MA/MSc MT-06

Vardhaman Mahaveer Open University, Kota

Analysis and Advanced Calculus

Vardhaman Mahaveer Open University, Kota

Analysis and Advanced Calculus

Course Development Committee

Chairman
Prof. (Dr.) Naresh Dadhich
Vice-Chancellor
Vardhaman Mahaveer Open University, Kota

Co-ordinator/Convener and Members

Subject Convener
Prof. D.S. Chauhan
Department of Mathematics
University of Rajasthan, Jaipur

Co-ordinator
Dr. Anuradha Sharma
Assistant Professor
Department of Botany, V.M.O.U., Kota

Members :

1. **Prof. V.P. Saxena**
   Ex Vice-Chancellor
   Jiwaji University,
   Gwalior (MP)
2. **Prof. S.C. Rajvanshi**
   Department of Mathematics
   Institute of Eng. & Tech.
   Bhaddal, Ropar (Punjab)
3. **Prof. P.K. Banerjee**
   Emeritus Fellow (UGC)
   Department of Mathematics
   J.N.V. University, Jodhpur
4. **Prof. S.P. Goyal**
   Emeritus Scientist (CSIR)
   Department of Mathematics
   University of Rajasthan, Jaipur
5. **Dr. A.K. Mathur**
   Associate Prof. (Retired)
   Department of Mathematics
   University of Rajasthan, Jaipur
6. **Dr. K.N. Singh**
   Associate Prof. (Retired)
   Department of Mathematics
   University of Rajasthan, Jaipur
7. **Dr. Paresh Vyas**
   Assistant Professor
   Department of Mathematics
   University of Rajasthan, Jaipur
8. **Dr. Vimlesh Soni**
   Lecturer
   Department of Mathematics
   Govt. PG College, Kota (Raj.)
9. **Dr. K.K. Mishra**
   Lecturer
   Department of Mathematics
   M.S.J. College, Bharatpur (Raj.)
10. **Dr. K.S. Shekhawat**
   Lecturer, Department of Mathematics
   Govt. Shri Kalyan College, Sikar (Raj.)

Editing and Course Writing

Editor
Prof. D.S. Chauhan
Department of Mathematics
University of Rajasthan, Jaipur

Writers

1. **Dr. A.K. Goyal**
   Department of Mathematics
   M.S.J. College, Bharatpur
2. **Prof. S.P. Goyal**
   Emeritus Scientist (CSIR)
   Department of Mathematics
   University of Rajasthan, Jaipur
3. **Dr. Keshav Sharma**
   Department of Mathematics
   R.R. Government College, Alwar

Academic and Administrative Management

Prof. (Dr.) Naresh Dadhich
Vice-Chancellor
Vardhaman Mahveer Open University, Kota

Prof. B.K. Sharma
Director (Academic)
Vardhaman Mahveer Open University, Kota

Mr. Yogendra Goyal
Incharge
Material Production and Distribution Department

Course Material Production

Mr. Yogendra Goyal
Assistant Production Officer
Vardhaman Mahaveer Open University, Kota

PREFACE

The present book entitled **"Analysis and Advanced Calculus"** has been designed so as to cover the unit-wise syllabus of MA/MSc MT-06 course for M.A./ M.Sc. Mathematics (Final) students of Vardhaman Mahaveer Open University, Kota. It can also be used for competitive examinations. The basic principles and theory have been explained in a simple, concise and lucid manner. Adequate number of illustrative examples and exercises have also been included to enable the students to grasp the subject easily. The units have been written by various experts in the field. The unit writers have consulted various standard books on the subject and they are thankful to the authors of these reference books.

Unit - 1
Normed Linear Spaces

Structure of the Unit

1.0 Objectives

In this unit, we introduce the concept of a norm over a linear space. A Banach space is a normed linear space which is complete metric space. The theory of normed linear spaces and Banach spaces, and the theory of linear operators defined on them are the fundamental of functional analysis. In this unit, we discuss basic propeties of normed linear spaces and Banach spaces and give some examples of these spaces.

1.1 Introduction

Usefull and important spaces are obtained if we take a linear space and define on it a metric by means of a norm. The resulting space is called a normed linear space. Normed spaces and metric spaces are special enough to provide a basis for a rich theory in functional analysis.

1.2 Linear (Vector) Spaces

A linear space (or vector spaces) is an additive abelian group L (whose elements are called vectors) with the property that any scalar α and any vector x can be combined by an operation called scalar multiplication to yield a vector αx in such a manner that

(i) $\alpha(x+y) = \alpha x + \alpha y$;

(ii) $(\alpha+\beta)x = \alpha x + \beta x$;

(iii) $(\alpha\beta)x = \alpha(\beta x)$;

(iv) $1.x = x$

A linear space is thus an additive abelian group whose elements can be multiplied by numbers in a reasonable way. The two primary operations in a linear space-vector addition and scalar multiplication are called the linear operations.

A linear space is called a real linear space or a complex linear space according as the scalars are the real numbers or the complex number.

1.3 Basic Concepts of Norm and Normed Spaces

1.3.1 Norm and Normed Space :

If N be a real or complex linear (vector) space and $\|.\|$ be a function from N into R (set of reals) i.e. $\|.\| : N \to R$ or $x \to \|x\|$ with $x \in N$,

then the non-negative real number $\|x\|$ regarded as the length of the vector x and said to be the **Norm** on N and the pair $(N, \|.\|)$ is called as **Normed linear space**, provided for all $x, y \in N$ and all $\alpha \in R$ (or C), the following axioms are satisfied :

$$N_1 : \|x\| \geq 0, \quad \text{if } x \neq 0$$

$$N_2 : \|x\| = 0 \Leftrightarrow x = 0$$

$$N_3 : \|x + y\| \leq \|x\| + \|y\|$$

$$N_4 : \|\alpha x\| = |\alpha| \, \|x\|$$

The function $\|.\|$ becomes a **semi-norm** and the corresponding space becomes **semi-normed linear space** if N_2 is replaced by

$$N_2(p) : x = 0 \Rightarrow \|x\| = 0$$

Example : The metric space induced by the metric $d(x, y) = \|x - y\|$ is a normed linear space.

1.3.2 Convergence in Normed Linear Space

Definition : A sequence $< x_n >$ in N i.e., normed linear space $(N, \|.\|)$ is said to **converge** to an element $x_0 \in N$ if given arbitrary $\in > 0$, $\exists$ a positive number (integer) n_0 s.t

$$n \geq n_0 \Rightarrow \|x_n - x_0\| < \in$$

and we write $\lim\limits_{n\to\infty} x_n = x_0$ or $x_n \to x_0$ as $n \to \infty$ i.e.,

Thus $x_n \to x_0$ iff $\|x_n - x_0\| \to 0$.

