| < \delta &\Rightarrow \left| Df(x) \frac{u}{|u|} - Df(y) \frac{u}{|u|} \right| < \varepsilon \\ \Rightarrow \quad |x - y| < \delta &\to |Df(x)u - Df(y)u| < \varepsilon|u|. \end{aligned}$$

Corollary 5.6.1. (Approximation Theorem). Let $f : A \subset R^n \to R^m$ be of class C^1. Then for every compact set $K \subset A$ and every $\varepsilon > 0$, there exists $\delta > 0$ such that $\forall x, y, z \in K : x, y \in B_\delta(z) \Rightarrow |f(x) - f(y) - Df(z)(x-y)| \leq \varepsilon |x - y|$.

Proof. Let $K \subset A$ be compact and $\varepsilon > 0$. For each $a \in K$ there exists $\delta_1 > 0$ such that $\overline{B}_{\delta_1}(a) \subset A$. Since $\overline{B}_{\delta_1}(a)$ is compact by Theorem 5.6.1. there exists δ_2 such that $x, y \in \overline{B}_{\delta_1}(a)$ and $|x - y| < \delta_2 \Rightarrow |Df(x)u - Df(y)u| < \varepsilon |u| \forall u$. Choose $\delta(a) = \min\left\{\delta_1, \frac{\delta_2}{4}\right\}$. Now let $x, y \in \overline{B}_{\delta(a)}(a)$. Then $x, y \in \overline{B}_{\delta_1}(a)$ and $|x - y| \leq |x - a| + |a - y| \leq 2\left(\frac{\delta_2}{4}\right) = \frac{\delta_2}{2} < \delta_2$. Thus we have $x, y \in B_{\delta(a)}(a) \Rightarrow |Df(x)u - Df(y)u| < \frac{\varepsilon}{m}|u|$. Now let $x, y, z \in B_{\delta(a)}(a)$ and $V \in R^m$.

$$
\begin{aligned}
& < f(x) - f(y) Df(z)(x-y), V > \\
= \; & < f(x) - f(y), V > + < -Df(z)(x-y), V > \\
= \; & \sum_{i=1}^{m} (f^i(x) - f^i(y)) V^i + < -Df(z)(x-y), V > .
\end{aligned}
$$

By M.V.T. for each $i = 1, 2, \ldots, m$, there exists c_i on the line joining x and y such that

$$f^i(x) - f^i(y) = Df^i(c_i)(x - y).$$

Thus the above expression is equivalent to

$$
\begin{aligned}
= \; & \sum_{i=1}^{m} Df^i(c_i)(x-y) V^i + < -Df(z)(x-y), V > \\
= \; & < (Df^1(c_1), \ldots, Df^m(c_m))(x-y), V > + < -Df(z)(x-y), V > \\
= \; & < (Df^1(c_1), \ldots, Df^m(c_m))(x-y) - Df(z)(x-y), V > .
\end{aligned}
$$

Now if we take $V = f(x) - f(y) - Df(z)(x-y)$.

$$
\begin{aligned}
& |f(x) - f(y) - Df(z)(x-y)| \\
= \; & |(Df^(c_1) \ldots Df^m(c_m)(x-y) - Df(z)(x-y)| \\
\leq \; & \sum_{i=1}^{m} |(Df^i(c_i) - Df^i(z))(x-y)| \\
< \; & \sum_{i=1}^{m} \frac{\varepsilon}{m} |x-y| = \varepsilon |x-y| \quad (*, *).
\end{aligned}
$$

Now we cover K by $\left\{B_{\frac{\delta(a)}{2}}(a) : a \in K\right\}$.

Since K is compact there exists $B_{\frac{\delta_1}{2}}(a_1) \ldots B_{\frac{\delta_k}{2}}(a_k)$ that covers K. Choose $\delta = \min\left\{\frac{\delta_1}{2}, \frac{\delta_2}{2}, \ldots, \frac{\delta_k}{2}\right\}$. Then for $x, y, z \in K$.

Claim : $x, y \in B(z, \delta) \Rightarrow |f(x) - f(y) - Df(z)(x-y)| < \varepsilon|x - y|, z \in B_{\frac{\delta_i}{2}}(a_i)$ for some i.

Then $|x - a_i| \leq |x - z| + |z - a_i| < \delta + \frac{\delta_i}{2} < \delta_i$, i.e. $x \in B_{\delta_i}(a)$.

Similarly $y \in B_{\delta_i}(a)$. Thus by $(**)$ are yet $|f(x) - f(y) - Df(z)(x-y)| < \varepsilon|x-y|$.

Corollary 5.6.2. Let $A \subset R^n$ be open and $f : A \to R^m$ be a function of class C^1. Then if $Df(a)$ is injective for $a \in A$, then there exists $B_\delta(a) \subseteq A$ and $m > 0$ such that $\forall x \in B_\delta(a)$ and $u \in R^n : |Df(x)u| \geq m|u|$. Moreover if $df(x)$ is injective for every x in a compact set $K \subset A$ then there exists $m > 0$ such that

$$|Df(x)u| \geq m|u|, \forall x \in K, \forall u \in R^n.$$

Proof. Since A is open there exists $\delta_1 > 0$ such that $\overline{B}_{\delta_1}(a) \subset A$. Since $Df(a)$ is injective there exists $l > 0$ such that $|Df(a)u| \geq l|u|$. Since $\overline{B}_{\delta_1}(a)$ is compact by Theorem 5.6.1 there exists $\delta, 0 < \delta < \delta_1$ such that

$$x \in B_\delta(a) \Rightarrow |Df(a)u - Df(x)u| \leq \frac{l}{2}|u|.$$

Then we have

$$\begin{aligned} l|u| \leq |Df(a)u| &= |Df(a)u - Df(x)u + Df(x)u| \\ &\leq |Df(a)u - Df(x)u| + Df(x)|u| \\ &\leq \frac{l}{2}|u| + Df(x)|u| \\ \Rightarrow \frac{l}{2}|u| \leq |Df(x)|u| & \quad . \end{aligned}$$

Thus we choose $m = \frac{l}{2}$. Now if $K \subset A$ is compact and $Df(a)u$ is continuous for each $u \in K$. By the above technique for each $a \in K$. Construct $B_{\delta(a)}(a) \subset A$ such that $\forall x \in B_{\delta(a)}(a), |Df(x)u| \geq m_a|u|$.

But $\left\{B_{\frac{\delta(a)}{2}}(a)\right\}$ is an open cover of K and there exists $B_{\frac{\delta_1(a)}{2}}(a_1) \ldots B_{\frac{\delta_n(a)}{2}}(a_n)$ cover K. Choose $m = \min\{ma_1, \ldots, ma_n\}$.

Then

$$|Df(x)||u| \geq m|u|.$$

Corollary 5.6.3. (Injective Function Theorem) Let $A \subset R^n$ be open and $f : A \to R^m$ be of class C^1. If $Df(a)$ is injective at $a \in A$ then there exists $B_\delta(a) \subset A$ and $m > 0$ such that $Df(x)$ is injective $\forall x \in B_\delta(a)$ and $\forall x, y \in B_\delta(a), |f(x) - f(y)| \geq M|x - y|$.

Proof. By corollary 5.6.2, there exists $\overline{B}_{\delta_1}(a) \subset A$ and $l > 0$, there exists $\forall z \in \overline{B}_{\delta_1}(a) Df(z)$ is injective and

$$|Df(z)(x - y)| \geq l|x - y| \forall x, y \in R^n.$$

By approximation theorem there exists $\delta_2 > 0$ such that $\forall x, y, z \in \overline{B}_{\delta_1}(a)$.

$$x, y, \in B_{\delta_2}(z) \Rightarrow |f(x) - f(y)| - Df(z)(x - y)| \leq \frac{l}{2}|x - y|$$
$$x, y, \in B_{\delta_2}(a) \Rightarrow |f(x) - f(y)| - Df(a)(x - y)| \leq \frac{l}{2}|x - y|.$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Then for $x, y \in B_\delta(a)$

$$|f(x) - f(y) - Df(a)(x - y)| \leq \frac{l}{2}|x - y|$$
$$\Rightarrow \quad \frac{l}{2}|x - y| \leq |Df(a)(x - y)| - \frac{l}{2}|x - y| \leq |f(x) - f(y)|,$$

set $m = \frac{l}{2}$

$$\Rightarrow |f(x) - f(y)| \geq M|x - y|.$$

Corollary 5.6.4. Let $A \subset R^n$ be open, $f : A \to R^m$ be of class C^1 and K be a compact subset of A. Then there exists$M > 0 \ni \forall x \in K$

$$|Df(x)u| \leq M|u|.$$

Proof. By Theorem 5.6.1 for every $a \in K$, there exists $\delta(a) > 0$ such that $\forall x \in B_{\delta(a)}(a) \cap K : |Df(x)u - Df(a)u| \leq |u|$ (i.e. chosen $\varepsilon = 1$).

$$\Rightarrow |Df(x)u| \leq |u| + |Df(a)u|.$$

But there exists $m_0 > 0$ such that $|Df(a)u)| \leq m_0|u|, m = 1 + m_0$.

Cover K by $\{B_{\delta(a)}(a)\}_{a\in K}, K$ is compact therefore there exists a finite subcover $\{B_{\delta(a_1)}(a_1)\dots D_{\delta(a_n)}(a_n)\}$ which cover K. Choose $M = \max\{ma_1, ma_2, \dots, ma_n\}$ then $|Df(x)u| \leq M|u|$.

Theorem 5.6.2. (Surjective Function Theorem) Let $A \subset R^n$ be open, $f : A \to R^m$ be of class C^1. If $Df(a)$ for $a \in A$ is surjective then there exist positive numbers α and β such that $\overline{B}_\alpha(a) \subset A, \overline{B}_{\frac{\alpha}{2\beta}}(f(a)) \subseteq f(\overline{B}_\alpha(a))$.

Proof. Since $Df(a)$ is surjective, there exists $u_1, u_2, \dots, u_m$ in R^n such that $Df(a)u_i = e_i$ where e_i's are the usual basis of R^m. Let $M : R^m \to R^n$ be the linear map

$$M\left(\sum_{i=1}^m a_ie_i\right) = \sum_{i=1}^m a_iu_i.$$

Then clearly $Df(a)oM : R^m \to R^m$ is the identity map. Set

$$\beta = \left(\sum_{i=1}^m |u_i|^2\right)^{1/2}.$$

For $y = \sum_{i=1}^m a_ie_i \in R^m$, we have

$$\begin{aligned}|M(y)| &= \left|\sum a_iu_i\right| \leq \sum |a_i||u_i| \\ &\leq \left(\sum |a_i|^2\right)^{1/2}\left(\sum |u_i|^2\right)^{1/2} = \beta|y|.\end{aligned}$$

$$\therefore |M(y)| \leq \beta|y|. \tag{5.6.1}$$

Now for $\varepsilon = \frac{1}{2\beta}$. By approximation theorem there exists $\alpha > 0$ such that

$$x, y \in B_\alpha(a) \Rightarrow |f(x) - f(y) - Df(a)(x-y)| \leq \frac{1}{2\beta}|x-y|.. \tag{5.6.2}$$

Now suppose $y \in \overline{B}_{\frac{\alpha}{2\beta}}(f(a))$. We claim there exists $x \in \overline{B}_\alpha(a)$ such that $y = f(x)$.

To show this we shall construct a sequence $\{x_p\}_{p=0}^\infty$ with the property that

(a) $|x_p - x_{p-1}| \leq \frac{\alpha}{2}$

(b) $|x_p - a| \leq \left(1 - \frac{1}{2}\right) \alpha$.

Let $x_0 = a$ and $x_1 = x_0 + M(y - f(a))$.
Then

$$\begin{aligned} |x_1 - x_0| &= |M(y - f(a))| \leq \beta |y - f(a)| \leq \beta \frac{\alpha}{2\beta} = \frac{\alpha}{2} \\ |x_1 - a| &= |x - x_0| \leq \frac{\alpha}{2} = \left(1 - \frac{1}{2}\right) \alpha \quad \texttt{i.e.} \quad (b). \end{aligned}$$

Suppose $x_0, x_1, \ldots, x_p$ satisfied the property $(a)\&(b)$. Define x_{p+1} by

$$x_{p+1} = x_p - |M|f(x_p) - f(x_{p-1}) - Df(a)(x_p - x_{p-1})||.$$

Then by (5.6.1) and (5.6.2) it follows that
Claim: $\{x_p\}_{p=0}^{\infty}$ is a Cauchy sequence.

$$\begin{aligned} |x_{p+k} - x_p| &\leq |x_{p+k} - x_{p+k-1}| + |x_{p+k-1} - x_{p+k-2}| + \ldots \\ &\quad + |x_{p+1} - x_p| \\ &\leq \frac{\alpha}{2^{p+k}} + \frac{\alpha}{2^{p+k-1}} + \cdots + \frac{1}{2^{p+1}} \\ &= \frac{\alpha}{2^p} \left(\frac{1}{2^k} + \frac{1}{2^{k-1}} + \cdots + \frac{1}{2} \right) \leq \frac{\alpha}{2^p} \to 0 \ \texttt{as} \ p \to \infty. \end{aligned}$$

$\therefore \{x_p\}$ is a Cauchy sequence.
i.e., $x_p \to x$ for some x, but since $|x_p - a| \leq \left(1 - \frac{1}{2^p}\right) \alpha$

$$|x - a| \leq \alpha \Rightarrow x \in B_\alpha(a).$$

Claim $Df(a)(x_{p+1} - x_p) = y - f(x_p)$.
We will show it by induction, for $p = 0$, by definition

$$x_1 - x_0 = M(y - f(a))$$

$$\Rightarrow Df(a)(x_1 - x_0) = Df(a)M(y - f(a)) = y - f(a).$$

By the definition of x_{p+1}, we have

$$\begin{aligned} Df(a)(x_{p+1}-x_p) &= -Df(a).M(f(x_p)-f(x_{p-1})- \\ &\quad Df(a)(x_p-x_{p-1})) \\ &= Df(a)(x_p-x_{p-1})-f(x_p)-f(x_{p-1}) \\ &= y-f(x_{p-1})-(f(x_p)-f(x_{p-1})) \\ &= y-f(x_p). \end{aligned}$$

Now since $Df(a)$ is continuous and f is continuous

$$\begin{aligned} \lim_{p\to\infty} Df(a)(x_{p+1}-x_p) &= \lim_{p\to\infty}(y-f(x_p)) \\ 0 &= y-f(x) \\ \Rightarrow y &= f(x). \end{aligned}$$

Corollary 5.6.5. (Open Mapping Theorem) Let $A \subset R^n$ be open, $f : A \to R^m$ be of class C^1, if for each $x \in A, Df(x)$ is surjective then $G \subset A$ is open implies $f(G)$ is open in R^m.

Proof. Suppose G is open. Let $c \in f(G)$ then there exists $u \in G$ such that $f(u) = c$. By surjective function theorem there exist $\alpha > 0, \beta > 0$ such that $\overline{B}_\alpha(u) \subset G$ and $ovB_{\frac{\alpha}{2\beta}}(f(u)) \subseteq f(\overline{B}_\alpha(u)) \Rightarrow \overline{B}_{\frac{\alpha}{2\beta}}(f(u)) \subset f(G)$ i.e., $\overline{B}_{\frac{\alpha}{2\beta}}(c) \subseteq f(G) \Rightarrow f(G)$ is open.

Corollary 5.6.6. Let $A \subset R^n$ be open, $f : A \to R^n$ be of class C^1. If $Df(x)$ is injective $\forall x \in A$. Then for $G \subset A$ is open, $f(G)$ is open $\Rightarrow f(A)$ is open.

Proof. $Df(x)$ is injective $\Rightarrow Df(x)$ is surjective and in view of Corollary 5.6.5 we get the proof.

5.7 The Inverse and The Implicit Function Theorem

As an application of surjective function theorem and open mapping theorem we will prove two famous and fundamental theorems.

Lemma 5.7.1. $f : R^n \to R^m, g : R^m \to R^k$ such that $D(gof)(a)$

and $Dg(f(a))$ exists, $Dg(f(a))$ is injective, f is continuous. Then $Df(a)$ exists and is given by

$$Df(a) = (Dg(f(a)))^{-1}.D(gof)(a).$$

Proof. By Injective function theorem there exists $B_\delta(a) \subset A$ such that $Df(x)$ is injective $\forall x \in B_\delta(a)$ and $f|_{B_delta(a)}$ is one to one. Let $U = B_\delta(a)$ and $V = f(U)$ by Corollary 5.6.6, V is open. Since $f : U \to V$ is one one onto it has inverse $g : V \to U$, again by Corollary 5.6.6 g is continuous $\Rightarrow$ ($g^{-1}(G) = f(G)$ is open if G is open in U).

Now $fog : V \to V$ is the identity map. Thus $D(fog)(x) = I$, $\forall x \in V$. But $Df(gf(x)) = Df(x) \forall x \in U \Rightarrow Df(g(y))$ exists and is injective $\forall y \in V$. Thus it follows that $Dg(y)$ exists $\forall y \in V$ and

$$Dg(y) = D(f(g(y)))^1 oD(f0g)(y) = (Df(x))^{-1} \texttt{ where } y = f(x).$$

Theorem 5.7.1. (Inverse Function) Let $A \subset R^n$ be an open set, $f : A \to R^n$ be a function of class C^1 and $Df(a)$ is injective at $a \in A$. Then there are two open sets $U \& V$ in R^n such that $a \in U \subset A, f(a) \in V$, and $f|_U$ is one to one with range V whose inverse is also of class C^1 and

$$Df^{-1}(y) = (Df(x))^{-1} \texttt{ where } y = f(x).$$

Proof. To obtain the complete proof in view of above Lemma 5.7.1 it is remains to prove that g is also of class C^1.

$$g^1(y) = (f^1 g(y))^{-1}.$$

Let $f^1(x) = a_{ij}(x) = A(\chi), \quad \chi = g(y)$
$\Rightarrow g^1(y) = a_{ij}(g(y))^{-1}$. Now if $g^1(y) = (b_{ij}(y))_{n \times n}$.
By the Cramer's rule

$$b_{ij}(y) = (-1)^{i+j} \frac{detA^{ij}(g(y))}{detA(g(y))},$$

where $A^{ij}(y)$ is $(n-1) \times (n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column from $A(y)$. But *determinant* is continuous, a_{ij}'s are continuous, g is continuous, $detA(g(y)) \neq 0$. Thus $b_{ij}(y)$ is continuous $\Rightarrow g$ is of class C^1.

Theorem 5.7.2. (Implicit Function Theorem)

Let $C \subset R^n \times R^m$ be open, $f : C \to R^m$ be a function of class C^1 such that for $(a, b) \in C$, $f(a, b) = 0$, the $m \times m$ matrix $(D_{n+j} f^i(a, b) 1 \leq i, j \leq m$ has a non-zero determinant. Then there is an open set $A \subset R^n$ containing a and an open set $B \subset R^n$ containing b and a unique C^1 function $g : A \to B$ such that

$$f(x, y(x)) = 0, \forall x \in A.$$

Proof. Define $F : C \to R^n \times R^m$ by

$$F(x, y) = (x, f(x, y)) \forall (x, y) \in C.$$

Since f and identity functions are of class C^1, so is F, and

$$F^1(a, b) = \begin{pmatrix} 1 & & \\ \vdots & & \\ 0 & & \\ D_1 f^1(a, b) & & \\ \vdots & & \\ D_1 f^m(a, b) & & \end{pmatrix}$$

$\Rightarrow det F^1(a, b) = det(D_{n+j} f^i(a, b))_{n \times n} \neq 0.$

Hence by I.F.T. there exists $W \subset R^n \times R^m$ open set containing $F(a, b) = (a, 0)$ and an open set $V \subset C$: containing (a, b), $F : V \to W$ has an inverse i.e., $h : W \to V$ which is of class C^1. Without loss of generality assume that $V = A \times B$ where A is open in R^n and B is open in R^m. Note that there exists $N \subset A$ such that $N \times \{0\} \subset W$. Clearly $h : W \to A \times B$ has the form

$h(x, y) = (x, K(x, y))$ where $K : W \to B$ is of class C^1, since F is of this form.

Let $\pi : R^n \times R^m \to R^m$ given by $\pi(x, y) = y$. Then $\pi o f = f$. And for $(x, y) \in W$,

$$\begin{aligned} f(x, K(x, y)) &= (foh)(x, y) = (\pi o f) o h(x, y) \\ &= \pi o(F o h)(x, y) = \pi(x, y) = y. \end{aligned}$$

Hence $f(x, K(x, y)) = 0, \forall x \in N$. Now let $g : N \to B$ be given by

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

Then g is the required function.

Exercises 5.7

1. Let $f : R^n \to R^n$ be given by the equation $f(x) = \|x\|^2.x$. Show that f is of class C^∞ and that f carries the unit ball $B(0,1)$ onto itself in a one-to-one fashion. Show, however, that the inverse function is not differentiable at 0.

2. Let $g : R^2 \to R^2$ be given by the equation

$$g(x,y) = (2ye^{2x}, xe^y).$$

   Let $f : R^2 \to R^3$ be given by the equation

$$f(x,y) = (3x - y^2, 2x + y, xy + y^3).$$

   (a) Show that there is a neighborhood of $(0,1)$ that g carries in a one-to-one fashion onto a neighborhood of $(2,0)$.

   (b) Find $D(fog^{-1})$ at $(2,0)$.

3. Let $f : R^3 \to R^2$ be of class C^1, write f in the form $f(x, y_1, y_2)$. Assume that $f(3,-1,2) = 0$ and

$$Df(3,-1,2) = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

   (i) Show there is a function $g : B \to R^2$ of class C^1 defined on an open set B in R such that

$$f(x, g_1(x), g_2(x)) = 0$$

   for $x \in B$, and $g(3) = (-1,2)$.

   (ii) Find $Dg(3)$.

   (iii) Discuss the problem of solving the equation $f(x, y_1, y_2) = 0$ for an arbitrary pair of the unknowns in terms of the third, near the point $(3,-1,2)$.

4. Let $f : R^{k+n} \to R^n$ be of class C^1; suppose that $f(a) = 0$ and that $Df(f)$ has rank n. Show that if c is a point of R^n sufficiently close to zero, then the equation $f(x) = c$ has a solution.

5. Let $f : R^2 \to R$ be of class C^1, with $f(2,-1) = -1$. Set

$$\begin{aligned} G(x,y,u) &= f(x,y) + u^2, \\ H(x,y,u) &= ux + 3y^3 + u^3. \end{aligned}$$

The equations $G(x,y,u) = 0$ and $H(x,y,u) = 0$ have the solution $(x,y,u) = (2,-1,1)$.

   (a) What conditions on Df ensure that there are C^1 functions $x = g(y)$ and $u = h(y)$ defined on an open set in R that satisfy both equations, such that $g(-1) = 2$ and $h(-1) = 1$?

   (b) Under the conditions of (a), and assuming that $Df(2,-1) = [1,-3]$, find $g'(-1)$ and $h'(-1)$.

6. Let A be open in R^n; let $f : A \to R^n$ be of class C^r, assume $Df(x)$ is non-singular for $x \in A$. Show that even if f is not one-to-one on A, the set $B = f(A)$ is open in R^n.

7. Define the function $f : R \to R$ by

$$f(x) = x^3 - 3x + 1 \texttt{ for } x \texttt{ in } R.$$

At what points x in R does the Inverse Functions Theorem apply ?

8. For each of the following mapping $F : R^2 \to R^2$, apply the Inverse Function Theorem at the point $(x_0, y_0) = (0,0)$ and calculate the partial derivatives of the components of the inverse mapping at the point $(u_0, \nu_0) = F(0,0)$:

   (a) $F(x,y) = (x + x^2 + e^{x^2y^2}, -x + y + \sin(xy))$ for (x,y) in R^2.

   (b) $F(x,y) = (e^{x+y}, e^{x-y})$ for (x,y) in R^2.

9. Define the mapping $F : R^2 \to R^2$ by

$$F(r,\theta) = (r\cos\theta, r\sin\theta) \texttt{ for } (r,\theta) \texttt{ in } R^2.$$

   (a) At what points (r_0, θ_0) in R^2 can we apply the Inverse Function Theorem to this mapping ?

(b) Find some explicit formula for the local inverse about the point $(r, \theta) = (1, \pi/2)$.

10. Suppose that the function $\varphi : R^2 \to R$ is continuously differentiable and define the mapping $F : R^2 \to R^2$ by

$$F(x, y) = (\varphi(x, y), -\varphi(x, y)) \texttt{ for } (x, y) \texttt{ in } R^2.$$

(a) Explain analytically why the hypotheses of the Inverse Function Theorem fail at each point (x_0, y_0) in R^2.

(b) Explain geometrically why the conclusion of the Inverse Function Theorem must fail at each point (x_0, y_0) in R^2.

11. For a point (ρ, θ, ϕ) in R^3, define

$$F(\rho, \theta, \phi) = (\rho \sin\phi \cos\theta, \rho \sin\theta \sin\theta, \rho \cos\phi).$$

At what points $(\rho_0, \theta_0, \phi_0)$ in R^3 does the Inverse Function Theorem apply to the mapping $F : R^3 \to R^3$?

12. Use the Implicit Function Theorem to analyze the solutions of the given systems of equations near the solution 0.

1. $$\left\{\begin{array}{ll} (x^2 + y^2 + z^2)^3 & -x + z = 0 \\ \cos(x^2 + y^4) & +e^c - 2 = 0, \quad (x, y, z) \texttt{ in } R^3. \end{array}\right.$$

2. $$\left\{\begin{array}{ll} (u\nu)^4 & +(u + s)^3 + t = 0 \\ \sin(u\nu) & +e^{\nu + t^2} - 1 = 0, \quad (u, \nu, s, t) \texttt{ in } R^4. \end{array}\right.$$

13. In the proof of the Implicit Function Theorem, it was asserted that the invertibility of the $k \times k$ matrix $D_yF(x_0, y_0)$ implies the invertibility of the $(n + k) \times (n + k)$ matrix $DH(x_0, y_0)$. Verify this assertion.

14. Graph the solutions of the equation

$$y^3 - x^2 = 0, (x, y) \texttt{ in } R^2.$$

Does the Implicit Function Theorem apply at the point $(0, 0)$? Does this equation define one of the components of a solution (x, y) as a function of other component ?

15. Suppose that the function $f : R^2 \to R$ is continuously differentiable and that there is a positive number C such that

$$\frac{\partial f}{\partial y}(x, y) \geq c \texttt{ for every } (x, y) \texttt{ in } R^2.$$

Prove that there is a continuously differentiable function $g : R \to R$ with $f(x, g(x)) = 0$ for every x in R and that if $f(x, y) = 0$, then $y = g(x)$.

Chapter 6

Multiple Integrals

In this chapter we have discussed the double integral over a rectangle, double integral over general region, double integral in polar coordinates, application to center of mass and surface area, triple integral, triple integral in cylindrical coordinates, triple integral in spherical coordinates and change of variables in multiple integrals.

6.1 The Double Integral Over a Rectangle

Let R be a rectangle in the xy plane and f a continuous function on R. To evaluate the double integral over the rectangle R, consider the partition P of R into sub-rectangle $R_1, R_2, \ldots, R_k, \ldots, R_n$.

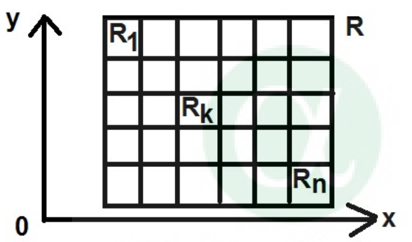

Fig. 6.1.1.

Since f is continuous on R and hence on each sub-rectangle R_k, f attains its maximum and minimum value on R_k.

Let m_k be the minimum value of f on R_k
M_k be the maximum value of f on R_k
$\triangle A_k$ be the area of the sub-rectangle R_k.

Then $m_k \triangle A_k \leq M_k \triangle A_k \forall k \in \{1, 2, \ldots, n\}$.

The sum $L_f(P) = \sum_{k=1}^{n} m_k \triangle A_k$ is called the lower sum of f with respect to P.

The sum $U_f(P) = \sum_{k=1}^{n} M_k \triangle A_k$ is called the upper sum of f with respect to P and we have

$$L_f(P) \leq U_f(P).$$

Example 6.1.1. Let $f(x) = x + y$ be defined on the rectangle

$$R = \{(x, y) : 2 \leq x \leq 5, 1 \leq y \leq 3\} = [2, 5] \times [1, 3].$$

Find the lower and upper sum with respect to partition

$$\begin{aligned} P = \{ \quad & R_1 = [2,3] \times [1, 3/2], R_2 = [2,3] \times [3/2, 3], \\ & R_3 = [3,4] \times [1, 3/2], R_4 = [3,4] \times [3/2, 3], \\ & R_5 = [4,5] \times [1, 3/2], R_6 = [4,5] \times [3/2, 3]\}. \end{aligned}$$

Solution. We have

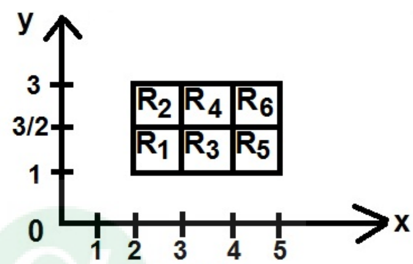

Fig. 6.1.2

$$\begin{aligned} m_1 &= f(2,1) = 2 + 1 = 3, M_1 = f(3, 3/2) = 3 + 3/2 = 9/2 \\ m_2 &= f(2, 3/2) = 2 + \frac{3}{2} = \frac{7}{2}, M_2 = f(3,3) = 3 + 3 = 6 \\ m_3 &= f(3,1) = 3 + 1 = 4, M_3 = f(4, 3/2) = 4 + 3/2 = \frac{11}{2} \\ m_4 &= f(3, 3/2) = 3 + 3/2 = \frac{9}{2}, M_4 = f(4,3) = 4 + 3 = 7 \\ m_5 &= f(4,1) = 4 + 1 = 5, M_5 = f(5, 3/2) = 5 + \frac{3}{2} = \frac{13}{2} \\ m_6 &= f(4, 3/2) = 4 + \frac{3}{2} = \frac{11}{2}, M_6 = f(5,3) = 5 + 3 = 8. \\ L_f(P) &= \sum_{k=1}^{6} m_k \triangle A_k = m_1 \triangle A + m_2 \triangle A_2 + \cdots + m_6 \triangle A_6 \end{aligned}$$

$$= 3(\frac{1}{2}) + \frac{7}{2}(\frac{3}{2}) + 4(\frac{1}{2}) + \frac{9}{2}(\frac{3}{2}) + 5(\frac{1}{2}) + \frac{11}{2}(\frac{3}{2}) = \frac{115}{4}.$$

$$\begin{aligned} U_f(P) &= \sum_{k=1}^{6} M_k \triangle A_k = M_1 \triangle A_1 + M_2 \triangle A_2 + \cdots + M_6 \triangle A_6 \\ &= \frac{9}{2}(\frac{1}{2}) + 6(3/2) + \frac{11}{2}(\frac{1}{2}) + 7(\frac{3}{2}) + \frac{13}{2}(\frac{1}{2}) + 8(\frac{3}{2}) = \frac{96}{4}. \end{aligned}$$

Definition 6.1.1.(i) The unique number I satisfying the inequality $L_f(P) \leq I \leq U_f(P)$ for every partition P of R is called the double integral of f on R and denoted by $\int\int_R f(x,y)dA$.

(ii) If f is nonnegative and integrable on the rectangle R, then the volume V of the solid region between the graph of f and R is given by

$$v = \int\int_R f(x,y)dA.$$

Example 6.1.2. Evaluate $\int\int_R KdA$, where $R = [a,b] \times [c,d]$ and K is a constant.

Solution. Consider $f(x,y) = K$, and partition $P = \{R_1, R_2, \ldots, R_n\}$ of R.

$$m_i = M_i = K \forall i \in \{1, 2, \ldots, n\}$$

$$\begin{aligned} L_f(P) &= \sum_{i=1}^{n} K \triangle A_i = K \sum_{i=1}^{n} \triangle A_i = K(\texttt{area of} R) = K(b-a)(d-c) \\ U_f(P) &= \sum_{i=1}^{n} K \triangle A_i = K(\texttt{area of} R) = K(b-a)(d-c) \end{aligned}$$

$\Rightarrow L_f(P) = U_f(P)$ for every partition P of R.

$\Rightarrow \int\int_R KdA = K(b-a)(d-c)$ since P is arbitrary.

Remark 6.1.1.(i) If $K > 0$, then $\int\int_R KdA$ is the volume of the rectangular parallelopiped with hight K and base R.

(ii) $\int\int_R KdA = \int\int_R Kdxdy = \int_c^d(\int_a^b Kdx)dy = \int_c^d K(b-a)dy = K(b-a)(d-c)$.

Exercises 6.1

1. For the generalized rectangle $I = [0,1] \times [0,1]$ in the plane R^2, define

$$f(x,y) = \begin{cases} 5 & if \quad (x,y) \texttt{ is in } I \texttt{ and } x > \frac{1}{2} \\ 1 & if \quad (x,y) \texttt{ is in } I \texttt{ and } x \leq \frac{1}{2}. \end{cases}$$

   Show that the function $f : I \to R$ is integrable.

2. For the rectangle $I = [0,1] \times [0,1]$ in the plane R^2, define the function $f : I \to R$ by $f(x,y) = xy$ for $(x,y) \in I$. Evaluate $\int_I f$.

3. For the rectangle $I = [0,1] \times [-1,0]$ in the plane R^2, define the function $f : I \to R$ by $f(x,y) = x^2y$ for $(x,y) \in I$. Evaluate $\int_I f$.