Definition : A sequence $< x_n >$ in N is said to be a **Cauchy sequence** if given $\in > 0 \exists$ a positive integer n_0 such that

$$m, n \geq n_0 \Rightarrow \|x_m - x_n\| < \in.$$

Definition : A sequence $< x_n >$ in N is said to be **bounded** if $\exists$ a real constant $K > 0$ s.t. $\|x_n\| \leq K$ for all n.

Definition : If every Cauchy sequence $< x_n >$ in N is convergent i.e. if $\forall$ Cauchy sequence $< x_n >$ in N $\exists$ an element $x_0 \in N$ s.t. $x_n \to x_0$, then the normed linear space is said to be **complete**.

1.3.3 Summability in Normed Linear Spaces

A series $\sum f_n$ of functions in a normed linear space N is **summable** to a sum s in N, if the sequence of partial sums of the series converges, s.t.

$$\left\| s - \sum_{i=1}^{n} f_i \right\| \to 0 \quad \text{as} \quad n \to \infty$$

i.e. $$s = \sum_{i=1}^{\infty} f_i$$

The series $\sum f_n$ is asbolutely summable if $\sum_{i=1}^{\infty} \|f_n\| < \infty$

1.3.4 Continuity in Normed Linear Space

If N, M be two normed linear spaces, then a function $f : N \to M$ is **continuous** at $x_0 \in N$ iff $\forall \in > 0$, $\exists$ a $\delta > 0$ s.t.

$$\|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \in.$$

The function f is continuous on N iff f is continuous at each point of N.

In other words, $f : N \to M$ is continuous at $x_0 \in N$ iff $\forall$ sequence $< x_n >$ in N converging to $x_0 \in N$, the sequence $< f(x_n) >$ in M converges to $f(x_0) \in M$ i.e., iff $x_n \to x_0 \Rightarrow f(x_n) \to f(x_0)$.

In case of three topological spaces X, Y, Z the continuous function $f : X \times Y \xrightarrow{\text{into}} Z$ is **jointly continuous** in x and y if $f(x, y) = z$. In other words, if

$f(x_n, y_n) \to f(x, y)$ whenever $x_n \to x$, $y_n \to y$ as $n \to \infty$.

1.3.5 Allied Spaces to Normed Linear Spaces

Banach Space : A complete normed linear space is known as a Banach space.

Function Space : A function space is the metric space which is linear space with elements as functions defined as $X(\neq \phi)$ with addition and multiplication, i.e., $f : X \to R : (f + g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha f(x)$.

***n*-Dimensional Enclidean Space :** If R^n be a set of all ordered n-tuples $x = (x_1, x_2, ...x_n)$ of real numbers, s.t. R^n is a real linear space with additive and multiplicative operations such as

$x+y=(x_1+y_1,\, x_2+y_2,\ldots,x_n+y_n)$ where $y=(y_1,y_2,\ldots\ldots,y_n)$ and $\alpha\, x=(\alpha\, x_1,\alpha\, x_2,\ldots,\alpha\, x_n)$ so that $\mathbf{0}=(0,\,0,\ldots\ldots,0)$ and $-x=(-x_1,-x_2,\ldots,-x_n)$ etc.

then R^n is a ***n*-dimensional space.** We can regard R^n as composed of real functions f defined on $(1,2,\ldots\ldots,n)$ s.t. $\|f\|=\left[\sum_{i=1}^{n}|f(i)|^2\right]^{1/2}$ known as **Euclidean norm,** then normed linear space R^n is called n-dimensional Euclidern space.

***n*-Dimensional Unitary Space :** The set C^n of all n-tuples $z=(z_1,z_2,\ldots,z_n)$ of complex numbers constitutes a complex Banach space w.r.t. operations of addition and scalar multiplication and the norm given by

$$\|z\|=\left[\sum_{i=1}^{n}|z_i|^2\right]^{1/2}$$

It is known as an **n-dimensional unitary space.**

1.4 Theorems on Normed Spaces

Theorem 1 : If N be a normed linear space and $x,\, y\in N$, then

$$\big|\,\|x\|-\|y\|\,\big|\le\|x-y\|$$

Proof : We can write

$$\|x\|=\|(x-y)+y\|\le\|x-y\|+\|y\| \qquad \text{by } N_3$$

giving $\|x\|-\|y\|\le\|x-y\|$ ...(1)

and $\|y\|=\|(y-x)+x\|\le\|y-x\|+\|x\|$, giving

$$\|y\|-\|x\|\le\|y-x\|=\|-(x-y)\|$$

$$=|-1|\|x-y\| \qquad \text{by } N_4$$

$$=\|x-y\|$$

or $\|y\|-\|x\|\le\|x-y\|$ ...(2)

$\therefore$ (1) and (2) $\Rightarrow\big|\,\|x\|-\|y\|\,\big|\le\|x-y\|$.

Theorem 2 : Every normed linear space is a metric space.

Proof : Let N be a normed linear space and let

$$d:N\times N\to R \text{ defined by } d(x,y)=\|x-y\|.$$

$[M_1]$ $x, y \in N \Rightarrow x - y \in N \Rightarrow \|(x-y)\| \geq 0$ (by N_1)

$$\Rightarrow d(x,y) \geq 0$$

$[M_2]$ $d(x,y) = 0 \Leftrightarrow \|x-y\| = 0$

$$\Leftrightarrow x - y = 0 \quad \text{(by } N_2\text{)}$$

$$\Leftrightarrow x = y$$

$[M_3]$ $d(x,y) = \|x-y\| = \|(-1)(y-x)\|$

$$= |-1|\|y-x\| \quad \text{(by } N_4\text{)}$$

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

$[M_4]$ $d(x,y) = \|x-y\| = \|x-z+z-y\|$

$$\leq \|x-z\| + \|z-y\| \quad \text{(by } N_3\text{)}$$

$$= d(x,z) + d(z,y)$$

It follows that d is a metric and hence N is a metric space.

Theorem 3 : If N be a normed linear space with the norm $\|.\|$, then the mapping $f : N \to R$ s.t. $f(x) = \|x\|$ is continuous. In other words, the norm $\|.\|$ on N is a continuous function.

Proof : Taking a sequence $< x_n >$ in N s.t. $x_n \to x \in N$, as $n \to \infty$, we have by Theorem 1,

$$|f(x_n) - f(x)| = |\,\|x_n\| - \|x\|\,|$$

$$\leq \|x_n - x\| \to 0 \quad \text{as} \quad n \to \infty$$

$\therefore$ $f(x_n) \to f(x)$ as $n \to \infty \Rightarrow f$ is continuous.

Theorem 4 : Every convergent sequence in a normed linear space is a Cauchy sequence.

Proof : Assuming that a sequence $< x_n >$ in a normed linear space N converges to $x_0 \in N$. We claim that $< x_n >$ is a Cauchy sequence.

Given $\in > 0$, and the sequence $< x_n > \to x_0$, $\exists$ a positive integer n_0 s.t.

$$n \geq n_0 \Rightarrow \|x_n - x_0\| < \frac{\in}{2}$$

so that for all $m, n > n_0$, we have

$$\|x_m - x_n\| = \|x_m - x_0 + x_0 - x_n\|$$

$$\leq \|x_m - x_0\| + \|x_0 - x_n\| \qquad \text{by } N_3.$$

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

i.e., $\|x_m - x_n\| < \in \Rightarrow$ the sequence $< x_n >$ is a Cauchy sequence.

Note : Its converse is not true, i.e., every Cauchy sequence (particularly in a metric space) is not convergent.

Consider a metric $d(x, y) = |x - y|$ in a space $X = (0, 1)$. Then the sequence $< x_n > = < \frac{1}{n} > \in X$, is clearly a Cauchy sequence, since $d(x, y) = |x_n - x_m| = \left|\frac{1}{n} - \frac{1}{m}\right| \to 0$ as $m, n \to \infty$, but $d\left(\frac{1}{n}, 0\right) = \left|\frac{1}{n}\right| \to 0$ as $n \to \infty$ with $0 \notin X$ shows that $< x_n >$ in X is not necessarily a convergent sequence.

Theorem 5 : The limit of a convergent sequence is unique.

Proof : Consider a convergent sequence $< x_n >$ in a normed linear space N, converging to two limits x, y s.t. $x \neq y$ i.e., $< x_n > \to x$ as well as $< x_n > \to y$. Then $\|x_n - x\| \to 0$ and $\|x_n - y\| \to 0$ as $n \to \infty$. ...(1)

Now $\|x - y\| = \|x - x_n + x_n - y\|$

$$\leq \|x - x_n\| + \|x_n - y\| \qquad \text{by } N_3$$

$$\leq |-1| \|x_n - x\| + \|x_n - y\| \qquad \text{by } N_4$$

$$\leq 0 \text{ by (1) as } n \to \infty$$

$\therefore \quad \|x - y\| = 0 \Rightarrow x - y = 0 \qquad$ by N_2

$\Rightarrow x = y$ i.e., the limit of $< x_n >$ in N is unique.

Lemma 6 : If $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, then $a, b \geq 0 \Rightarrow a^{1/p} b^{1/q} \leq \frac{a}{p} + \frac{b}{q}$, where the sign of equality holds iff $a^p = b^q$.

Proof : For $a = 0$ or $b = 0$, the result is obvious. Therefore taking $a > 0$, $b > 0$ and $k \in (0, 1)$, i.e., $0 < k < 1$, set a function

$$f(t) = 1 - k + kt - t^k \text{ for } t \geq 0 \qquad ...(1)$$

and $k = \frac{1}{p},\ t = \frac{a}{b}$

$$f(t) = k(t - 1) - t^k + 1 \Rightarrow f(1) = 0 \qquad ...(2)$$

$$f'(t) = k - kt^{k-1} = k\left(1 - t^{k-1}\right)$$

so that $f'(t) < 0$, for $0 < t < 1$

$$f'(t) > 0, \quad t > 1$$

For $0 < t < 1$ and some c s.t. $t < c < 1$, the mean value theorem of differential calculus yeilds

$$\frac{f(1) - f(t)}{1 - t} = f'(c) \Rightarrow f(1) - f(t) = f'(c)(1 - t) < 0 \text{ for } 1 - t > 0 \text{ and } f'(c) < 0$$

$$\Rightarrow f(t) > f(1) \qquad ...(3)$$

for $t > 1$ and some d s.t., $1 < d < t$, the mean value theorem gives

$$\frac{f(t) - f(1)}{t - 1} = f'(d) \Rightarrow f(t) - f(1) = (t - 1)f'(d) > 0 \text{ for } t > 1 \text{ and } f'(d) > 0$$

$$\Rightarrow f(t) > f(1) \qquad ...(4)$$

Thus (3) and (4) $\Rightarrow f(t) > f(1)$ either $t < 1$ or $t > 1$ but $t \neq 1$ ...(5)

and $f(t) = k(t - 1) - t^k + 1 \Rightarrow f(t) = 0$ for $t = 1$ ...(6)

Also (2) and (5) $\Rightarrow f(t) > 0$ for $t \neq 1$ ...(7)

$\therefore$ (6) and (7) $\Rightarrow f(t) \geq 0$ for $t \geq 0$

$$\Rightarrow (1 - k) + kt - t^k \geq 0$$

$$\Rightarrow t^k \leq kt + (1 - k) \text{ for } t \geq 0 \qquad ...(8)$$

$$\Rightarrow \left(\frac{a}{b}\right)^{1/p} \leq \frac{1}{p}\frac{a}{b} + 1 - \frac{1}{p}$$

$$\Rightarrow b\left(\frac{a}{b}\right)^{1/p} \leq \frac{a}{p} + b\left(1 - \frac{1}{p}\right)$$

$$\Rightarrow a^{1/p} b^{1/q} \leq \frac{a}{p} + \frac{b}{q} \text{ as } 1 - \frac{1}{p} = \frac{1}{q} \qquad ...(9)$$

Corollary : If we set $t = a^p b^{-q}$ in (8), other assumptions being the same, then we get

$$\left(a^p b^{-q}\right)^{\frac{1}{p}} \leq \frac{1}{p} a^p b^{-q} + 1 - \frac{1}{p}$$

or $\quad a\,b^{-\frac{q}{p}} \le \frac{1}{p}a^p b^{-q} + \frac{1}{q}$ as $\frac{1}{p}+\frac{1}{q}=1$

Multiplying both sides by b^q, this reduces to

$$a\,b^{q\left(1-\frac{1}{p}\right)} \le \frac{a^p}{p}+\frac{b^q}{q}$$

or $$a\,b \le \frac{a^p}{p}+\frac{b^q}{q} \qquad ...(10)$$

Now to show that the sign of equality holds iff $a^p = b^q$, we have

$$a^p = b^q \Rightarrow a^{\frac{1}{p}} = b^{\frac{1}{q}} \Rightarrow a^{\left(1-\frac{1}{p}\right)} = b^{\frac{1}{p}} \quad \text{as } \frac{1}{q} = 1-\frac{1}{p}$$

or $$a.a^{-\frac{1}{p}} = b^{\frac{1}{p}} \Rightarrow a = (ab)^{\frac{1}{p}} \Rightarrow a^p = ab$$