4. For the rectangle $I = [0,2] \times [0,1]$ in the plane R^2, define the function $f : I \to R$ by $f(x,y) = x^2y$ for $(x,y) \in I$. Evaluate $\int_I f$.

6.2 The Double Integral Over General Regions

Suppose R is any bounded region, that is contained in a rectangle R' and f is a function continuous on R. Extend f to all of R' by setting f equals to zero in R'/R, i.e., defined g by

$$g(x,y) = \begin{cases} f(x,y) & if \quad (x,y) \in R \\ 0 & if \quad (x,y) \in R'/R. \end{cases}$$

g is continuous on R' except possibly at the boundary of R. Then

$$\int\int_R f(x,y)dA = \int\int_{R'} g(x,y)dA \texttt{ if right hand side exists.}$$

Evaluation of double integral by iterated integral

1. Vertically simple (plane) region

$$R_1 = \{(x, y) : a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\}$$

where g_1 and g_2 are continuous functions over $[a, b]$.

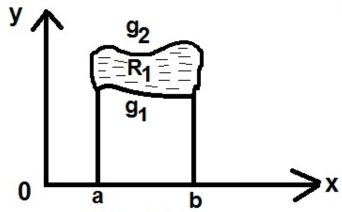

Fig. 6.2.1

2. Horizontally simple region

$$R_2 = \{(x, y) : c \leq y \leq d \texttt{ and } h_1(y) \leq x \leq h_2(y)\}$$

where h_1 and h_2 are continuous functions over $[c, d]$.

A region R is simple if R is both vertically and horizontally simple.

eg. Let R be the region between the graphs of $y = x^2, y = z - x$ and $x = 0$. Show that R is simple.

Simple $R = \{(x, y) : 0 \leq x \leq 1, x^2 \leq y \leq 2 - x\}$

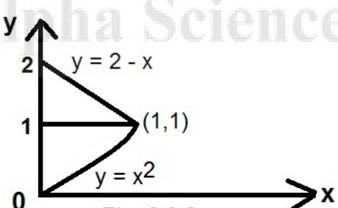

Fig. 6.2.2

it is vertically simple and

$$\begin{aligned} R &= \{(x, y) : 0 \leq y \leq 1, 0 \leq x \leq \sqrt{y}\} \cup \{(x, y) \\ &\quad : 1 \leq y \leq 2, 0 \leq x \leq 2 - y\} \\ &= \{(x, y) : 0 \leq y \leq 2, h_1(y) \leq x \leq h_2(y)\} \end{aligned}$$

where $h_1(y) = 0$ and $h_2(y) = \begin{cases} \sqrt{y} & for \quad 0 \leq y \leq 1 \\ 2 - y & for \quad 1 \leq y \leq 2, \end{cases}$,

is horizontally simple. So R is simple.

3. $\int\int_{R_1} f(x,y)dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)dydx$

and

$$\int\int_{R_2} f(x,y)dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)dxdy$$

where f is continuous on R_1 and R_2, respectively.

These integrals are called iterated integrals. For simple region R

$$\int\int_R f(x,y)dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)dydx$$
$$= \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)dxdy.$$

Note 6.2.1. Constants of integration must be from outside.

Note 6.2.2. If both x and y range over constant limits of integration, one can interchange freely.

Definition 6.2.1. The area A of a plane region R is given by $A = \int_R 1dA$.

$$A = \int\int_R dA = \int_a^b \int_{g_1(x)}^{g_2(x)} dydx = \int_a^b [y]_{g_1(x)}^{g_2(x)} = \int_a^b (g_2(x) - g_1(x))dx.$$

Example 6.2.1. Let $R = [-1,3] \times [2,4]$. Evaluate the double integrals $\int\int_R (x^2 - y^2)dydx$.

Solution. We have

$$\int\int_R (x^2 - y^2)dydx = \int_{-1}^{3}\int_{2}^{4}(x^2 - y^2)dydx$$

$$= \int_{-1}^{3}[x^2y - \frac{y^3}{3}]_2^4 dx$$

$$= \int_{-1}^{3}[4x^2 - \frac{64}{3} - 2x^2 + \frac{6}{3}]dx$$

$$= \int_{-1}^{3}(2x^2 - \frac{56}{3})dx$$

$$= [\frac{2x^3}{3} - \frac{56}{3}x]_{-1}^{3} = -56.$$

Example 6.2.2. Find the volume of the solid bounded by the graph of $z = 4x^2 + y^2$ and over the rectangular region R in the xy plane having vertices $(0,0,0)(0,1,0)(2,0,0)(2,1,0)$.

Solution. We have

$$R = \{(x, y) : 0 \le x \le 2, 0 \le y \le 1\}$$

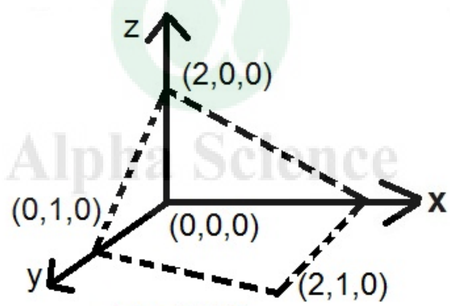

Fig. 6.2.3

$$v = \int\int_R (4x^2 + y^2)dxdy$$

$$= \int_0^1\int_0^2 (4x^2 + y^2)dxdy = \int_0^1 [4\frac{x^3}{3} + y^2x]_0^2 dy$$

$$= \int_0^1 (\frac{32}{3} + 2y^2)dy$$

$$= [\frac{32}{3}y + \frac{2y^3}{3}]_0^1 = \frac{34}{3}.$$

Note 6.2.3. If R is a union of non overlapping regions $R_i, i = 1, \ldots, n$ i.e, $R = \cup_{i=1}^{n} R_i$, then

$$\int\int_R f(x,y)dA = \sum_{i=1}^{n} \int\int_{R_i} f(x,y)dA.$$

Example 6.2.3. Evaluate $\int\int_R ydA$ where R is the region bounded by the graphs of $y = \cos x, y = \sin x, x = 0, x = \frac{3\pi}{4}$.

Solution. We have

$$\int\int_R ydA = \int\int_{R_1} ydA + \int\int_R ydA$$

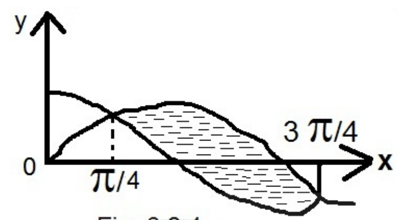

Fig. 6.2.4

$$\begin{aligned}
&= \int_0^{\pi/4}\int_{\sin x}^{\cos x} ydydx + \int_{\pi/4}^{3\pi/4}\int_{\cos x}^{\sin x} ydydx \\
&= \int_0^{\pi/4} \frac{1}{2}(\cos^2 x - \sin^2 x)dx + \int_{\pi/4}^{3\pi/4} \frac{1}{2}(\sin^2 x - \cos^2 x)dx \\
&= \frac{1}{2}\int_0^{\pi/4} \cos 2xdx + \frac{1}{2}\int_{\pi/4}^{3\pi/4} -\cos 2xdx \\
&= \frac{1}{2}[\frac{1}{2}\sin 2x]_0^{\pi/4} - \frac{1}{2}[\frac{1}{2}\sin 2x]_{\pi/4}^{3\pi/4} = \frac{3}{4}.
\end{aligned}$$

Example 6.2.4. Find the volume of the solid between the cylinder $x^2 + z^2 = 9$ and the planes $y = 0$ and $y + z = 4$.

Solution. Consider y to play the role of z and $f(x,z) = y = 4 - z$

$$\begin{aligned}
R &= \left\{(x,z) : -3 \le x \le 3, -\sqrt{9-x^2} \le z \le \sqrt{9-x^2}\right\} \\
V &= \int\int_R f(x,z)dzdx = \int_{-3}^{3}\int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} (4-z)dzdx
\end{aligned}$$

$$= \int_{-3}^{3}[4z - \frac{z^2}{2}]_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}}$$
$$= \int_{-3}^{3}[4\sqrt{9-x^2} - (\frac{9-x^2}{2}) - (-4\sqrt{9-x^2} - \frac{9-x^2}{2})]dx$$
$$= \int_{-3}^{3} 8\sqrt{9-x^2}dx.$$

Put $x = 3\sin\theta, dx = 3\cos\theta d\theta$

$$= 72\int_{-\pi/2}^{\pi/2}\cos^2\theta d\theta = 72\int_{-\pi/2}^{\pi/2}\frac{1}{2}(1+\cos 2\theta)d\theta$$
$$= 36[\theta + \frac{1}{2}\sin 2\theta]_{-\pi/2}^{\pi/2} = 36\pi.$$

Exercises 6.2

1. Let $R = [0,1] \times [\pi/2, \pi]$. Evaluate the double integral

$$\int\int_R x\sin\frac{y}{2}dydx.$$

2. Evaluate the iterated integrals

   (a) $\int_{1/2}^{1}\int_0^{2x}\cos(\pi x^2)dydx.$

   (b) $\int_1^3\int_{-y}^{2y} xe^{y^3}dxdy.$

3. Evaluate

   $\int\int_R (x+y-1)dxdy$, where R is given by

$$R = \{(x,y) : -1 \le x \le 2 \text{ and } x^2 \le y \le x+2\}.$$

4. Evaluate

$$\int\int_R \frac{x}{y}dydx, \quad R = \{(x,y) : \frac{1}{2} \le x \le 1, x^2 \le y \le e^x\}.$$

6.3 Double Integral in Polar Coordinates

Suppose that h_1 and h_2 are continuous on $[\alpha, \beta]$, where $0 \leq \beta - \alpha \leq 2\pi$ and that $0 \leq h_1(\theta) \leq h_2(\theta)$ for $\alpha \leq \theta \leq \beta$. Let R be the region between the polar graphs of $r = h_1(\theta), r = h_2(\theta)$ for $\alpha \leq \theta \leq \beta$.

If f is continuous on R, then

$$\int\int_R f(x,y)dA = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta)rdrd\theta.$$

Hence the transformation is $x = r\cos\theta, y = r\sin\theta$ and

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

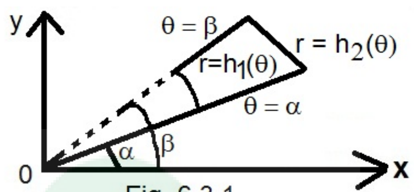

Fig. 6.3.1

Note 6.3.1. If f is nonnegative on R, the volume v of the region between the graphs of f and R is

$$v = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta)rdrd\theta$$

and the area of R is

$$A = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} rdrd\theta.$$

Example 6.3.1. Use polar coordinates to evaluate

$$\int_{-a}^{a} \int_0^{\sqrt{a^2-x^2}} (x^2+y^2)^{3/2}dydx.$$

Solution. We have

$$\begin{aligned} R &= \{(x,y) : -a \leq x \leq a, 0 \leq y \leq \sqrt{a^2-x^2}\} \\ R &= \{(r,\theta) : 0 \leq r \leq a, 0 \leq \theta \leq \pi\}. \end{aligned}$$

$$\int\int_R (x^2+y^2)^{3/2} dydx = \int_0^{\pi}\int_0^a (r^2)^{3/2} r dr d\theta = \frac{a^5}{5}\pi.$$

Example 6.3.2. Express $\int\int_R x^2 dA$ as an iterated integral in polar coordinates and evaluate it where R is the region bounded by $r = 2\cos\theta$.

Solution. We have

$$\int\int_R x^2 dA = \int_{-\pi/2}^{\pi/2}\int_0^{2\cos\theta} (r^2\cos^2\theta) r dr d\theta = 5\pi.$$

Example 6.3.3. Find the area of the region R that lies outside the graph of $r = a$ and inside the graph of $r = 2a\sin\theta$.

Solution. First determine the intersection points

$$\begin{aligned} r &= 2a\sin\theta \\ r^2 &= 2ar\sin\theta \\ x^2+y^2 &= 2ay \\ x^2+(y-a)^2 &= a^2 \\ r &= 2a\sin\theta = a \\ &\sin\theta = \frac{1}{2} \\ &\theta = \pi/6 \texttt{ or } \frac{5\pi}{6} \end{aligned}$$

$$\begin{aligned} R &= \left\{(r,\theta) : a \le r \le 2a\sin\theta, \pi/6 \le \theta \le \frac{5\pi}{6}\right\} \\ A &= \int\int_R 1.dA = \int_{\pi/6}^{5\pi/6}\int_a^{2a\sin\theta} r dr d\theta = \int_{\pi/6}^{5\pi/6} [\frac{r^2}{2}]_a^{2a\sin\theta} d\theta \\ &= \int_{\pi/6}^{5\pi/6} \frac{1}{2}[4a^2\sin^2\theta - a^2]d\theta = \frac{1}{2}[4a^2(\frac{\theta}{2} - \frac{\sin 2\theta}{4}) - a^2\theta]_{\pi/6}^{5\pi/6} \\ &= \frac{1}{2}[(4a^2(\frac{1}{2}\frac{5\pi}{6} - \frac{1}{4}(-\frac{\sqrt{3}}{2})) - a^2\frac{5\pi}{6}) - 4a^2(\frac{1}{2}\frac{\pi}{6} - \frac{1}{4}\frac{\sqrt{3}}{2})] \end{aligned}$$

$$-\frac{a^2\pi}{6}$$

$$= a^2[\pi/3 + \sqrt{3}/2].$$

Example 6.3.4. Evaluate

$$\int\int_R (y-2x)dxdy, R: 1 \le x \le 2, 3 \le y \le 5.$$

Solution. Consider

$P_1 = (x_0, x_1, \ldots, x_m)$ as an partition of $[1,2]$

$P_2 = (y_0, y_1, \ldots, y_n)$ as an partition of $[3,5]$

$P = P_1 \times P_2 = \{(x_i, y_i) : x_i \in P_1, y_i \in P_2\}$ as an partition of R. On each rectangle $R_{ij} : x_{i-1} \le x \le x_i, y_{j-1} \le y \le y_j$, the function $f(x,y) = y - 2x$ has a maximum $M_{ij} = y_j - 2x_{i-1}$ (y maximized and x is minimized) and minimum $m_{ij} = y_{j-1} - 2x_i$ (y minimized and x maximized).

Thus

$$\begin{aligned} U_f(P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}(y_j - 2x_{i-1})\triangle x_i \triangle y_j \\ L_f(P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}(y_{j-1} - 2x_i)\triangle x_i \triangle y_j. \end{aligned}$$

For each pair of integer i and j

$$y_{j-1}2x_i \le \frac{1}{2}(y_j + y_{j-1}) - (x_i + x_{i-1}) \le y_j - 2x_{i-1}$$

for arbitrary P we have

$$L_f(P) \le \sum_{i=1}^{m}\sum_{j=1}^{n}[\frac{1}{2}(y_j + y_{j-1}) - (x_i + x_{i-1})]\triangle x_i \triangle y_j \le U_f(P)$$

or

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{1}{2}(y_j + y_{j-1})\triangle x_i \triangle y_j - \sum_{i=1}^{m}\sum_{j=1}^{n}(x_i + x_{i-1})\triangle x_i \triangle y_j$$

Ist sum

$$\begin{aligned} \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{1}{2}\triangle x_i(y_j^2 - y_{j-1}^2) &= \frac{1}{2}\sum_{i=1}^{m}\triangle x_i(\sum_{j=1}^{n}(y_j^2 - y_{j-1}^2)) \\ &= \frac{1}{2}(2-1)(25-9) = 8. \end{aligned}$$

IInd sum

$$-\sum_{i=1}^{m}\sum_{j=1}^{n}(x_i^2 - x_{i-1}^2)\triangle y_i = -\sum_{i=1}^{m}(x_i^2 - x_{i-1}^2)\sum_{j=1}^{n}\triangle y_j$$
$$= -(4-1)(5-3) = -6.$$

$$I = 8 + (-6) = 2,$$
$$L_f(P) \leq 2 \leq U_f(P)$$

or

$$\int\int(y-2x)dxdy = 2.$$

Example 6.3.5. Integrate $f(x,y) = \sqrt{x^2+y^2}$ over the triangle with vertices $(0,0)(1,0)(1,\sqrt{3})$.

Solution. We have

$$\tan\theta = \sqrt{3} \Rightarrow \theta = \pi/3$$

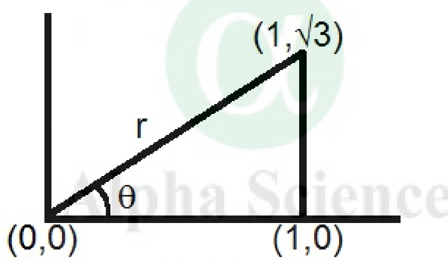

Fig. 6.3.2

$$r\cos\theta = 1, \quad 0 \leq r \leq \frac{1}{\cos\theta}$$

$$I = \int_0^{\pi/3}\int_0^{1/\cos\theta} r.rdrd\theta = \frac{1}{3}\int_0^{\pi/3}\sec^3\theta d\theta$$
$$= \frac{1}{3}[\sec\theta\tan\theta +$$
$$\frac{1}{3}\log|\sec\theta+\tan\theta|]_0^{\pi/3}$$
$$= \frac{1}{\sqrt{3}} + \frac{1}{6}\log[2+\sqrt{3}].$$

Exercises 6.3

1. Integrate $f(x, y) = \cos(x^2 + y^2)$.
   (a) Over closed unit disc.
   (b) the annular region $1 \leq x^2 + y^2 \leq 4$.