Similarly $b^q = ab$

$$\therefore \quad \frac{a^p}{p}+\frac{b^q}{q} = \frac{ab}{p}+\frac{ab}{q} = ab\left(\frac{1}{p}+\frac{1}{q}\right) = ab \quad \text{as } \frac{1}{p}+\frac{1}{q}=1$$

i.e. $$ab = \frac{a^p}{p}+\frac{b^q}{q},$$

which follows that the sign of equality holds if $a^p = b^q$

Theorem 7 [Holder's Inequality] :

If $x=(x_1, x_2, \ldots, x_n)$ and $y=(y_1, y_2, \ldots, y_n)$ be two $n-$ tuples of scalars (real or complex), then under the norm

$$\|x\|_p = \left[\sum_{i=1}^{n}|x_i|^p\right]^{1/p} \left[\sum_{i=1}^{n}|y_i|^p\right]^{1/q}$$, we have the inequality

$$\sum_{i=1}^{n}|x_i y_i| \le \left[\sum_{i=1}^{n}|x_i|^p\right]^{1/p} \left[\sum_{i=1}^{n}|y_i|^p\right]^{1/q}$$

$$= \|x\|_p \, \|y\|_q$$

where $1 < p < \infty$ and $\frac{1}{p}+\frac{1}{q}=1$.

Proof : For $x=0$, $y=0$, the result is obvious. We therefore consider the case when $x \ne 0$, $y \ne 0$.

The inequality (10) of Lemma 6 for $a_i \geq 0$, $b_i \geq 0$ yields

$$a_i\, b_i \leq \frac{a_i^p}{p} + \frac{b_i^q}{q} \qquad ...(1)$$

Setting $a_i = \frac{|x_i|}{\|x\|_p}$ and $b_i = \frac{|y_i|}{\|y\|_q}$, (1) reduces to

$$\frac{|x_i|}{\|x\|_p} \frac{|y_i|}{\|y\|_q} \leq \frac{1}{p} \frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{q} \frac{|y_i|^q}{\|y\|_q^q}$$

summing over i from 1 to n, we find

$$\frac{1}{\|x\|_p \|y\|_q} \sum_{i=1}^{n} |x_i||y_i| \leq \frac{1}{p\,\|x\|_p^p} \sum_{i=1}^{n} |x_i|^p + \frac{1}{q\,\|y\|_q^q} \sum_{i=1}^{n} |y_i|^q$$

$$= \frac{1}{p\,\|x\|_p^p} \|x\|_p^p + \frac{1}{q\,\|y\|_q^q} \|y\|_q^q$$

$$\left\{ as \;\; \|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p \right]^{1/p} \Rightarrow \|x\|_p^p = \sum_{i=1}^{n} |x_i|^p \right\}$$

$$= \frac{1}{p} + \frac{1}{q} = 1$$

or $\displaystyle\sum_{i=1}^{n} |x_i\, y_i| \leq \|x\|_p \|y\|_q$, since $|x_i||y_i| = |x_i\, y_i|$ ...(2)

Note : The theorem is also true for sequance $x = < x_n >$, $y = < y_n >$ s.t.

$$\sum_{n=1}^{\infty} |x_n|^p < \infty, \quad \sum_{n=1}^{\infty} |y_n|^p < \infty \;\; \text{for } p \geq 1$$

Corollary : For $p = 2$, $q = 2$ the inequality (2) reduces to

$$\sum_{i=1}^{n} |x_i\, y_i| \leq \left\{ \sum_{i=1}^{n} |x_i|^2 \right\}^{1/2} \left\{ \sum_{i=1}^{n} |y_i|^2 \right\}^{1/2}$$

$$= \|x\|_2 \|y\|_2 \qquad ...(3)$$

Theorem 8 [Minkowski's Inequality] :

If $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ be two $n-$tuples of real or complex numbers, then under the norm

$$\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p}, \quad p \geq 1$$

We have the inequality

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p$$

i.e., $$\left[\sum_{i=1}^{n} |x_i + y_i|^p\right]^{1/p} \leq \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p} + \left[\sum_{i=1}^{n} |y_i|^p\right]^{1/p}$$

where $1 < p < \infty$.

Proof : For $p = 1$, $\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right] \Rightarrow \|x\|_1 = \sum_{i=1}^{n} |x_i|$

so that $\|x+y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i|$ by N_3

or $\|x+y\|_1 \leq \|x\|_1 + \|y\|_1$,

which shows that the inequality holds for $p = 1$.

Taking $p > 1$ and setting $\frac{1}{q} = 1 - \frac{1}{p}$ so that $q > 1$, we have

$$\|x+y\|_p^p = \sum_{i=1}^{n} |x_i + y_i|^p$$

$$= \sum_{i=1}^{n} |x_i + y_i| \, |x_i + y_i|^{p-1}$$

$$= \sum_{i=1}^{n} \{|x_i + y_i|\} \, |x_i + y_i|^{p-1}$$

$$\leq \sum_{i=1}^{n} |x_i| \, |x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i| \, |x_i + y_i|^{p-1}$$ by N_3

$$\leq \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p} \left[\sum_{i=1}^{n} |x_i + y_1|^{q(p-1)}\right]^{1/q}$$

$$+ \left[\sum_{i=1}^{n} |y_i|^p\right]^{1/p} \left[\sum_{i=1}^{n} |x_i + y_i|^{q(p-1)}\right]^{1/q}$$ by Holder's inequality

$$= \|x\|_p \left[\sum_{i=1}^{n} |x_i + y_i|^p\right]^{1/q} + \|y\|_p \left[\sum_{i=1}^{n} |x_i + y_1|^p\right]^{1/q}$$

$$\text{since } \frac{1}{q} = 1 - \frac{1}{p} \Rightarrow q(p-1) = p$$

$$= \left\{\|x\|_p + \|y\|_p\right\} \left[\sum_{i=1}^{n} |x_i + y_i|^p\right]^{\frac{1}{q}}$$

$$= \left\{\|x\|_p + \|y\|_p\right\} \|x+y\|_p^{\frac{p}{q}}$$

or $\qquad \|x+y\|_p^{p-\frac{p}{q}} \le \|x\|_p + \|y\|_p$, where $p - \dfrac{p}{q} = p\left(1 - \dfrac{1}{q}\right) = p\,\dfrac{1}{p} = 1$

or $\qquad \|x+y\|_p \le \|x\|_p + \|y\|_p$

Note : The Theorem is also true for sequences $x = < x_n >$, $y = < y_n >$

s.t. $\qquad \sum_{i=1}^{n} |x_n|^p < \infty$, $\sum_{i=1}^{n} |y_n|^p < \infty$ for $p \ge 1$.