2. Let $f(x, y) = \sin(x + y)$ on $R : 0 \leq x \leq 1, 0 \leq y \leq 1$ show that
$$0 \leq \int\int_R \sin(x + y)dxdy \leq 1$$

3. Let $f = f(x, y)$ be continuous on the rectangle $R : a \leq x \leq b, c \leq y \leq d$. Suppose that $L_f(P) = U_f(P)$ for some partition P of R. What can you conclude about f ? what is $\int\int_R f(x, y)dxdy$?

   Hint. $\int\int_R f(x, y)dxdy = f(a, c)(b - a)(d - c)$. f is constant on each sub rectangle.

6.4 Applications to Center of Mass

1. Consider a point mass m located at a point (x, y). The moments of the point mass about the x axis and y axis are

   $\mu_x = my$ and $\mu_y = mx$, respectively.

   (measure of the tendency of the point mass to rotate about the axis)

2. Consider two point masses, m_1 and m_2 located at points (x_1, t) and (x_2, t).

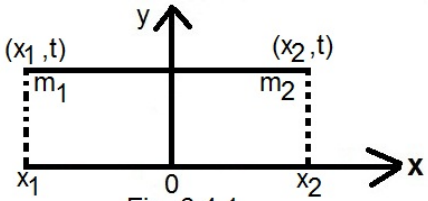

Fig. 6.4.1

   The moments of the masses about both axis:
$$\mu_1 x = m_1 t \quad \mu_2 x = m_2 t$$
$$\mu_1 y = m_1 x_1 \quad \mu_1 y = m_2 x_2$$

if $m_1x_1 + m_2x_2 = 0$ then the point masses are at equilibrium and $(0, t)$ is the point of equilibrium.

3. Suppose several point masses with masses $m_1, m_2, \ldots, m_n$ are located at the points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. The moment of the system of the point masses with respect to axis are

$$\mu_x = m_1y_1 + m_2y_2 + \cdots + m_ny_n$$
$$\mu_y = m_1x_1 + m_2x_2 + \cdots + m_nx_n$$

respectively.

The system is at equilibrium with respect to the x axis (or the y axis) if $\mu_x = 0$ (or $\mu_y = 0$).

Now let $m = m_1 + m_2 + \cdots + m_n$. Then the center of gravity $(\overline{x}, \overline{y})$ of the system is given by

$$\overline{x} = \frac{\mu_y}{m}, \overline{y} = \frac{\mu_x}{m}.$$

Note 6.4.1. The center of mass of a lamina (plate) in the geometrical center (like intersection point of the diagonals of a rectangle).

6.5 Application of Double Integral to Surface Area

Definition 6.5.1. Let R be a vertically or horizontally simple region and let f have continuous partial derivatives on R. The surface area of the portion of the graph of f over R is

$$S = \int\int_R (\sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 + 1})dA.$$

Example 6.5.1. Find the area of the part of the surface that lies over the given region

$$z = x + 2y, R = \{(x, y) : 0 \leq x \leq y, 0 \leq y \leq 2\}.$$

Solution. We have

$$f(x,y) = x + 2y \qquad S = \int\int(\sqrt{f_x^2 + f_y^2 + 1})dA$$
$$f_x(x,y) = 1 \qquad = \int_0^2\int_0^y(\sqrt{1^2 + 2^2 + 1})dxdy = 2\sqrt{6}.$$
$$f_y(x,y) = 2$$

Example 6.5.2. Find the surface area of the portion of the plane $x + y + z = 4$ which is inside the cylinder $x^2 + y^2 = 1$.

Solution. We have to find the area of the shaded region above the circle $x^2 + y^2 = 1$.

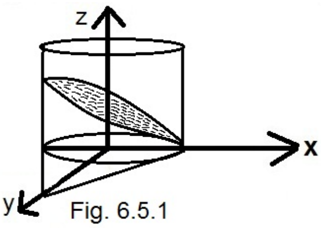

Fig. 6.5.1

We have $f(x,y) = z = 4 - x - y, R = \{(x,y) : -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}, -1 \leq y \leq 1\}$

$$f_x(x,y) = -1, f_y(x,y) = -1$$

$$\begin{aligned} S = \int\int_R(\sqrt{1 + f_x^2 + f_y^2})dxdy &= \int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \sqrt{3}dxdy \\ &= \int_{-1}^{1} 2\sqrt{3}\sqrt{1-y^2}dy \\ &= 2\sqrt{3}\int_{-1}^{1}\sqrt{1-y^2}dy = \pi\sqrt{3}. \end{aligned}$$

Example 6.5.3. Find the area of the portion of the sphere $x^2 + y^2 + z^2 = 16$ lying within the circular parabolid $z = x^2 + y^2$.

Solution. Determine the intersection

$$\begin{aligned} x^2 + y^2 + z^2 &= 16z \\ z &= x^2 + y^2 \\ x^2 + y^2 + (z-8)^2 &= 8^2, \\ z + z^2 &= 16z \\ z(z-15) &= 0 \\ & z = 0, 15, \end{aligned}$$

if $z = 0$ then $x^2 + y^2 = 0$,
$z = 15$ then $x^2 + y^2 = 15$,

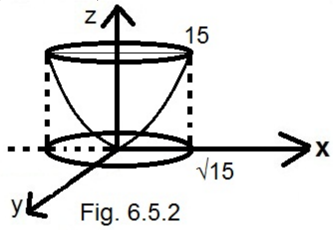

Fig. 6.5.2

$f(x, y) = z = 8 + \sqrt{64 - x^2 - y^2}$
$S = \int\int_R (\sqrt{f_x^2 + f_y^2 + 1})dA$
$f_x(x, y) = \frac{-x}{\sqrt{64-x^2-y^2}} f_y(x, y) = \frac{-y}{\sqrt{64-x^2-y^2}}$
$\sqrt{f_x^2 + f_y^2 + 1} = \sqrt{\frac{64}{64-x^2-y^2}} = \frac{8}{\sqrt{64-x^2-y^2}}$
$S = \int\int_R \frac{8}{\sqrt{64-x^2-y^2}} dxdy = \int_0^{2\pi} \int_0^{\sqrt{15}} \frac{8}{\sqrt{64-r^2}} rdrd\theta.$
Put $u = 64 - r^2, du = -2rdr$

$$= \int_0^{2\pi} \int_{64}^{49} \frac{-4}{\sqrt{u}} dud\theta$$
$$= \int_0^{2\pi} [-8\sqrt{u}]_{64}^{49} d\theta = \int_0^{2\pi} 8d\theta = 16\pi.$$

Exercises 6.5

1. Find the area of the portion of the circular parabolid $z = x^2+y^2$ within the sphere $x^2+y^2+z^2 = 16z$. [Ans. $\left(\frac{61^{3/2}-1}{6}\right)\pi$].

2. Calculate the area of the region enclosed by the curve $r = 2(1 + \sin\theta)$. [Ans. 6π]

6.6 Triple Integral

In R^3 consider the parallelopiped D written as

$$D = \{(x, y, z) : a_1 \le x \le a_2, b_1 \le y \le b_2, c_1 \le z \le c_2\}$$

and

$P = P_1 \times P_2 \times P_3$ `where` $P_1 = (x_1, \ldots, x_m), P_2 = (y_1, \ldots, y_n),$

$P_3 = (z_1, \dots, z_l)$ partition of $[a_1, a_2], [b_1, b_2]$ and $[c_1, c_2]$ respectively. We know that

$$L_f(P) \leq I \leq U_f(P)$$

for any partition of D.

The triple integral of f over D is defined by

$$I = \int\int\int_D f(x,y,z)dv = \int_{a_1}^{a_2}\int_{b_1}^{b_2}\int_{c_1}^{c_2} f(x,y,z)dzdydx.$$

Now let D take the form

$$D = \{(x,y,z) : a_1 \leq x \leq a_2, g_1(x) \leq y \leq g_2(x), h_1(x,y) \leq z \leq h_2(x,y)\}$$

then

$$\int\int\int_D f(x,y,z)dzdydx = \int_{a_1}^{a_2}\int_{g_1(x)}^{g_2(x)}\int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)dzdydx$$

and volume of the region D is given by

$$v = \int\int\int_D 1.dzdydx.$$

Example 6.6.1. Find the volume of the tetrahedron formed by the planes $x = 0, y = 0, z = 0$ and $2x + y + 3z = 6$.

Solution. We have

$$D = \{(x,y,z) : 0 \leq x \leq 3, 0 \leq y \leq 6 - 2x, 0 \leq z \leq \frac{6-2x-y}{3}\}.$$

$$v = \int_0^3\int_0^{6-2x}\int_0^{\frac{6-2x-y}{3}} dzdydx$$

$$= \int_0^3\int_0^{6-2x}[z]\int_0^{\frac{6-2x-y}{3}} dydx = \int_0^3\int_0^{6-2x}\frac{6-2x-y}{3}dydx$$

$$= \int_0^3 \frac{1}{3}[18 - 12x + 2x^2]dx = 6.$$

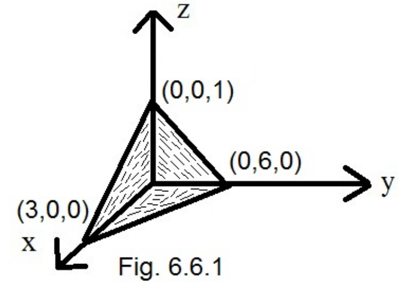

Fig. 6.6.1

Example 6.6.2. Find the volume of the solid that is bounded by the cylinder $y = x^2$ and by the planes $y + z = 4$ and $z = 0$.

Solution. We have

$$D = \{(x, y, z) : -2 \leq x \leq 2, x^2 \leq y \leq 4, 0 \leq z \leq 4 - y\}$$

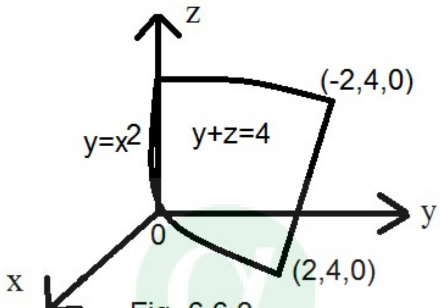

Fig. 6.6.2

$$\begin{aligned} v = \int\int\int_D 1.dv &= \int_{-2}^{2}\int_{x^2}^{4}\int_{0}^{4-y} dzdydx \\ &= (\int_{0}^{4}\int_{-\sqrt{y}}^{\sqrt{y}}\int_{0}^{4-y} dzdxdy) \\ &= \int_{-2}^{2}\int_{x^2}^{4}[z]_0^{4-y}dydx = \int_{-2}^{2}\int_{x^2}^{4}(4-y)dydx \\ &= \int_{-2}^{2}[4y - \frac{y^2}{2}]_{x^2}^{4}dx = \int_{-2}^{2}(8 - 4x^2 + \frac{x^4}{2})dx \\ &= \frac{256}{15}. \end{aligned}$$

6.6.1 Triple Integral in Cylindrical Coordinates

In the cylindrical coordinate system a point is given by $P = (r, \theta, z)$.

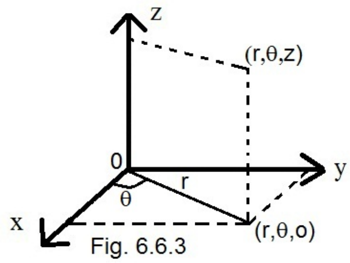

Fig. 6.6.3

$r \geq 0, 0 \leq \theta \leq 2\pi, x = r\cos\theta, y = r\sin\theta, z = z, r = \sqrt{x^2+y^2}$, $\theta = \tan^{-1}\frac{y}{x}$. Consider the cylindrical parallelopiped given by

$$D_c = \{(r,\theta,z) : \theta_1 \leq \theta \leq \theta_2, g_1(\theta) \leq r \leq g_2(\theta), h_1(\sigma,\theta) \leq z \leq h_2(\sigma,\theta)\}$$

and if f is continuous on D_c, then

$$\begin{aligned}\int\int\int_{D_c} f(x,y,z)dxdydz &= \int_{\theta_1}^{\theta_2}\int_{g_1(\theta)}^{g_2(\theta)}\int_{F_1(r\cos\theta, r\sin\theta)}^{F_2(r\cos\theta, r\sin\theta)} f(r\cos\theta, r\sin\theta, z)rdzdrd\theta \\ &= \int_{\theta_1}^{\theta_2}\int_{g_1(\theta)}^{g_2(\theta)}\int_{(r,\theta)}^{(r,\theta)} f(r\cos\theta, r\sin\theta, z)rdzdrd\theta.\end{aligned}$$

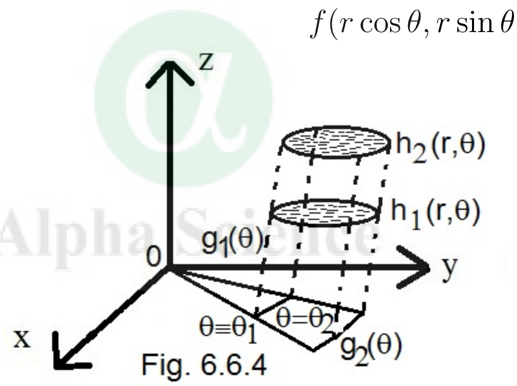

Fig. 6.6.4

$$v = \int_{z_1}^{z_2}\int_{\theta_1}^{\theta_2}\int_{r_1}^{r_2} rdrd\theta dz.$$

For cylindrical transformation, the magnification factor is

$$\begin{aligned}\frac{\partial(x,y,z)}{\partial(r,\theta,z)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} \\ &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.\end{aligned}$$

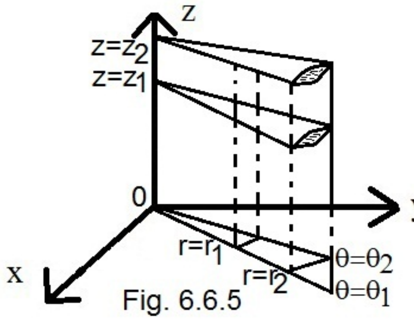

Fig. 6.6.5

Example 6.6.3. Find the volume of the solid bounded by the cone $z = x^2 + y^2$ and the parabolid $4z = x^2 + y^2$.

Solution. Intersection of two surfaces as

$$z^2 = x^2 + y^2 = 4z \Leftrightarrow z(z-4) = 0 \Rightarrow z = 0, z = 4.$$

The intersection is in a point $(0,0,0)$ and in a circle whose projection onto $xy-$plane is $x^2 + y^2 = 16$.

$$\begin{aligned} D_c &= \{(r,\theta,z) : 0 \le r \le 4, 0 \le \theta \le 2\pi, \frac{r^2}{4} \le z \le r\} \\ v &= \int\int\int_{D_c} r dz dr d\theta \\ &= \int_0^{2\pi}\int_0^4 r(r - \frac{r^2}{4}) dr d\theta \\ &= \int_0^{2\pi} \left[\frac{r^3}{3} - \frac{r^4}{16}\right]_0^4 d\theta = \int_0^{2\pi} \frac{16}{3} d\theta = \frac{32\pi}{3}. \end{aligned}$$

Example 6.6.4. Evaluate the integral by changing to cylindrical coordinates

$$\int_0^2 \int_{-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} \int_0^{x^2+y^2} \sqrt{x^2+y^2} dz dy dx.$$

Solution. We have

$$D_c = \{(x,y,z) : 0 \le x \le 2, -\sqrt{2x-x^2} \le y \le \sqrt{2x-x^2}, \\ 0 \le z \le x^2 + y^2\}$$

$$\begin{aligned}
2x - x^2 = y^2 \quad \Leftrightarrow \quad & 2x = x^2 + y^2 \\
& 2r\cos\theta = r^2 \\
& (2\cos\theta - r)r = 0 \\
& r = 0, r = 2\cos\theta \\
D_c \quad = \quad & \{(r,\theta,z) : -\frac{\pi}{2} \le \theta \le \frac{\pi}{2}, 0 \le r\cos\theta, 0 \le z \le r\}
\end{aligned}$$

$$\begin{aligned}
\int\int\int_{D_c} \sqrt{x^2+y^2}\,dzdydx &= \int_{-\pi/2}^{\pi/2}\int_0^{2\cos\theta}\int_0^{r^2} (r)r\,dzdrd\theta \\
&= \int_{-\pi/2}^{\pi/2}\int_0^{2\cos\theta} [r^2 z]_0^{r^2} drd\theta \\
&= \int_{-\pi/2}^{\pi/2}\int_0^{2\cos\theta} r^4 drd\theta \\
&= \int_{-\pi/2}^{\pi/2} \left[\frac{r^5}{5}\right]_0^{2\cos\theta} d\theta \\
&= \frac{64}{5}\int_0^{\pi/2} \cos^5\theta d\theta \\
&= \frac{64}{5}\frac{\sqrt{3}\sqrt{\frac{1}{2}}}{2\sqrt{\frac{7}{2}}} = \frac{512}{75}.
\end{aligned}$$

6.6.2 Triple Integral in Spherical Coordinates

Let (x, y, z) and (r, θ, z) be sets of rectangular and cylindrical coordinates for a point P in space, with $r \ge 0, 0 \le \theta \le 2\pi$.