1.5 Factor (quotient) Spaces

If M be a subspace of a vector space N, then $\exists$ an equivalence relation between any two vectors $x, y \in N$ i.e., $x \sim y$ iff $x - y \in M$, since this relation is :

Reflexive i.e., $\quad x \sim x$ as $x - x = 0 \in M$

Symmetric i.e., $\quad x \sim y \Rightarrow y \sim x$ as $x - y \in M$

$$\Rightarrow -(x-y) = y - x \in M$$

Transitive i.e., $\quad x \sim y$, $y \sim z \Rightarrow x \sim z$ as

$$x - y \in M \text{ and } y - z \in M \Rightarrow x - y + y - z = x - z \in M$$

$\therefore$ Vectors x, y being equivalent under '~' $\Rightarrow x - y \in M$.

Thus N is divided into mutually disjoint equivalence classes. We denote the set of all such equivalence classes by $\dfrac{N}{M}$.

Let $[x]$ denote the equivalence class which contains the element x. Thus

$$[x] = \{y : y \sim x\} = \{y : y - x \in M\}$$

$$= \{y : y - x = m \;\; for\ some\ m \in M\}$$

$$= \{ y : y = x + m \ \ for\ some\ \ m \in M \} = \{ x + m : m \in M \}$$

Thus $[x]$ is the set of all sums of x and elements of M. The set $[x]$ is called the coset of M determined by x and is usually written as $x + M$. In $\frac{N}{M}$, we define addition and scalar multiplication by

$$(x + M) + (y + M) = (x + y) + M \ ; \ x, y \in N$$

$\alpha (x + M) = (\alpha x) + M$, $\alpha \in F$ over which N is defined.

Here $\frac{N}{M}$ is a vector (linear) space w.r.t. addition and scalar multiplication. Also N is a normed linear spce and exihibits a norm for $\frac{N}{M}$. The zero element of N/M is $0 + M = M$.

The set of all such equivalence classes $\{ x + M : x \in N \}$ referred as $\frac{N}{M}$ is known as the **Factor space or Quotient space** of N w.r.t. N.

Our next theorem shows that if M be a closed linear suspance in a normed linear space N, then $\frac{N}{M}$ can be made into a normed linear space.

Theorem 9 : If M be a closed subspace of a normed linear space N and if the norm of a coset $x + M$ is the quotient space $\frac{N}{M}$ is defined by

$$\| x + M \| = \text{Inf.} \{ \| x + m \| : m \in M \},$$

then $\frac{N}{M}$ is a normed linear space. Also if N is complex (Banach space), then so is $\frac{N}{M}$.

Proof : We verify all the postulates for a norm. $[N_1]$ since $\| x + m \|$ is a non-negative real number and every set of non-negative real numbers is bounded below, it follows that inf $\{ \| x + m \| : m \in M \}$ exists and is non-negative, that is

$$\| x + M \| \geq 0 \ \forall \ x \in N .$$

$[N_2]$: Let $x + M = M$ (the zero element of $\frac{N}{M}$). Then $x \in M$.

Hence $\| x + M \| = \inf \{ \| x + m \| : m \in M, \, x \in M \}$

$$= \inf \{ \| y \| : y \in M \} = 0$$

[$\because$ M being a subspace contains zero vector whose norm is real number 0]

Thus $x + M = M \Rightarrow \|x + M\| = 0$

Conversely, we have

$$\|x + m\| = 0 \Rightarrow \inf \{\|x + m\| : m \in M\} = 0$$

$$\Rightarrow \text{there exists a sequence } < m_k >_{k=1}^{\infty} \text{ in } M$$

Such that $\|x + m_k\| \to 0$ as $k \to \infty$

$$\Rightarrow \lim_{k \to \infty} m_k = -x$$

$\Rightarrow -x \in M$ [Since M is closed and $< m_k >$ is sequence in M converging to $-x$]

$\Rightarrow x \in M$ [$\because$ M is a subspace]

$\Rightarrow x + M = M$ (the zero element of $\frac{N}{M}$)

Thus we have shown that

$\|x + M\| = 0 \Rightarrow x + M = M$ (the zero element of N/M)

$[N_3]$: Let $x + M$, $y + M \in N/M$, then

$\|(x + M) + (y + M)\| = \|(x + y) + M\|$ by definition of addition of coset.

$$= \inf \{\|x + y + m\| : m \in M\} \quad ...(1)$$

$$= \inf \{\|x + y + m + m\| : m \in M, m' \in M\} \quad ...(2)$$

[$\because$ M is a subspace, the sets in (1) and (2) are the same]

$$= \inf \{\|(x + m) + (y + m')\| : m, m' \in M\}$$

$$\leq \inf \{\|x + m\| + \|y + m'\| : m, m' \in M\}$$

[Using N_3 for N, since $x + m$, $y + m' \in N$]

$$= \inf \{\|x + m\| : m \in M\} + \inf \{\|y + m'\| : m \in M\}$$

$$= \|x + M\| + \|y + M\|$$

$[N_4]$: $\|\alpha (x + M)\| = \inf\{\|\alpha x + m\| : m \in M\}$ since $\alpha (x + M) = \alpha x + M$ in $\frac{N}{M}$

$= \inf\{\|\alpha x + m\| : m \in M\}$ if $\alpha \neq 0$

$$= \inf\{|\alpha| \, \|x+m\| : m \in M\}$$

$$= |\alpha| \, \inf\{\|x+m\| : m \in M\}$$

$$= |\alpha| \, \|x+M\|$$

For $\alpha = 0$, the result is obvious.

Hence $\frac{N}{M}$ is a normed linear space.

We now prove that if N is complete, then so is $\frac{N}{M}$. Suppose that $< x_n + M >$ is a Cauchy sequence in $\frac{N}{M}$. Then to show that $< x_n + M >$ is convergent, it is sufficient to prove that this sequence has convergent subsequance.

we can easily find a subsequence of the original Cauchy sequence for a fixed n s.t.

$$\|(x_1 + M) - (x_2 + M)\| < \frac{1}{2}$$

$$\|(x_2 + M) - (x_3 + M)\| < \frac{1}{2^2}$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\|(x_n + M) - (x_{n+1} + M)\| < \frac{1}{2^n}$$

We prove that this sequence is convergent in $\frac{N}{M}$. We begin by choosing any vectory y_1 in $x_1 + M$, and we select y_2 in $x_2 + M$ such that $\|y_1 - y_2\| < \frac{1}{2}$. We next select a vector y_3 in $x_3 + M$. Such that $\|y_2 - y_3\| < \frac{1}{2^2}$ containing in this way, we obtain a sequence $\{y_n\}$ in N such that $\|y_n - y_{n+1}\| < \frac{1}{2^n}$.

Thus for $m < n$, we have

$$\|y_m - y_n\| = \|(y_m - y_{m+1}) + (y_{m+1} - y_{m+2}) + \ldots + (y_{n-1} - y_n)\|$$

$$\leq \|y_m - y_{m+1}\| + \|y_{m+1} - y_{m+2}\| + \ldots + \|y_{n-1} - y_n\|$$

$$< \frac{1}{2^m} + \frac{1}{2^{m+1}} + \ldots + \frac{1}{2^{n-1}}$$

$$= \frac{1}{2^m}\left[1 + \frac{1}{2} + \ldots + \frac{1}{2^{n-m-1}}\right]$$

$$= \frac{1}{2^m}\left[\frac{1-\left(\frac{1}{2}\right)^{n-m}}{1-\frac{1}{2}}\right] = \frac{1}{2^{m-1}}\left[1 - \frac{1}{2^{n-m}}\right]$$

$$< \frac{1}{2^{m-1}} \to 0 \quad \text{as} \quad m \to \infty,$$

which follows that $< y_n >$ is a Cauchy sequence in N.

Since N is complete, there exists a vector y in N such that $y_n \to y$. It now follows from

$$\left\|(y_n + M) - (y + M)\right\| \le \left\|y_n - y\right\| \to 0 \quad \text{as} \qquad n \to \infty$$

that $y_n + M \to y + M$. i.r., $y_n + M$ converges to $y + M$ in $\frac{N}{M}$. Hence $\frac{N}{M}$ is complete.

1.6 Examples of Banach Spaces

We now describe some of the main examples of Banach spaces. In each of these, the linear operations are understood to be defined eithr co-ordinatewise or pointwise, which ever is appropriate in the circumstances.

Example 1 : Show that the linear spaces R (real) and C (complex) are normed linear spaces under the norm $\|x\| = |x|$, $x \in R$ or C as the case may be.

Also show that these spaces are complete and hence Banach spaces.

Solution : R is a normed linear space, since

N_1 : $\|x\| \ge 0 \Rightarrow |x| \ge 0$, which is so, $\forall \ x \in R$

N_2 : $\|x\| = 0 \Leftrightarrow |x| = 0 \Leftrightarrow x = 0, \ \forall \ x \in R$

N_3 : $\|x + y\| = |x + y| \le |x| + |y| = \|x\| + \|y\| \ \forall \ x, \ y \in R$

N_4 : $\|\alpha\, x\| = |\alpha\, x| = |\alpha||x| = |\alpha| \ \|x\|$, α being real or complex.

Similarly C is a normed linear space, since

N_1 : $\|x\| \ge 0 \Rightarrow |x| \ge 0$, $\forall \ x \in C$

N_2 : $\|x\| = 0 \Leftrightarrow |x| = 0 \Leftrightarrow x = 0 \quad \forall \ x \in C$

N_2 : $x, \ y \in C$ and $\bar{x}$, $\bar{y}$ being thier conjugates (complex), We have

$$|x+y|^2 = (x+y)(\overline{x+y}) = (x+y)(\bar{x}+\bar{y}) = x\,\bar{x} + y\,\bar{y} + x\bar{y} + \bar{x}y$$

$$\leq |x|^2 + |y|^2 + 2|x\bar{y}|, \text{ by properties of complex quantities.}$$

$$= |x|^2 + |y|^2 + 2|x||y| \quad \text{as } |\bar{y}| = |y|$$

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

giving $|x+y| \leq |x| + |y|$

$\Rightarrow \quad \|x+y\| \leq \|x\| + \|y\|$

N_4 : $\|\alpha\, x\| = |\alpha\, x| = |\alpha||x| = |\alpha|\;\|x\|$, α being real or complex.

By Theorem 4, every convergent sequence in a normed linear space being a Cauchy sequence, the real (R) or complex (C) normed linear space is complete and hence a Banach space.

Example 2 : Show that the linear spaces R^n (Euclidean) and C^n (Unitary) of n-tuples $x = (x_1, x_2, ..., x_n)$ of real and complex numbers are Banach space under the norm

$$\|x\| = \left\{ \sum_{i=1}^{n} |x_i|^2 \right\}^{\frac{1}{2}}$$

Solution : N_1 : Since each $|x_i| \geq 0$, we have $\|x\| \geq 0$

$$N_2 \;:\; \|x\| = 0 \Leftrightarrow \sum_{i=1}^{n} |x_i|^2 = 0 \Leftrightarrow x_i = 0,\; i = 1,2,......,n$$

$$\Leftrightarrow (x_1, x_2, ..., x_n) = 0$$

$$\Leftrightarrow x = 0$$

N_3 : Let $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ be any two members of C^n (or R^n). Then

$$\|x+y\|^2 = \|(x_1, x_2, ..., x_n) + (y_1, y_2, ..., y_n)\|^2$$

$$= \|(x_1 + y_1,\; x_2 + y_2, ..., x_n + y_n)\|^2$$

$$= \sum_{i=1}^{n} |x_i + y_i|^2 = \sum_{i=1}^{n} |x_i + y_i|\;|x_i + y_i|$$

$$\leq \sum_{i=1}^{n} |x_i + y_i|\;(|x_i| + |y_i|)$$

$$= \sum_{i=1}^{n} |x_i + y_i| \, |x_i| + \sum_{i=1}^{n} |x_i + y_i| \, |y_i|$$

$$= \|x + y\| \, \|x\| + \|x + y\| \, \|y\|, \text{ by Cauchy's inequality.}$$

$$= \|x + y\| \left(\|x\| + \|y\| \right)$$

If $\|x + y\| = 0$, then the above is evidently true. If $\|x + y\| \neq 0$

Then we can divide both sides by it to obtain

$$\|x + y\| \leq \|x\| + \|y\|$$

$$N_4 \; : \; \|\alpha x\| = \|\alpha (x_1, x_2, \ldots., x_n)\|$$

$$= \|\alpha x_1, \alpha x_2, \ldots., \alpha x_n\|$$

$$= \left\{ \sum_{i=1}^{n} |\alpha x_i|^2 \right\}^{1/2} = \left\{ \sum_{i=1}^{n} |\alpha|^2 |x_i|^2 \right\}^{1/2}$$