Transformation from spherical to rectangular coordinates are as

$$\begin{aligned}
z &= \rho\cos\phi = z(\rho,\phi,\theta) \\
r &= \rho\sin\phi \\
y &= r\sin\theta = \rho\sin\phi\sin\theta = y(\rho,\phi,\theta) \\
x &= r\cos\theta = \rho\sin\phi\cos\theta = x(\rho,\phi,\theta) \\
& \quad 0 \le \phi \le \pi.
\end{aligned}$$

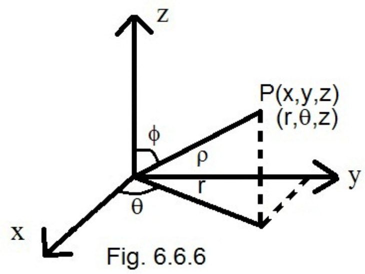

Fig. 6.6.6

In spherical transformation, the magnification factor

$$\begin{aligned}
\frac{\partial(x,y,z)}{\partial(\rho,\phi,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{vmatrix} \\
&= \begin{vmatrix} \sin\phi\cos\theta & \rho\cos\phi\cos\theta & -\rho\sin\phi\sin\theta \\ \sin\phi\sin\theta & \rho\cos\phi\sin\theta & \rho\sin\phi\cos\theta \\ \cos\phi & -\rho\sin\phi & 0 \end{vmatrix} \\
&= \rho^2\sin\phi.
\end{aligned}$$

Let h_1, h_2, F_1, F_2 be continuous functions with $0 \leq \beta - \alpha \leq 2\pi, 0 \leq h_1 \leq h_2 \leq \pi$ and $0 \leq F_1 \leq F_2$ and the region D_S be $D_S = \{(\rho,\phi,\theta) : \alpha \leq \theta \leq \beta, h_1(\theta) \leq \phi \leq h_2(\theta)$ and $F_1(\phi,\theta) \leq \rho \leq F_2(\phi,\theta)\}$ then $\int\int\int_{D_s} f(x,y,z)dv = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} \int_{F_1(\phi,\theta)}^{F_2(\phi,\theta)} f(\rho\sin\phi\cos\theta,$ $\rho\sin\phi\sin\theta, \rho\cos\phi).\rho^2\sin\phi d\rho d\phi d\theta.$

Example 6.6.5. Find the volume of the solid inside the sphere $x^2 + y^2 + z^2 = 9$ and outside the cone $z^2 = x^2 + y^2$.

Solution. We have

$$\begin{aligned}
D_s &= \{(\rho,\phi,\theta) : 0 \leq \rho \leq 3, \frac{\pi}{4} \leq \phi \leq \frac{3\pi}{4}, 0 \leq \theta \leq 2\pi\} \\
v = \int\int_D 1.dv &= \int_{\pi/4}^{3\pi/4}\int_0^{2\pi}\int_0^3 \rho^2\sin\phi d\rho d\theta d\phi \\
&= \int_{\pi/4}^{3\pi/4}\int_0^{2\pi} [\frac{\rho^3}{3}]_0^3 \sin\phi d\theta d\phi \\
&= \int_{\pi/4}^{3\pi/4}\int_0^{2\pi} 9\sin\phi d\theta d\phi = \int_{\pi/4}^{3\pi/4} [9\theta]_0^{2\pi}\sin\phi d\phi \\
&= 18\pi\int_{\pi/4}^{3\pi/4}\sin\phi d\phi = 18\sqrt{2}.
\end{aligned}$$

Example 6.6.6. Evaluate the integral by changing to spherical coordinates.

(a) $\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\int_{\sqrt{x^2+y^2}}^{\sqrt{8-x^2-y^2}}(x^2+y^2+z^2)dzdydx.$

(b) $\int_0^{\sqrt{2}}\int_0^{\sqrt{4-y^2}}\int_0^{\sqrt{4-x^2-y^2}}\sqrt{x^2+y^2+z^2}_0\sqrt{x^2+y^2+z^2}dzdxdy.$

Solution.

(a) We have $D=\{(x,y,z):-2\le x\le 2,-\sqrt{4-x^2}\le y\le \sqrt{4-x^2},\sqrt{x^2+y^2}\le z=\sqrt{8-x^2-y^2}\}$

$$\begin{array}{rcl} x & : & x=-2 \texttt{ to } x=2 \\ y & : & y^2=4-x^2\Rightarrow x^2+y^2=4 \\ z & : & z=\sqrt{x^2+y^2} \texttt{ to } z=\sqrt{8-x^2-y^2} \\ & & z^2=x^2+y^2 \texttt{ to } z^2+x^2+y^2=8 \\ & & \texttt{cone } z\ge 0,\ \begin{array}{c}\rho^2=8\Rightarrow\rho=2\sqrt{2}\\ \texttt{sphere }.\end{array} \end{array}$$

Intersection of the cone and sphere

$8=z^2+x^2+y^2=z^2+z^2=2z^2\Rightarrow z=\pm 2$, since $z\ge 0, z=2$, i.e., the projection of the intersection of the two surfaces is the circle $x^2+y^2=4$,

$D_\rho=\{(\rho,\phi,\theta):0\le\rho\le 2\sqrt{2},0\le\phi\le\pi/4,0\le\theta\le 2\pi\}$,

$f(x,y,z)=x^2+y^2+z^2=\rho^2$,

$$\begin{aligned} &\int\int\int_D(x^2+y^2+z^2)dzdydx \\ =&\int_0^{2\pi}\int_0^{\pi/4}\int_0^{2\sqrt{2}}(\rho^2)\rho^2\sin\phi d\rho d\phi d\theta \\ =&\int_0^{2\pi}\int_0^{\pi/4}[\frac{\rho^5}{5}]_0^{2\sqrt{2}}\sin\phi d\phi d\theta \\ =&\int_0^{2\pi}\int_0^{\pi/4}\frac{128\sqrt{2}}{5}\sin\phi d\phi d\theta \\ =&\frac{-128\sqrt{2}}{5}\int_0^{2\pi}[\cos\phi]_0^{\pi/4}d\theta=\frac{2^8}{5}(\sqrt{2}-1)\pi. \end{aligned}$$

(b) We have

$$\begin{aligned} D &= \{(x, y, z) : 0 \leq y \leq \sqrt{2}, 0 \leq x \leq \sqrt{4 - y^2}, \\ &\quad 0 \leq z \leq \sqrt{4 - x^2 - y^2}\} \\ &\quad x = \sqrt{4 - y^2} \Leftrightarrow x^2 + y^2 = 4 \\ &\quad z = \sqrt{4 - x^2 - y^2} \Leftrightarrow x^2 + y^2 + y^2 = 4 \\ D_s &= \{(\rho, \phi, \theta) : 0 \leq \rho \leq 2, \frac{\pi}{4} \leq \phi \leq \pi/2, 0 \leq \theta \leq \pi/2\}. \end{aligned}$$

Here y takes the place of z, x takes the place of y and z takes the place of x, hence $y = \rho \cos \phi, x = \rho \sin \phi \sin \theta, z = \rho \sin \phi \cos \theta, x^2 + y^2 + z^2 = \rho^2$.

$$\begin{aligned} \int\int\int_{D_s} \sqrt{x^2 + y^2 + z^2} dv &= \int_0^{\pi/2} \int_{\pi/4}^{\pi/2} \int_0^2 (\rho) \rho^2 \sin \phi d\rho d\phi d\theta \\ &= \int_0^{\pi/2} \int_{\pi/4}^{\pi/2} [\frac{\rho 4}{4}]_0^2 \sin \phi d\phi d\theta \\ &= \int_0^{\pi/2} \int_{\pi/4}^{\pi/2} 4 \sin \phi d\phi d\theta = \pi\sqrt{2}. \end{aligned}$$

6.6.3 Change of Variables in Multiple Integrals

Let R be a region in the $xy-$plane, and suppose that x and y are functions of a new set of variables u and ν, i.e.,

$$x = g_1(u, \nu), y = g_2(u, \nu)$$

where g_1 and g_2 have continuous partial derivatives in some region S in the $u\nu$ plane ($(x, y) \in R$).

In the above transformation assume that each point (u, ν) in S is taken to exactly one point (x, y) in R.

Then

$$\int\int_R f(x, y) dA = \int\int_S f(g_1(u, \nu), g_2(u, \nu)) |\frac{\partial(x, y)}{\partial(u, \nu)}| dA$$

where

$$\frac{\partial(x,y)}{\partial(u,\nu)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial \nu} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial \nu} \end{vmatrix}$$

is the area magnification factor and is called the Jacobian of the transformation.

Example 6.6.7. Evaluate $\int\int_R (y-x)dA$, where R is the region bounded by the lines $y = x+1, y = 2-x$, and $y = 2x - y$.

Solution.

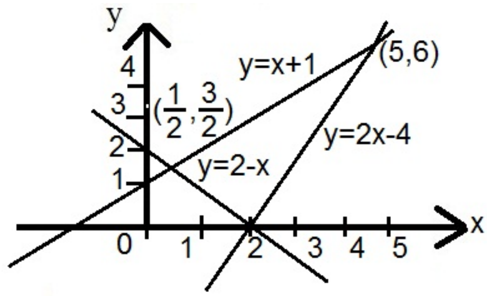

Fig. 6.6.7

$$u = y - x$$

$$\nu = y + x.$$

Then the lines $y = x+1, y = 2-x$ are transformed into the lines $u = 1, \nu = 2$.

Since

$$\begin{aligned} u + \nu &= y - x + y + x = 2y \\ u - \nu &= y - x - y + x = -2x \\ x &= -\frac{u}{2} + \frac{\nu}{2}, \quad y = 2x - y \\ y &= \frac{u}{2} + \frac{\nu}{2}, \quad \frac{u+\nu}{2} = -u + \nu - y \\ &\nu = 3u + 8, \\ \frac{\partial(x,y)}{\partial(u,\nu)} &= \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial \nu} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial \nu} \end{vmatrix} = \begin{vmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{vmatrix} = -\frac{1}{2}. \end{aligned}$$

$$y - x = \frac{u}{2} + \frac{\nu}{2} - \left(-\frac{u}{2} + \frac{\nu}{2}\right) = u,$$

$$\int\int_R (y-x)dA = \int\int_S u| - \frac{1}{2}|dA = \int_{-2}^{1}\int_{2}^{3u+8} \frac{u}{2} d\nu du$$
$$= \int_{-2}^{1} [\frac{u\nu}{2}]_2^{3u+8} du = 0.$$

Example 6.6.8. Evaluate $\int\int_R (x-y)^2 \sin^2(x+y)dA$ by suitable change of variables where R is the region bounded by the parallelogram with vertices $(\pi,0)(2\pi,\pi),(\pi,2\pi)$ and $(0,\pi)$.

Solution. Let the transformation be

$$u = x - y, \nu = x + y.$$

Then if

$$y = x - \pi \text{ then } u = x - x + \pi = \pi$$
$$y = \pi - x \text{ then } v = x + \pi - x = \pi$$
$$y = x + \pi \text{ then } u = x - (x + \pi) = -\pi$$
$$y = 3\pi - x \text{ then } \nu = x + (3\pi - x) = 3\pi.$$

Now

$$(x-y)^2 = u^2, \sin^2(x+y) = \sin^2\nu,$$

then

$$\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial \nu}{\partial x} & \frac{\partial \nu}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2,$$

$$\int\int_R (x-y)^2(x+y)dA = \int\int_S u^2 \sin^2\nu \left|\frac{\partial(x,y)}{\partial(u,\nu)}\right| dA$$
$$= \int_{\pi}^{3\pi}\int_{-\pi}^{\pi} u^2 \sin^2\nu |\frac{1}{2}| du d\nu = \frac{\pi^4}{3}.$$

Note 6.6.1. When solving for x and y in the integrand use the relation

$$\frac{\partial(x,y)}{\partial(u,\nu)} = \frac{\frac{1}{\partial(u,\nu)}}{\partial(x,y)}.$$

Example 6.6.9. Evaluate $\int\int_R (x^2+y^2)dxdy$, where R is a region bounded by $x^2 - y^2 = 1, x^2 - y^2 = 9, xy = 2, xy = 4$.

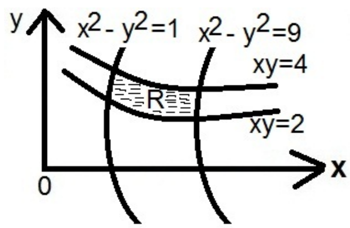

Fig. 6.6.8

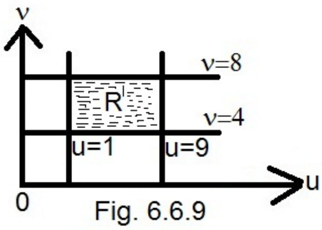

Fig. 6.6.9

Solution. Let $x^2 - y^2 = u, 2xy = \nu$

$$\int\int_R (x^2+y^2)dxdy = \int\int_{R^1}(x^2+y^2)\left|\frac{\partial(x,y)}{\partial(u,\nu)}\right| dud\nu$$

$$\frac{\partial(u,\nu)}{\partial(x,y)} = \begin{vmatrix} u_x & u_y \\ \nu_x & \nu_y \end{vmatrix} = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix} = 4(x^2+y^2),$$

$$\begin{aligned} (x^2+y^2)^2 &= (x^2-y^2)+(2xy)^2 \\ &= u^2+\nu^2 \\ (x^2+y^2) &= \sqrt{u^2+\nu^2} \end{aligned}$$

$$\frac{\partial(x,y)}{\partial(u,\nu)} = \frac{1}{\frac{\partial(u,\nu)}{\partial(x,y)}} = \frac{1}{4(x^2+y^2)} = \frac{1}{4\sqrt{u^2+\nu^2}},$$

$$\begin{aligned} &\int\int_{R^1}\sqrt{u^2+\nu^2}\frac{dud\nu}{4\sqrt{u^2+\nu^2}} \\ &= \frac{1}{4}\int_{u=1}^{9}\int_{\nu=4}^{8} dud\nu = 8. \end{aligned}$$

Remark 6.6.1. If a transformation is defined by

$$x = g_1(u,\nu,w), y = g_2(u,\nu,w), z = g_3(u,\nu,w)$$

and it maps a region E in u,ν,w space onto a region D in xyz space, then

$$\begin{aligned} \int\int\int_D f(x,y,z)dv &= \int\int\int_E f(g_1(u,\nu,w), g_2(u,\nu,w), \\ &\quad g_3(u,\nu,w))\left|\frac{\partial(x,y,z)}{\partial(u,\nu,w)}\right| dv, \end{aligned}$$

where

$$\frac{\partial(x,y,z)}{\partial(u,\nu,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial \nu} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial \nu} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial \nu} & \frac{\partial z}{\partial w} \end{vmatrix}$$

is the Jacobian of the transformation.

Remark 6.6.2. Suppose T is an object in the form of basic solid. If T has constant mass density λ, then the mass of T is $M = \lambda v$. If the mass density varies continuously over T say $\lambda = \lambda(x,y,z)$ then

$$M = \int\int\int_T \lambda(x,y,z)dxdydz,$$

coordinates of centroid are given by

$$\overline{x}v = \int\int\int_T xdxdydz, \overline{y}v = \int\int\int_T ydxdydz,$$

and

$$\overline{z}v = \int\int\int_T zdxdydz,$$

moment of inertia of T about a line is given by

$$I = \int\int\int_T \lambda(x,y,z)[r(x,y,z)]^2 dxdydz$$

where $r(x,y,z)$ is the distance of (x,y,z) from the line.

Example 6.6.10. Find the mass of a solid right circular cylinder of radius r and height H given that the mass density is directly proportional to the distance from the lower base.