$$= |\alpha| \left\{ \sum_{i=1}^{n} |x_i|^2 \right\}^{1/2} = |\alpha| \, \|x\|$$

Thus C^n and R^n are normed linear spaces. Again to show that the normal linear spaces R^n and C^n are complete, consider a Cauchy sequence $< x_i >_{i=1}^{\infty}$, i.e., $< x_1, x_2, \ldots., x_n \ldots >$ of points in R^n or C^n, so that x_i being an n-tuple of real or complex numbers, we can write

$$x_m = \left(x_1^{(m)}, x_2^{(m)}, \ldots., x_n^{(m)} \right)$$

so that $x_k^{(m)}$ is the k^{th} co-ordinate of x_m. Let $\in > 0$ be given. Since $< x_m >$ is a Cauchy sequence, there exists a positive integer m_0 such that

$$l_1 m \geq m_0 \Rightarrow \|x_m - x_l\| < \in \Rightarrow \|x_m - x_l\|^2 < \in^2$$

$$\Rightarrow \sum_{i=1}^{n} \left| x_i^{(m)} - x_i^{(l)} \right| < \in^2 \qquad \ldots(1)$$

$$\Rightarrow \left| x_i^{(m)} - x_i^{(l)} \right|^2 < \in^2 \quad (i = 1, 2, \ldots., n)$$

$$\Rightarrow \left| x_i^{(m)} - x_i^{(l)} \right| < \in$$

This shows that the sequence $< x_i^{(m)} >_{m=1}^{\infty}$ is a Cauchy sequence of complex (or real) numbers for each fixed but arbitrary i.

Since C (or R) is complete, each of these sequences converges to a point, say z_i in C (or R) so that

$$\lim_{m\to\infty} x_i^{(m)} = z_i \quad (i = 1,2,....,n) \qquad ...(2)$$

We now show that the Cauchy sequence $< x_m >$ converges to the point $z = (z_1, z_2, ..., z_n) \in C^n$ (or R^n). To prove this, we let $l \to \infty$ in (1). Then by (2), for $m \geq m_0$, we obtain

$$\sum_{i=1}^{n} \left| x_i^{(m)} - z_i \right| < \in^2 \Rightarrow \|x_m - z\|^2 < \in^2 \Rightarrow \|x_m - z\| < \in$$

It follows that the Cauchy sequence $< x_m >$ converges to $z \in C^n$ (or R^n). Hence C^n and R^n are complete spaces and consequently they are Banach spaces.

Example 3 : Let p be a real number such that $1 \leq p < \infty$. Show that the space l_p^n of all n-tuples of scalars with the norm defined by

$$\|x\|_p = \left\{ \sum_{i=1}^{n} |x_i|^p \right\}^{\frac{1}{p}}$$

is a Banach space.

Solution : Let $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ and let α be any scalar. Then it is understood here that l_p^n is a linear space with respect to the operations, $x + y = (x_1 + y_1, ..., x_n + y_n)$ and $\alpha\, x = (\alpha\, x_1, ..., \alpha\, x_n)$. We now show that l_n^p is a normed linear space.

N_1 : $\|x\|_p \geq 0$, obvious since $|x_i| \geq 0$ for each i

$$N_2 \ : \ \|x\|_p = 0 \Leftrightarrow \left\{ \sum_{i=1}^{n} |x_i|^p \right\}^{\frac{1}{p}} = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} |x_i|^p = 0$$

$$\Leftrightarrow |x_i| = 0, \qquad (i = 1,2,....,n)$$

$$\Leftrightarrow x_i = 0, \qquad i = 1,2,....,n$$

$$\Leftrightarrow x = (x_1, x_2, ..., x_n) = 0$$

N_3 : $\|x + y\|_p \leq \|x\|_p + \|y\|_p$, by Minkowski's inequality.

$$N_4 \ : \ \|\alpha\, x\|_p = \|\, \alpha\, (x_1, x_2, ..., x_n)\, \|_p$$

$$=\left\|(\alpha x_1,\alpha x_2,\ldots.,\alpha x_n)\right\|_p$$

$$=\left\{\sum_{i=1}^{n}|\alpha x_i|^p\right\}^{\frac{1}{p}}=\left\{\sum_{i=1}^{n}|\alpha|^p|x_i|^p\right\}^{\frac{1}{p}}$$

$$=\left\{|\alpha|^p\sum_{i=1}^{n}|x_i|^p\right\}^{\frac{1}{p}}$$

$$=|\alpha|\left\{\sum_{i=1}^{n}|x_i|^p\right\}^{\frac{1}{p}}=|\alpha|\,\|x\|_p$$

Thus l_p^n is a normed linear space.

Again to show that l_p^n is complete, let $< x_m >_{m=1}^{\infty}$ be a Cauchy sequence in l_p^n. Since each x_m is an n-tuple of scalars, for convenience, we shall write

$$x_m=\left(x_1^m,x_2^m,\ldots.,x_n^m\right).$$

Let $\in> 0$ be given. Since $< x_m >$ is a Cauchy sequence, there exists a positive integer m_0 such that

$$l,\, m\geq m_0\Rightarrow\|x_m-x_l\|_p<\in\Rightarrow\|x_m-x_l\|_p^p<\in^p$$

$$\Rightarrow\sum_{i=1}^{n}\left|x_i^{(m)}-x_i^{(l)}\right|^p<\in^p \qquad ...(1)$$

$$\Rightarrow\left|x_i^{(m)}-x_i^{(l)}\right|^p<\in^p\quad(i=1,2,\ldots,n)$$

$$\Rightarrow\left|x_i^{(m)}-x_i^{(l)}\right|<\in$$

This shows that for fixed but arbitrary i, the sequence $< x_i^{(m)} >_{m=1}^{\infty}$ is a Cauchy sequence in C (or R) is complete, each of these sequences converges to a point, say z_i, in C (or R) so that

$$\lim_{m\to\infty}x_i^{(m)}=z_i\quad(i=1,2,\ldots,n) \qquad ...(2)$$

It will now be shown that the Cauchy sequence $< x_m >$ converses to the point $z=(z_1,z_2,\ldots,z_n)\in l_p^n$. To prove this, we let $l\to\infty$ (1). Then by (2), for $m\geq m_0$, we obtain

$$\sum_{i=1}^{n}\left|x_i^{(m)}-z_i\right|^p<\in^p\Rightarrow\|x_m-z\|_p^p<\in^p$$

$$\Rightarrow\|x_m-z\|<\in$$

It follows that the Cauchy sequence $< x_m >$ converges to $z \in l_p^n$. Hence l_p^n is complete and therefore it is a Banach spaces.