Solution. We have

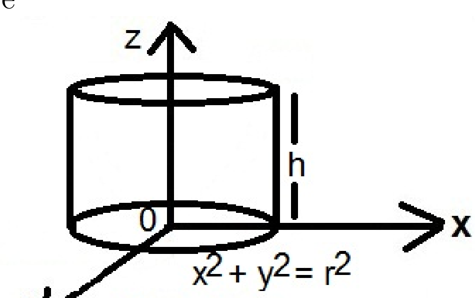

Fig. 6.6.10

$-r \leq x \leq r, -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}, 0 \leq z \leq h, \lambda(x, y, z) = Kz, K > 0 M = \int\int\int Kzdxdydz$

$$\begin{aligned} &= \int_{-r}^{R} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \int_{0}^{h} Kzdxdydz \\ &= \int_{-r}^{r} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} K\frac{h^2}{2}dydx \\ &= 4\int_{0}^{r} \int_{0}^{\sqrt{r^2-x^2}} \frac{Kh^2}{2}dydx \\ &= 2Kh^2 \int_{0}^{r} \sqrt{r^2 - x^2}dx. \end{aligned}$$

Putting $x = r\sin\theta, dx = r\cos\theta d\theta$,

$$\begin{aligned} &= 2Kh^2 \int_{0}^{\pi/2} r^2 \cos^2\theta d\theta, \\ &= \frac{Kh^2r^2\pi}{2}. \end{aligned}$$

Example 6.6.11. Set $f(x, y, z) = xyz$ on $II : 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1$ and take P as the partition $P_1 \times P_2 \times P_3$.

(a) Find $L_f(P)$ and $U_f(P)$ given that
$P_1 = (x_0, x_1, \ldots, x_m), P_2 = (y_0, y_1, \ldots, y_n), P_3 = (z_0, z_1, \ldots, z_q)$ are all arbitrary partitions of $[0, 1]$.

(b) Use your answer to (a) to calculate

$$\int\int\int_{II} xyzdxdydz.$$

Solution.

(a) Let

$$\begin{aligned} L_f(P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{q} x_{i-1}y_{j-1}z_{k-1}\triangle x_i \triangle y_j \triangle z_k, \\ U_f(P) &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{q} x_i y_j z_k \triangle x_i \triangle y_j \triangle z_k. \end{aligned}$$

(b) $x_{i-1}y_{j-1}z_{k-1} \leq \left(\frac{x_i+x_{i-1}}{2}\right)\left(\frac{y_j+y_{j-1}}{2}\right)\left(\frac{z_k+z_{k-1}}{2}\right) \leq x_i y_j z_k$
$x_{i-1}y_{j-1}z_{k-1}\triangle x_i \triangle y_j \triangle z_k \leq \frac{1}{8}\left(x_i^2 - x_{i-1}^2\right)\left(y_j^2 - y_{j-1}^2\right)$
$\left(z_k^2 - z_{k-1}^2\right) \leq x_i y_j z_k \triangle x_i \triangle y_j \triangle z_k,$
$L_f(P) \leq \frac{1}{8}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{q}(x_i^2 - x_{i-1}^2)(y_j^2 - y_{j-1}^2)(z_k^2 - z_{k-1}^2)$
$\leq U_f(P)$.

The middle term can be written as
$\frac{1}{8}\sum_{i=1}^{m}(x_i^2-x_{i-1}^2)\sum_{j=1}^{n}(y_j^2-y_{j-1}^2)\sum_{k=1}^{n}(z_k^2-z_{k-1}^2) = \frac{1}{8}.1.1.1 = \frac{1}{8}$.

$\Rightarrow I = \frac{1}{8}$.

Example 6.6.12. Find the mass of the solid bounded above by the parabolic cylinder $z = u - y^2$ and bounded below by the elliptic parabolid $z = x^2 + 3y^2$, given that the density varies directly with $|x|$.

Solution. We have

$$\begin{aligned}
4 - y^2 &= x^2 + 3y^2 \Rightarrow 4y^2 = 4 - x^2, \\
M &= \int_{-2}^{2}\int_{-\sqrt{\frac{4-x^2}{2}}}^{\sqrt{\frac{4-x^2}{2}}}\int_{x^2+3y^2}^{4-y^2} k|x|dzdydx \\
&= 4\int_{0}^{2}\int_{0}^{\sqrt{\frac{4-x^2}{2}}}\int_{x^2+3y^2}^{4-y^2} Kxdzdydx \\
&= 4K\int_{0}^{2}\int_{0}^{\sqrt{\frac{4-x^2}{2}}}(4x - x^3 - 4xy^2)dydx = \frac{128}{15}K.
\end{aligned}$$

Exercises 6.6

1. Express $\int\int\int_D f(x,y,z)dv$ as an iterated integral where D is the region in the first octant bounded by the coordinate planes and the graphs of $z - 2 = x^2 + \frac{y^2}{4}$ and $x^2 + y^2 = 1$.

   Hint. $D = \{(x,y,z) : 0 \leq x \leq 1, 0 \leq y \leq \sqrt{1-x^2}, 0 \leq z \leq x^2 + \frac{y^2}{4} + 2\}$

$$\int\int\int_D f(x,y,z)dv = \int_0^1\int_0^{\sqrt{1-x^2}}\int_0^{2+x^2+\frac{y^2}{4}} f(x,y,z)dzdydx.$$

2. Evaluate $\int\int\int_D xzdv$, where D is the solid region in the first octant bounded above by the sphere $x^2+y^2+z^2=4$, below by the plane $z=0$ and on the sides by the planes $x=0, y=0$ and the cylinder $x^2+y^2=1$.

3. Express $\int\int\int_D (x^2+y^2)dv$ as an iterated integral in cylindrical coordinates and evaluate it, where D is the solid region bounded by the cylinder $x^2+y^2=1$ and the planes $z=0$ and $z=4$.

   Hint. $D=\{(x,\theta,z): 0\le r\le 1, 0\le\theta\le 2\pi, 0\le z\le 4\}$, Ans. 2π.

4. Express $\int\int\int_D f(x,y,z)dv$ as an iterated integral, where D is the region in the I^{st} octant bounded by coordinate plane and the graphs of $z-2=x^2+\frac{y^2}{4}$ and $x^2+y^2=1$.

5. Find the volume of the solid that is bounded by the cylinder $y=x^2$ and by the planes $y+z=4$ and $z=0$.

6. Find the mass and center of mass of a solid hemisphere of radius a if λ at P is directly proportional to the distance from the center of the base to P.

   Hint. Center of mass is $\frac{29}{5}$ from base along the axis of symmetry, Ans. $\frac{Ka^4\pi}{2}$.

7. Evaluate the integrals by changing to cylindrical coordinates

   (a) $\int_0^1\int_0^{\sqrt{1-y^2}}\int_0^{\sqrt{4-x^2-y^2}} zdzdxdy$.

   (b) $\int_0^2\int_{-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}}\int_0^{x^2+y^2} \sqrt{x^2+y^2}dzdydx$.

8. Evaluate the integrals by changing to spherical coordinates

   (a) $\int_{-2}^2\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\int_{\sqrt{x^2+y^2}}^{\sqrt{8-x^2-y^2}}(x^2+y^2+z^2)dzdydx$.

   (b) $\int_0^{\sqrt{2}}\int_0^{\sqrt{4-y^2}}\int_0^{\sqrt{4-x^2-y^2}}\sqrt{x^2+y^2+z^2}dzdxdy$.

9. Sketch the region, use triple integrals and find v.

   (a) $z+x^2=4, y+z=4, y=0, z=0$, Ans. $\frac{128}{5}$.

   (b) $y=2-z^2, y=z^2, x+z=4, x=0$, Ans. $\frac{32}{3}$.

(c) $y^2+z^2=1, x+y+z=2, x=0$, Ans. 2π.

(d) $z=9-x^2, z=0, y=-1, y=2$, Ans. 108.

(e) $z=x^2, z=x^3, y=z^2, y=0$, Ans. $\frac{1}{70}$.

10. A region S is bounded by the surfaces $x^2+y^2-2x=0, 4z=x^2+y^2, z^2=x^2+y^2$, use cylindrical coordinates to find the volume $v(S)$.

    Hint. $v=\int\int\left[\sqrt{x^2+y^2}-\frac{x^2+y^2}{4}\right]dA=\int_{-\pi/2}^{\pi/2}\int_0^{2\cos\theta}(r^2-\frac{r^3}{4})drd\theta=\frac{32}{9}-\frac{3\pi}{8}$.

11. A region F above the x axis is bounded on the left by the line $y=-x$, and on the right by the curve $c: x^2+y^2=3\sqrt{x^2+y^2}-3x$, find the area.

    Hint. c : the cardoid $r=3(1-\cos\theta), y=-x$: the ray $\theta=\frac{3\pi}{4}, G=\{(r,\theta): 0\le r\le 3(1-\cos\theta), 0\le\theta\le\frac{3\pi}{4}\}, A(F)=\int\int_F dA=\int\int_G rdrd\theta=\int_0^{\pi/4}\int_0^{3(1-\cos\theta)} rdrd\theta=\frac{9}{8}(\frac{9\pi}{2}-4\sqrt{2}-1)$.

12. Use polar coordinates to evaluate $\int\int_F\sqrt{x^2+y^2}dA$, where F is the region inside the circle $x^2+y^2=2x$.

    Hint. $\int\int_F\sqrt{x^2+y^2}dA=\int_{-\pi/2}^{\pi/2}\int_0^{2\cos\theta} r^2drd\theta=\frac{32}{9}$.

13. Evaluate using polar coordinates

    (a) $\int_0^2\int_0^{\sqrt{4-y^2}}\sqrt{x^2+y^2}dxdy$, Ans. $\frac{4\pi}{3}$.

    (b) $\int_{-\sqrt{\pi}}^{\sqrt{\pi}}\int_{-\sqrt{\pi-y^2}}^{\sqrt{\pi-y^2}}\sin(x^2+y^2)dxdy$, Ans. 2π.

    (c) $\int_{-2}^{2}\int_{2-\sqrt{4-x^2}}^{2+\sqrt{4-x^2}}\sqrt{16-x^2-y^2}dydx$, Ans. $\frac{64\pi}{3}-\frac{256}{9}$.

    (d) $\int_0^1\int_y^{\sqrt{y}}(x^2+y^2)^{-1/2}dxdy$, Ans. $\sqrt{2}-1$.

14. Find the volume of the solid bounded above by plane $2z=4+x$ below by the $xy-$plane, and on the sides by the cylinder $x^2+y^2=2$.

    Hint. $r^2=2r\cos\theta, r=2\cos\theta, \cos\theta=0, \theta=\pm\pi/2, z+r/2\cos\theta$,

    $v=\int\pi/2_{???}\int_0^{2\cos\theta}\int_0^{2+r/2\cos\theta} rdzdrd\theta=\frac{5}{0}\pi$.

15. Find the volume of the solid that lies between the cylinder $x^2+y^2=1$ and $x^2+y^2=4$, and is bounded above by the ellipsoid $x^2+y^2+4z^2=36$ and below by the xy plane.

    Hint. $v=\int_0^{2\pi}\int_1^2\int_0^{\frac{1}{2}\sqrt{36-r^2}} r\,dz\,dr\,d\theta=\frac{\pi}{3}(35\sqrt{35}-128\sqrt{2})$.

16. Find the volume of the solid T enclosed by the surface $(x^2+y^2+z^2)^2=2z(x^2+y^2)$.

    Hint. $\rho^4=2\rho\cos\phi\rho^2\sin^2\phi\Rightarrow\rho=2\cos\phi\sin^2\phi$, here no restriction on θ thus $0\le\theta\le 2\pi$. Since ρ remains nonnegative, ϕ can range only $0\le\phi\le\pi/2, 0\le\rho\le 2\sin^2\phi\cos\theta$. $v=\int_0^{2\pi}\int_0^{\pi/2}\int_0^{2\sin^2\phi\cos\theta}\rho^2\sin\phi d\rho d\phi d\theta=\frac{2\pi}{15}$.

17. Find the mass of a right circular cone of radius r and height h given that the density varies directly with the distance from vertex.

    Hint. $D_s=\{(\rho,\phi,\theta):0\le\theta\le 2\pi, 0\le\phi\le\tan^{-1}r/h, 0\le\rho\le h\sec\phi\}$
    $M=\int\int\int K\rho^3\sin\phi d\rho d\phi d\theta$. Ans. $\frac{1}{6}K\pi h[(r^2+h^2)^{3/2}-h^3]$.

18. Locate the center of mass of the upper half ball, center of the ball is at the origin.

    Hint. $\overline{z}v=\int_0^{2\pi}\int_0^{\pi/2}\int_0^R(\rho\cos\phi)\rho^2\sin\phi d\rho d\phi d\theta=\frac{\pi R^4}{4}, v=\frac{2}{3}\pi R^3$, $\overline{z}=\frac{3}{8}R, (\overline{x},\overline{y},\overline{z})=(0,0,\frac{3R}{8})$.

19. Let Ω be the first quadrant region bounded by the curve $xy=1, xy=4, y=x, y=4x$.

    (a) Determine the area of Ω.

    (b) Locate the centroid.

    Hint.

    (a) $h=xy,\nu=y$, then $x=u/\nu, y=\nu, J(u,\nu)=1/\nu.xy=1, xy=4\Rightarrow u=1, u=4, y=x, y=4x\Rightarrow\nu^2=u,\nu^2=4u$. Ω is the set of all (x,y) with $u\nu$ coordinates in the set $\Gamma:1\le u\le 4,\sqrt{u}\le\nu\le 2\sqrt{u}, \int_1^4\int_{\sqrt{u}}^{2\sqrt{u}}\frac{1}{\nu}dud\nu=3\log 2$.

    (b) $\overline{x}A=\int_1^4\int_{\sqrt{u}}^{2\sqrt{u}}\frac{u}{\nu^2}d\nu du=\frac{7}{3},\overline{x}=\frac{7}{9\log 2},\overline{y}A=\int\int d\nu du=\frac{14}{3},\overline{y}=\frac{14}{9\log 2}$.

20. Show that the ellipse $b^2x^2 + a^2y^2 = a^2b^2$ has area πab setting $x = ar\cos\theta, y = br\sin\theta$.

    Hint. $J(r,\theta) = abr, 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi, A = ab\int_0^{2\pi}\int_0^1 rdrd\theta = \pi ab.$

21. Let T be the solid ellipsoid $x^2/a^2+y^2/b^2+z^2/c^2 \leq 1$, calculate volume of T by setting $x = a\rho\sin\phi\cos\theta, y = b\rho\sin\phi\sin\theta, z = c\rho\cos\phi$.

    Hint. $J = abc\rho^2\sin\phi, v = \int_0^{2\pi}\int_0^{\pi}\int_0^1 abc\rho^2\sin\phi d\rho d\phi d\theta = \frac{4}{3}\pi abc.$

Chapter 7

Integration

In this chapter we have studied the integral of real-valued function of several real variables, and its properties.The integral studied here is called Riemann integral; it is a direct generalization of the integral usually studied in a first course in single-variable analysis.

7.1 Basic Definitions

Definition 7.1.1. Let $I = [a, b] = [a^1, b^1] \times [a^2, b^2] \times \cdots \times [a^n, b^n]$ be a closed rectangle in R^n. Then the volume of I is defined by

$$V(I) = \prod_{i=1}^{n} (b^i - a^i).$$

Definition 7.1.2. Let $I = [a, b]$ be a closed rectangle in R^n. Then a partition P of I is a set $P = \{I_1, I_2, \ldots, I_n\}$ of closed sub rectangles such that $\cup_{i=1}^{n} I_i = I$ and I_i and I_j intersects at most on their boundaries.

Note that $P = \{I_1, I_2, \ldots, I_n\}$ is a partition of I (as in figure below) then

$$V(I) = \sum_{i=1}^{n} V(I_i).$$

Now suppose that $f : I \to R$ is a bounded function where I is a closed rectangle in R^n. Let $P = \{I_1, \ldots, I_n\}$ be a partition of I for each $i =, \ldots, n$, define

$$m_i = \inf_{x \in I_i} f(x), \quad M_i = \sup_{x \in I_i} f(x).$$

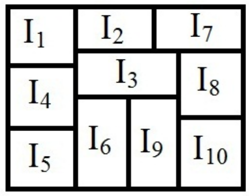

Fig. 7.1.1

($m_i \& M_i$ exists since function is bounded).

Then we define the upper and lower sum of f with respect to P respectively as

$$U_f(P) = \sum_{i=1}^{n} M_i V(I_i), \ L_f(P) = \sum_{i=1}^{n} m_i V(I_i)$$

clearly $L_f(P) \leq U_f(P)$.

Definition 7.1.3. Let P and P' be partitions of an rectangle I, we say P' is a refinement of P.

Lemma 7.1.1. Suppose P' is a refinement of P. Then

$$L_f(P) \leq L_f(P') \text{ and } U_f(P') \leq U_f(P).$$

Proof. Let I_i be a sub rectangle in P, then $I_i = \cup I_{i_j}$ where I_{i_j}'s are sub rectangles in P'.

$$\begin{aligned}
&\Rightarrow \quad m_i \leq m_{i_j} \text{ and } M_{i_j} \leq M_i \\
&\Rightarrow \quad \sum_j m_i V(J_{i_j}) \leq \sum_j m_{i_j} V(J_{i_j}) \\
&\Rightarrow \quad m_i V(I_i) \leq \sum_j m_{i_j} V(J_{i_j}) \\
&\Rightarrow \quad \sum_i m_i V(I_i) \leq \sum_i \sum_j m_{i_j} V(J_{i_j}) \\
&\Rightarrow \quad L_f(P) \leq L_f(P').
\end{aligned}$$

Similarly using $M_{i_j} \leq M_i$, one can show that $U_f(P') \leq U_f(P)$.

Corollary 7.1.1. If P and P' are any two partitions then $L_f(P) \leq$

$U_f(P')$.

Proof. Let P'' be a refinement of P and P'. Then by Lemma 7.1.1

$$L_f(P) \leq L_f(P'') \texttt{ and } U_f(P'') \leq U_f(P') \Rightarrow L_f(P) \leq U_f(P').$$

Note. From the above corollary it follows that the set of all upper sums is bounded below and hence it has infimum, defined by

$$\overline{\int}_I \, i.e., \overline{\int}_I f = \inf\{U_f(P) : P \texttt{ is a partition of } I\}.$$

Similarly the set of lower sums is bounded above and we define the lower integral by

$$\underline{\int}_I f = \sup\{L_f(P) : P \texttt{ is a partition of } I\}.$$

One can easily see that $\overline{\int}_I f \leq \overline{\int}_I f$.

Definition 7.1.4. A bounded function f from an interval I to R is said to be integrable if and only if

$$\underline{\int}_I f = \overline{\int}_I f \quad \texttt{or} \quad \int_I f.$$

Theorem 7.1.1. A bounded function $f : I \to R$ is integrable if and only if $\forall \varepsilon > 0$, there exists P a partition of I such that $U_f(P) - L_f(P) < \varepsilon$.

Proof. Suppose f is integrable. Then $\sup\{L_f(P) : P$ a partition$\} = \inf\{U_f(P) : P$ a partition $\}$. Let l be the common value, then there exists P_1 a partition such that $l - \frac{\varepsilon}{2} < L_f(P_1)$ and P_2 a partition such that $U_f(P_2) < l + \frac{\varepsilon}{2}$.

Let P be a refinement of P_1 and P_2.

$$l - \frac{\varepsilon}{2} < L_f(P_1) \leq L_f(P) \leq U_f(P) \leq U_f(P_2) < l + \frac{\varepsilon}{2}$$
$$\Rightarrow U_f(P) - L_f(P) < \varepsilon.$$

Now suppose $\forall \varepsilon > 0$, there exists $P : U_f(P) - L_f(P) < \varepsilon$.

Note that $\overline{\int}_I f \leq U_f(P)$ and $L_f(P) \leq \underline{\int}_I f$

$\overline{\int}_P f - \underline{\int}_I f \leq U_f(P) - L_f(P) < \varepsilon$, since ε was arbitrary $\Rightarrow \overline{\int} f = \underline{\int} f \Rightarrow f$ is integrable.

Exercises 7.1

1. Suppose $f : Q \to R$ is continuous. Show f is integrable over Q.

2. Let $[0,1]^2 = [0,1] \times [0,1]$. Let $f : [0,1]^2 \to R$ be defined by setting $f(x,y) = 0$ if $y \neq x$, and $f(x,y) = 1$ if $y = x$. Show that f is integrable over $[0,1]^2$.

3. Let $f : R \to R$ be defined by setting $f(x) = 1/q$ if $x = p/q$, where p and q are positive integers with no common factor, and $f(x) = 0$ otherwise. Show f is integrable over $[0,1]$.

4. For the rectangle $I = [0,1] \times [0,1]$ in the plane R^2, define

$$f(x,y) = \begin{cases} 5 & \texttt{if } (x,y) \texttt{ is in } I \texttt{ and } x > 1/2 \\ 1 & \texttt{if } (x,y) \texttt{ is in } I \texttt{ and } x \leq 1/2. \end{cases}$$

   Use the Integrability Criterian to show that the function $f : I \to R$ is integrable.

5. Let I be a generalized rectangle in R^n and suppose that the function $f : I \to R$ assumes the value 0 except at a single point x in I. Show that $f : I \to R$ is integrable. Then show that $\int_I f = 0$.

6. Let I be a generalized rectangle in R^n and suppose that the bounded function $f : I \to R$ has the value 0 on the interior of I. Show that $f : I \to R$ is integrable and that $\int_I f = 0$.

7. Let I be a generalized rectangle in R^n and suppose that the function $f : I \to R$ is integrable. Let the number M have the property that $|f(x)| \leq M$ for all x in I. Prove that

$$|\int_I f| \leq M. \quad \texttt{vol I.}$$

7.2 Measure Zero and Content Zero

A subset $A \subset R^n$ is said to be a set of measure zero if $\forall \varepsilon > 0$, there is a countable collection $\{I_n\}$ of closed rectangles such that $A \subset \cup I_n$ and $\sum_n V(I_n) < \varepsilon$.

Example 7.2.1. A subset of a set of measure zero is a set of measure zero.

Example 7.2.2. A one point set $\{u\}$ is a set of measure zero.

Note. In the definition of a set of measure zero, the closed rectangle can be replace by open or half open, half closed rectangle.

Theorem 7.2.1. If $A_1, A_2, \dots, A_n$ are set of measure zero, then $A = A_1 \cup A_i$ is a set of measure zero.

Proof. For each i, since A_i is a set of measure zero, there exists $\{I_{i_j}\}_{j=1}^{\infty}$ such that $A_i \subset \cup_j I_{i_j}$ and $\sum_j V(I_{i_j}) < \frac{\varepsilon}{2^i}$.

But $A \subset \cup I_{i_j}$ and $\sum_{i_j} V(I_{i_j}) = \sum_i \sum_j V(I_{i_j}) < \sum_i \frac{\varepsilon}{2^i} = \varepsilon \Rightarrow A$ is a set of measure zero.

Example 7.2.3. Any countable subset of R^n is a set of measure zero.

Proof. Let $A = \{u_1, u_2, \dots\}$ be a countable subset of R^n. Then $A = \cup A_i$ where $A_i = \{u_i\}$, since each A_i is a set of measure zero by the above Theorem 7.2.1 so is A.

Definition 7.2.1. A subset A of R^n is said to be a set of content zero if $\forall \varepsilon > 0$, there exist $I_1, I_2, \dots, I_n$ rectangles such that $A \subset I_1 \cup \cdots \cup I_n$ and $\sum_i V(I_i) < \varepsilon$.

Clearly any set of content zero is of measure zero.

Theorem 7.2.2. If A is compact and has measure zero then A has content zero.

Proof. Let A be compact set and of measure zero. Let $\varepsilon > 0$, there exists $\{I_i\}$ open rectangle such that $A \subset \cup_{i=1}^{\infty} I_i$ and $\sum_{i=1}^{\infty} V(I_i) < \varepsilon$.

But since A is compact there exists n such that

$$A \subset \cup_{i=1}^{n} I_i, \ \texttt{also} \ \sum_{i=1}^{n} V(I_i) \leq \sum_{i=1}^{\infty} V(I_i) < \varepsilon$$

$\Rightarrow A$ is a set of content zero.

Exercises 7.2

1. Show that an unbounded set can not have content zero.

2. Q^n is a set of measure zero in R^n.

3. Show that if C is a set of content zero boundary$C$ is also of content zero.

4. Show that if A has measure zero in R^n, the set $\overline{A}$ and boundary A need not have measure zero.

5. Show that no open set in R^n has measure zero in R^n.

6. Show that the set $R^{n-1} \times 0$ has measure zero in R^n.

7. Let $f : [a, b] \to R$. The graph of f is the subset

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

   of R^2. Show that if f is continuous, G_f has measure zero in R^2.

8. Show that the set of irrationals in $[0, 1]$ does not have measure zero in R.

7.3 Integrable Functions

In this section we shall see the condition for integrability of a bounded function on a closed rectangle.

Lemma 7.3.1. Let I be a closed rectangle and $f : I \to R$. If $O(f, x) < \varepsilon, \forall x \in I$. Then there exists a partition P of I such that $U_f(P) - L_f(P) < \varepsilon V(I)$.

Proof. For each $x \in I$, there exists I_x a closed rectangle which contains x in its interior and $M_{I_x}(f) - m_{I_x}(f) < \varepsilon (M_{I_x} = \sup_{t \in I_x} f(t), m_{I_x} = \inf_{t \in I_x} f(t))$, since I is compact , finite rectangles $I_{x_1}, I_{x_2}, \ldots, I_{x_n}$ covers it. Let P be a partition of I where each sub rectangle S in P is contained in some I_{x_i}. Then for every $S \in P, M_S(f) - m_S(f) < \varepsilon$

$$\begin{aligned} \Rightarrow U_f(P) - L_f(P) &= \sum_{S \in P} [M_S(f) - m_S(f)] V(S) \\ &< \varepsilon \sum_{S \in P} V(S) = V(I)\varepsilon. \end{aligned}$$

Theorem 7.3.1. Let I be a closed rectangle and $f : I \to R$ in bounded function. Let $B = \{x : f \text{ is not continuous at } x\}$. Then f is integrable if and only if B is a set of measure zero.

Proof. Suppose that B is a set of measure zero. Let $\varepsilon > 0$ and $B_\varepsilon = \{x : O(f, x) \geq \varepsilon\}$. Then $B_\varepsilon \subset B$ and hence a set of measure zero. But B_ε is closed. Thus B_ε is a set of content zero. Thus there exist $I_1, \ldots, I_n$ closed rectangles whose interior covers B_ε and $\sum_{i=1}^n V(I_i) < \varepsilon$. Let P be a partition of I such that every sub rectangle of P is in one of two groups.

1. P_1, which consists of sub rectangles S such that $S \subset I_i$ for some i.

2. P_2, which consists of sub rectangles S with $S \cap B_\varepsilon = \phi$.

Let $|f(x)| < M, \forall x \in I$. Then $M_S(f) - m_S(f) < 2M$ for any S. Thus

$$\sum_{S \in P_1} [M_S(f) - m_S(f)] V(S) < 2M \sum_{i=1}^{n} V(I_i) < 2M\varepsilon.$$

For $S \in P_2$ and $x \in S', O(f, x) < \varepsilon$. Thus by Lemma 7.1.1 there is a refinement P' of P (the refinement made only on S such that $S \in P_2$) such that

$$\sum_{\substack{S' \in P' \\ S' \subset S}} [M_{S'}(f) - m_{S'}(f)] V(S') < \varepsilon V(S).$$

Now

$$\begin{aligned} U_f(P') - L_f(P') &= \sum_{S\in P_1}[M_S(f) - m_{S'}(f)]V(S) \\ &\quad + \sum_{S'\subset S\in P_2}[M_{S'}(f) - m_{S'}(f)]V(S') \\ &< 2M_\varepsilon + \sum_{S\in P_2}\varepsilon V(S) \leq (2M + V(I))\varepsilon. \end{aligned}$$

Since M and $V(I)$ are fixed, this shows that f is integrable.

Now suppose that f is integrable. Since $B = B_1 \cup B_{1/2} \cup B_{1/3} \cup \ldots$, where $B_{1/n} = \{x : O(f,x) \geq 1/n\}$ it suffice to show that each $B_{1/n}$ is a set of measure zero. Let $\varepsilon > 0$, let P be a partition such that $U_f(P) - L_f(P) < \frac{\varepsilon}{n}$. Let P_1 be the collection of sub rectangles which intersects $B_{1/n}$. Then P_1 covers $B_{1/n}$. Moreover for $S \in P_1 M_S(f) - m_S(f) \geq 1/n$. Thus

$$\begin{aligned} \frac{1}{n}\sum_{S\in P_1} V(S) &\leq \sum_{S\in P_1}[M_S(f) - m_S(f)]V(S) \\ &\leq \sum_{S\in P}[M_S(f) - m_S(f)]V(S) < \frac{\varepsilon}{n}. \end{aligned}$$

$\Rightarrow \sum_{S\in P_1} V(S) < \varepsilon$. Hence $B_{1/n}$ is a set of measure zero.

Now we define integral over any bounded set C. Let C be a bounded set then $C \subset I$ where I is a closed rectangle. We define the characteristic function of C to be

$$\chi_C = \begin{cases} 1 & x \in C \\ 0 & x \notin C \end{cases}.$$

Then

$$\int_C f = \int_I f\chi_C.$$

Exercises 7.3

1. Let $f : [0,1] \to R$ be given by

$$f(x) = \begin{cases} 0 & x \quad \textit{irrational} \\ \frac{1}{f} & x = P/f. \end{cases}$$

(a) Show that f is discontinuous only at rational numbers and f is integrable on $[0,1]$.

(b) Find $\int_0^1 f$

2. Let $f : [0,1] \times [0,1] \to R$ be defined by

$$f(x,y) = \begin{cases} 0 & if \quad 0 \leq x \leq \frac{1}{2} \\ 1 & if \quad \frac{1}{2} \leq x < 1 \end{cases}.$$

Show that f is integrable and $\int_{[0,1]\times[0,1]} f = 1/2$.

3. Let f be integrable on I. Show that $|f|$ is also integrable and

$$|\int_I f| \leq \int_I |f|.$$

4. Let A be a rectangle in R^n; let B be a rectangle in R^n; let $Q = A \times B$. Let $f : Q \to R$ be bounded function. Show that if $\int_Q f$ exists, then

$$\int_{y \in B} f(x,y)$$

exists for $x \in A - D$, where D is a set of measure zero in R^k.

5. Let S_1 and S_2 be bounded sets in R^n; let $f : S \to R$ be a bounded function. Show that if f is integrable over S_1 and S_2, then f is integrable over $S_1 - S_2$, and

$$\int_{S_1 - S_2} f = \int_{S_1} f - \int_{S_1 \cap S_2} f.$$

6. Let $f, g : S \to R$; assume f and g are integrable over S.

(i) Show that if f and g agree except on a set of measure zero, then

$$\int_S f = \int_S g.$$

(ii) Show that if $f(x) \leq g(x)$ for $x \in S$ and $\int_S f = \int_S g$, then f and g agree except on a set of measure zero.

7. Let D be a compact, connected Jordan domain in R^n with positive volume, and suppose that the function $f : D \to R$ is continuous. Show that there is a point x in D at which

$$f(x) = \frac{1}{volD} \int_D f.$$

8. Let I be a rectangle in R^n and let the function $f : I \to R$ be integrable. Denote the interior of I by D. Show that the restriction $f : D \to R$ is integrable and that

$$\int_I f = \int_D f.$$

7.4 Iterated Integrals

Let $R : [a, b] \times [c, d]$ be a rectangle, the volume over R under the graph $z = f(x, y)$ is given by

$$\int_a^b (\texttt{cross-sectional area at } x)dx$$
$$= \int_a^b (\int_c^d f(x, y)dy)dx = \int_a^b \int_c^d f(x, y)dydx.$$

This expression is called an iterated integral.

Example 7.4.1. Consider the iterated integral

$$\begin{aligned} \int_0^1 \int_1^2 (1 + x^2 + xy)dydx &= \int_0^1 (1 + x^2)y + \frac{xy^2}{2}\Big]_{y=1}^2 dx \\ &= \int_0^1 (1 + x^2 + \frac{3}{2}x)dx \\ &= x + \frac{x^3}{3} + \frac{3}{4}x^2\Big]_0^1 = \frac{25}{12}. \end{aligned}$$

Example 7.4.2.

$$\int_0^2 \int_{-1}^1 xye^{x+y^2} dydx$$
$$= \int_0^2 \left(\int_{-1}^1 (xe^x)(ye^{y^2}) dy \right) dx$$
$$= \int_0^2 \left(\frac{1}{2}(xe^x)e^{y^2} \Big]_{y=-1}^1 \right) dx = 0[\ \texttt{Since}(xe^x)y(e^{y^2})$$

is odd function of y].

Example 7.4.3. Find the volume of the region lying over the triangle $\Omega \subset R^2$ with vertices at $(0,0)(0,1)$ and $(1,1)$ and bounded by $z = f(x,y) = xy$.

Solution. We consider Ω as a subset of the square $R = [0,1] \times [0,1]$ and define $\widetilde{f} : R \to R$ by

$$\widetilde{f}(x,y) = \begin{cases} xy & if (x,y) \in \Omega \\ 0 & otherwise \end{cases}.$$

Note that for x fixed, $\widetilde{f}(x,y) = xy$ when $0 \leq y \leq x$ and is 0 otherwise. So

$$\int_0^1 \widetilde{f}(x,y)dy = \int_0^x xydx + \int_0^1 0dy = \int_0^x xydy$$
$$\int_0^1 \int_0^1 \widetilde{f}(x,y)dydx = \int_0^1 (\int_0^x xydy)dx = \int_0^1 \left(\frac{1}{2}xy^2 \Big]_{y=0}^x \right) dx = \frac{1}{8}.$$

7.5 Fubini's Theorem, 2-Dimensional Case

Suppose f is integrable on a rectangle $R = [a,b] \times [c,d] \subset R^2$. Suppose that for each $x \in [a,b]$, the function $f(x,y)$ is integrable on $[c,d]$ i.e., $F(x) = \int_c^d f(x,y)dy$ exists. Suppose next that the function F is integrable on $[a,b]$, i.e.,

$$\int_a^b F(x)dx = \int_a^b \left(\int_c^d f(x,y)dy \right) dx$$

exists. Then we have

$$\int_R f dA = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

Proof. Let P be an arbitrary partition of R into rectangles $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j], i = 1, \ldots, k, j = 1, \ldots, l$.

When $(x, y) \in R_{ij}$, we have

$$\begin{aligned} m_{ij} \leq f(x, y) &\leq M_{ij} \\ m_{ij}(y_j - y_{j-1}) &\leq \int_{y_{j-1}}^{y_j} f(x, y) dy \leq M_{ij}(y_j - y_{j-1}). \end{aligned}$$

So now when $x \in [x_{i-1}, x_i]$, we have

$$\sum_{j=1}^{l} m_{ij}(y_j - y_{j-1}) \leq \int_c^d f(x, y) dy \leq \sum_{j=1}^{l} M_{ij}(y_j - y_{j-1})$$

or

$$\begin{aligned} \left(\sum_{j=1}^{l} m_{ij}(y_j - y_{j-1}) \right) (x_i - x_{i-1}) &\leq \int_{x_{i-1}}^{x_i} \left(\int_c^d f(x, y) dy \right) dx \\ &\leq (\sum_{j=1}^{l} M_{ij}(y_j - y_{j-1}))(x_i - x_{i-1}). \end{aligned}$$

Summing over i, we have

$$\begin{aligned} \sum_{i=1}^{k} \sum_{j=1}^{l} m_{ij}(y_j - y_{j-1})(x_i - x_{i-1}) &\leq \int_a^b \left(\int_c^d f(x, y) dy \right) dx \\ &\leq \sum_{i=1}^{k} \left(\sum_{j=1}^{l} M_{ij}(y_j - y_{j-1}) \right) \times \\ (x_i - x_{i-1}). \end{aligned}$$

But this can be rewritten as

$$\begin{aligned} \sum_{i=1}^{k} \sum_{j=1}^{l} m_{ij} \texttt{ area } (R_{ij}) &\leq \int_a^b \left(\int_c^d f(x, y) dy \right) dx \\ &\leq \sum_{i=1}^{k} \sum_{j=1}^{l} M_{ij} \texttt{ area } (R_{ij}). \end{aligned}$$

or

$$L_f(P) \le \int_a^b \left(\int_c^d f(x,y)dy \right) dx \le U_f(P).$$

Since f is integrable on R, if a number I satisfies $L_f(P) \le I \le U_f(P)$ for all partitions P of $[a,b]$ then

$$I = \int_R f dA.$$

Corollary 7.5.1. Suppose f is integrable on the rectangle $R = [a,b] \times [a,d]$ and the iterated integrals

$$\int_a^b \int_c^d f(x,y)dydx \text{ and } \int_c^d \int_c^b f(x,y)dxdy$$

both exist. Then

$$\int_a^b \int_c^d f(x,y)dydx = \int_R f dA = \int_c^d \int_a^b f(x,y)dxdy.$$

Example 7.5.1. Find a function f on the rectangle $R = [0,1] \times [0,1]$ that is integrable but whose iterated integral does't exist.

Solution. Take

$$f(x,y) = \begin{cases} 1, & x = 0, y \in Q \\ 0, & otherwise \end{cases}.$$

The integral $\int_0^1 f(0,y)dy$ does not exist, but f is integrable and $\int_R f dA = 0$.

Example 7.5.2. Find a function whose iterated integral exists but that is not integrable.

Solution. Let

$$f(x,y) = \begin{cases} 1 & , \quad y \in Q \\ 2x & , \quad y \notin Q \end{cases}.$$

Then $\int_0^1 f(x,y)dx = 1$ for every $y \in [0,1]$, so the iterated integral $\int_0^1 \int_0^1 f(x,y)dxdy$ exists and equals 1. Whether f is integrable on $R = [0,1] \times [0,1]$ we proceed as assume it is integrable then it would also be integrable on $R' = [0, \frac{1}{2}] \times [0,1]$. For any partition P of R', we have $U_f(P) = \frac{1}{2}$, whereas we can make $L_f(P)$ as close to $\int_0^1 \int_0^{1/2} 2x dx dy = \frac{1}{4}$ as we wish. Hence f is not integrable.

Example 7.5.3. Evaluate the iterated integral

$$\int_0^1 \int_y^1 \frac{\sin x}{x} dx dy.$$

Solution. It is a classical fact that $\int \frac{\sin x}{x} dx$ can not be evaluate in elementary terms, and so we define

$$f(x,y) = \begin{cases} \frac{\sin x}{x} & , \quad x \neq 0 \\ 1 & \quad x = 0 \end{cases} .$$

Then f is continuous and the iterated integral is equal to the double integral $\int_\Omega f dA$, where

$$\Omega = \{(x,y) : 0 \leq y \leq 1, y \leq x \leq 1\}.$$

Now changing the order of integration we get

$$\Omega = \{(x,y) : 0 \leq x \leq 1, 0 \leq y \leq x\}$$

and

$$\begin{aligned} \int_\Omega f dA &= \int_0^1 \left(\int_0^x \frac{\sin x}{x} dy \right) dx \\ &= \int_0^1 \left(\left(\frac{\sin x}{x} \right) y \Big]_{y=0}^x \right)^{dx} \\ &= \int_0^1 \left(\frac{\sin x}{x} . x \right) dx = 1 - \cos 1. \end{aligned}$$

Example 7.5.4. Let $\Omega =$

$\{x \in R^n : 0 \leq x_n \leq x_{n-1} \leq \cdots \leq x_2 \leq x_1 \leq 1\}$. Then

$$\begin{aligned} vol(\Omega) &= \int_0^1 \int_0^{x_1} \ldots \int_0^{x_{n-1}} dx_n \ldots dx_2 dx_1 \\ &= \int_0^1 \int_0^{x_1} \ldots \int_0^{x_{n-2}} x_{n-1} dx_{n-1} \ldots dx_2 dx_1 \\ &= \int_0^1 \int_0^{x_1} \ldots \int_0^{x_{n-3}} \frac{1}{2} x_{n-2}^2 dx_{n-2} \ldots dx_2 dx_1 \\ &= \cdots = \int_0^1 \frac{1}{(n-1)!} x_1^{n-1} dx_1 = \frac{1}{n!}. \end{aligned}$$

Fubini's Theorem in General Form

Let $I \subset R^n$ and $J \subset R^m$ be closed rectangles and let $f : I \times J \to R$ be integrable. For each $x \in I$ let $g_x : J \to R$ defined by $g_x(y) = f(x, y)$ and

$$\begin{aligned} \Im(x) &= L\int_J g_x = L\int_J f(x,y)dy \\ \aleph(x) &= U\int_J g_x = U\int_J f(x,y)dy. \end{aligned}$$

Then $\Im$ and $\aleph$ are integrable on I and

$$\begin{aligned} \int_{I\times J} f &= \int_I \Im \\ &= \int_I (L\int_J f(x,y)dy)dx. \end{aligned}$$

$$\int_{I\times J} f = \int_I \aleph = \int_I (U\int_J f(x,y)dy)dx.$$

Proof. Let P_I be a partition of I and P_J a partition of J. Together they give a partition P of $I \times J$ such that any sub rectangle S of P is $S_I \times S_J$ and $S_I \in P_I$ and $S_J \in P_J$.

$$\begin{aligned} L_f(P) = \sum_S m_S(f)V(S) &= \sum_{S_I}\sum_{S_J} m_{S_I\times S_J}(f)V(S_I \times S_J) \\ &= \sum_{S_I}\sum_{S_J} m_{S_I\times S_J}(f).V(S_J)V(S_I). \end{aligned}$$

Now if $x \in S_A$ then clearly

$$\begin{aligned} m_{S_I \times S_J}(f) &\leq m_{S_J}(g_x) \text{ and} \\ M_{S_I \times S_J}(f) &\geq M_{S_J}(g_x). \end{aligned}$$

Consequently for $x \in S_I$ we have

$$\sum_{S_J} m_{S_I \times S_J}(f) V(S_J) \leq \sum_{S_J} (g_x) V(S_J) \leq L \int_B g_x = \Im(x).$$

$\Rightarrow \sum_{S_I} \sum_{S_J} m_{S_I \times S_J}(f) V(S_J) V(S_I) \leq L_{\Im}(P_I).$

Similarly it can be shown that $U_{\aleph}(P_I) \leq U_f(P)$.
Thus we obtain

$$L_f(P) \leq L_{\Im}(P_I) \leq U_{\Im}(P_I) \leq U_{\aleph}(P_I) \leq U_f(P)$$

$\Rightarrow \Im$ is integrable and $\int_I \Im = \int_{I \times J} f$.

Similarly from the fact that

$$L_f(P) \leq L_{\Im}(P_I) \leq L_{\aleph}(P_I) \leq U_{\aleph}(P_I) = \leq U_f(P)$$

this implies that $\aleph$ is integrable and

$$\int_I \aleph = \int_{I \times J} f.$$

Remark 7.5.1. $f : I \times J \to R$ is continuous. Then $\aleph(x) = \Im(x)$ and hence

$$\int_{I \times J} f = \int_I (\int_J f(x,y) dy) dx.$$

Thus $f : J = [a^1, b^1] \times [a^2, b^2] \times \cdots \times [a^n, b^n] \to R$ is continuous, then

$$\int_I f = \int_{a^n}^{b^n} \left[\int_{a^{n-1}}^{b^{n-1}} \cdots \left(\int_{a^1}^{b^1} f(x^1 x^2 \ldots x^n) dx \right) dx^2 \right) \ldots dx^n.$$

Example 7.5.5. Construct a function $f : [0,1] \times [0,1] \to R$ such that $g_x : [0,1] \to R$ given by $g_x(y) = f(x,y)$ is integrable but f is not.

Hint. Construct a set $A \subset [0,1] \times [0,1]$ such that each vertical or horizontal line contains at most one point of $[0,1] \times [0,1]$ but $\partial A = [0,1] \times [0,1]$

Change of Variables

Theorem 7.5.1. Let $A \subset R^n$ be an open set $g : A \to R^n$ a one one continuously differentiable function such that $detg'(x) \neq 0, \forall x \in A$. If $f : g(A) \to R$ is integrable function. Then

$$\int_{g(A)} f = \int_A (fog)|detg'|.$$

Exercises 7.5

1. Let $g : \{r|r > 0\} \times (0, 2\pi) \to R^2$ be given by $g(r, \theta) = (r\cos\theta, r\sin\theta)$

   (i) Show that g is one one and continuously differentiable with $detg' \neq 0$.

   (ii) Let $C = \{(x, y) : r_1^2 < x^2 + y^2 < r_2^2\}$ where $r_1 < r_2, r_1, r_2 > 0$.
   If $f : C \to R$ is integrable show that

   $$\int_C f = \int_{r_1}^{r^2} \left(\int_0^{2\pi} f(r\cos\theta, r\sin\theta)d\theta \right) dr.$$

2. Show that $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

   Hint. Note that

   $$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-(x^2+y^2)} dydx = \sqrt{\int_{-\infty}^{\infty} e^{-x^2} dx}. \right.$$

   (a) Define $C_r = \{(x, y)/x^2 + y^2 \leq r^2\}$ and show that

   $$\int_{C_r} e^{-(x^2+y^2)} dydx = \int_0^r \int_0^{2\pi} e^{-r^2} r dr d\theta.$$

   (b) Verify that $\lim_{r\to\infty} \int_{C_r} e^{-(x^2+y^2)} dydx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f$.

3. Evaluate

   $$\int\int_{[0,1]\times[0,1]} \sin^2 x \sin^2 y dxdy.$$

4. Show that

$$\int_0^3 \left[\int_1^{\sqrt{4-y}} (x+y)dx \right] dy = \int_1^2 \left[\int_0^{4-x^2} (x+y)dy \right] dx = \frac{241}{60}.$$

5. Let A be open in R^2; let $f : A \to R$ be of class C^2. Let Q be a rectangle contained in A.

   Use Fubini's theorem and the fundamental theorem of calculus to show that

$$\int_Q D_2 D_1 f = \int_D D_1 D_2 f.$$

6. Give and example where $\int_Q f$ exists and one of the iterated integrals

$$\int_{x\in A} \int_{y\in B} f(x,y) \texttt{ and } \int_{y\in B} \int_{x\in A} f(x,y)$$

   exists, but the other does not, where $Q = A \times B$; A is rectangle in R^k and B is a rectangle in R^n and $f : Q \to R$ be a bounded function.

7. Let $I = [0,1]$; let $Q = I \times I$. Define $f : Q \to R$ by letting $f(x,y) = 1/q$ if y is rational and $x = p/q$, where p and q are positive integers with no common factor; let $f(x,y) = 0$ otherwise.

   (a) Show that $\int_Q f$ exists.

   (b) Compute

$$\int_{-y\in I} f(x,y) \texttt{ and } \overline{\int}_{y\in I} f(x,y).$$

   (c) Verify Fubini's theorem.

8. For a continuous function $f : [a.b] \times [a,b] \to R$, prove Dirichlet's formula

$$\int_a^b \left[\int_a^x f(x,y)dy \right] dx = \int_a^b \left[\int_y^b f(x,y)dx \right] dy.$$

9. Suppose that the function $\phi : R^{\rightarrow} R$ is continuous. Prove that for each $x \geq 0$,

$$\int_0^x \left[\int_0^t \phi(s)ds\right] dt = \int_0^x (x-s)\phi(s)ds.$$

10. Let $g : \{r|r > 0\} \times (0, 2\pi) \rightarrow R^2$ be given by $g(r, \phi) = (r\cos\theta, r\sin\theta)$. Show that g is one-one and continuously differentiable with determinant $g' \neq 0$.

11. Let $C = \{(x, y) : r_1^2 < x^2 + y^2 < r_2^2\}$ where $r_1 < r_2, r_1, r_2 > 0$. If $f : C \rightarrow$ is integrable show that

$$\int_C f = \int_{r_1}^{r_2} \left(\int_0^{2\pi} f(r\cos\theta, r\sin\theta)d\theta\right) dr.$$

12. Show that

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{x}.$$

Hint. Note that $\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} e^{-(x^2+y^2)}dydx\right) = \left[\int_{-\infty}^{\infty} e^{-x^2}dx\right]^2$.

13. Define $C_r = \{(x, y) : x^2 + y^2 \leq r^2\}$ and show that

$$\int_{C_r} e^{-(x^2+y^2)}dydx = \int_0^r \int_0^{2\pi} e^{-r^2} rdrd\theta.$$

Verify that

$$\lim_{r\rightarrow\infty} \int_{C_r} e^{-(x^2+y^2)}dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f.$$

14. Let S be the tetrahedron in R^3 having vertices $(0, 0, 0)$, $(1, 2, 3)$, $(0, 1, 2)$ and $(-1, 1, 1)$. Evaluate $\int_S f$, where $f(x, y, z) = x + 2y - z$.

Hint. Use a suitable linear transformation g as a change of variables.

15. Let B be the portion of the first quadrant in R^2 lying between the hyperbolas $xy = 1$ and $xy = 2$ and the two straight lines $y = x$ and $y = 4x$. Evaluate $\int_B x^2y^3$.

    Hint. Set $x = u/\nu$ and $y = u\nu$.

16. If $V = \{(x, y, z) : x^2 + y^2 + z^2 < a^2 \text{ and } z > 0\}$,

    use the spherical coordinate transformation to express $\int_V z$ as an integral over an appropriate set in (ρ, ϕ, θ) space. Justify your answer.

Bibliography

Here some references are given which are useful for further reading as well as to look into some deep results which have been left unproved in the text.

- Apostal, T.M, *Mathematical Analysis*, 2^{nd} edition, Addison Wesley, 1974.
- Boothby, W.M., *An Introduction to Differential Manifolds and Riemannian Geometry*, Academic Press, 1975.
- Devinatz, A., *Advanced Calculus*, Holt, Rinehart and Winston, 1968.
- Fleming, W., *Functions of Several Variables*, Addison - Wesley, 1965, Springer - Verlag, 1977.
- Fitzpatrick, Patrick, M., *Advanced Calculus*, PWS Pub. Co. 1996.
- Goldberg, R.P., *Methods of Real Analysis*, Wiley, 1976.
- Munkres, J.R., *Analysis on Manifolds*, Addison - Wesley, Pub. Co., 1991.
- Munkres, J.R., *Topology, A First Course*, Prentice - Hall, 1975.
- Nickerson, H.K., Spenser, D.C. and Steenrod, N.E., *Advanced Calculus*, Van Nostrand, 1959.
- Northcott, D.G., *Multi-linear Algebra*, Cambridge Univ. Press, 1984.
- Royden, H., *Real Analysis*, 3^{rd} edition, Macmillan, 1988.
- Rudin, W., *Principles of Mathematical Analysis*, 3^{rd} edition, McGraw-Hill, 1976.
- Spivak, M., *Calculus on Manifolds*, Addison-Wesley, 1965.

Index