Example 4 : Consider the linear space of all n-tuples $x = (x_1, x_2, \ldots, x_n)$ of scalars and define the norm by $\|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$

This space is usually denoted by the symbol l_∞^n. Show that $(l_\infty^n, \|.\|_\infty)$ is a Banach space.

Solution : We first prove that l_∞^n is a normed linear space.

N_1 : since each $|x_n| \geq 0$, we have $\|x\|_\infty \geq 0$.

N_2 : $\|x\|_\infty = 0 \Leftrightarrow \max\{|x_1|, |x_2|, \ldots, |x_n|\} = 0$

$$\Leftrightarrow |x_1| = 0, |x_2| = 0, \ldots, |x_n| = 0$$

$$\Leftrightarrow x_1 = 0, x_2 = 0, \ldots, x_n = 0$$

$$\Leftrightarrow (x_1, x_2, \ldots, x_n) = 0$$

$$\Leftrightarrow x = 0$$

N_3 : let $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$

Then $\|x + y\|_\infty = \max\{|x_1 + y_1|, |x_2 + y_2|, \ldots., |x_n + y_n|\}$

$$\leq \max\{|x_1| + |y_1|, |x_2| + |y_2|, \ldots., |x_n| + |y_n|\}$$

$$\leq \max\{|x_1|, |x_2|, \ldots., |x_n|\} + \max\{|y_1|, |y_2|, \ldots., |y_n|\}$$

$$= \|x\|_\infty + \|y\|_\infty$$

N_4 : If α is any scalar, then

$$\|\alpha x\|_\infty = \max\{|\alpha x_1|, |\alpha x_2|, \ldots., |\alpha x_n|\}$$

$$= \max\{|\alpha| |x_1|, |\alpha| |x_2|, \ldots., |\alpha| |x_n|\}$$

$$= |\alpha| \max\{|x_1|, |x_2|, \ldots., |x_n|\}$$

$$= |\alpha| \|x\|_\infty$$

Hence l_∞^n is a normed linear space. We now show that it is a complete space. Let $< x_m >_{m=1}^{\infty}$ be any Cauchy sequence in l_∞^n. Since each x_m is an n-tuple of scalars, we shall write.

$$x_m = \left(x_1^{(m)}, x_2^{(m)}, \ldots, x_n^{(m)}\right)$$

Let $\in > 0$ be given. Then there exits a positive integer m_0 such that $l,\, m \geq m_0 \Rightarrow \|x_m - x_l\|_\infty < \in$

$$\Rightarrow \qquad \max\left\{\left|x_1^{(m)} - x_1^{(l)}\right|, \left|x_2^{(m)} - x_2^{(l)}\right|, \ldots, \left|x_n^{(m)} - x_n^{(l)}\right|\right\} < \in \qquad \ldots(1)$$

$$\Rightarrow \qquad \left|x_i^{(m)} - x_i^{(l)}\right| < \in, \qquad i = 1,2,\ldots,n.$$

This shows that for fixed i, $< x_i^{(m)} >_{m=1}^{\infty}$ is a Cauchy sequence of complex or real numbers. Since C (or R) is complete, it must converges to some $z_i \in C$ (or R). We assert that the Cauchy sequence $< x_n >$ converges to $z = (z_1, z_2, \ldots, z_n)$. The prove this, we let $l \to \infty$ in (1). Then for $m \geq m_0$, we obtain

$\|x_m - z\| < \in$. Thus it follows that the Cauchy sequence $< x_m >$ converges to $z \in l_\infty^n$. Hence l_∞^n is a Banach space.

Example 5 : If $C(X)$ be a linear space of all bounded continous scalar valued function defined on a topological space X. Then show that $C(X)$ is a Banach space under the norm

$$\|f\| = \sup\left\{|f(x)| : x \in X\right\},\ f \in C(X).$$

Solution : Given that $C(X)$ is a linear space, means $C(X)$ is linear under the operations of vector addition and scalar multiplication i.e., $f,\ g \in C(X)$ and α being a scalar, we must have

$$(f+g)(x) = f(x) + g(x) \qquad \ldots(1)$$

$$(\alpha f)(x) = \alpha f(x) \qquad \ldots(2)$$

We now show that $C(X)$ is normed linear space.

N_1 : since $|f(x)| \geq 0\ \forall\ x \in X$, we have

$$\|f(x)\| \geq 0$$

N_2 : $\|f\| = 0 \Leftrightarrow \sup\left\{|f(x)| : x \in X\right\} = 0$

$$\Leftrightarrow |f(x)| = 0\ \forall\ x \in X$$

$$\Leftrightarrow f(x) = 0\ \forall\ x \in X$$

$\Leftrightarrow f$ is a zero function.

N_3 : $\|f+g\| = \sup\left\{|(f+g)(x)| : x \in X\right\}$

$$= \sup\left\{|f(x) + g(x)| : x \in X\right\}$$

$$\leq \sup\left\{|f(x)|+|g(x)| : x \in X\right\}$$

$$\leq \sup\left\{|f(x)| : x \in X\right\} + \sup\left\{|g(x)| : x \in X\right\}$$

$$= \|f\| + \|g\|$$

$$N_4 \; : \; \|\alpha f\| = \sup\left\{|(\alpha f)(x)| : x \in X\right\}$$

$$= \sup\left\{|\alpha f(x)| : x \in X\right\}$$

$$= \sup\left\{|\alpha| \, |f(x)| : x \in X\right\}$$

$$= |\alpha| \sup\left\{|f(x)| : x \in X\right\}$$

$$= |\alpha| \, \|f\|$$

Hence $C(X)$ is a normed linear space.

Finally we prove that $C(X)$ is complete as a metric space. Let $< f_n >$ be any Cauchy sequence in $C(X)$. Then for a given $\in > 0$, there exists a positive integer m_0 such that

$$m, n \geq m_0 \Rightarrow \|f_m - f_n\| < \in$$

$$\Rightarrow \sup\left\{|f_m - f_n(x)| : x \in X\right\} < \in$$

$$\Rightarrow \sup\left\{|f_m(x) - f_n(x)| : x \in X\right\} < \in$$

$$\Rightarrow \left\{|f_m(x) - f_n(x)|\right\} < \in \forall \; x \in X .$$

But this is the Cauchy's condition for uniform convergence of the sequence of bounded continous scalar valued functions. Hence the sequence $< f_n >$ must converge to a bounded continous function f on X. It follows that $C(X)$ is complete and hence it is a Banach space.

Self-Learning Exercise - I

1. Write whether the following statements are true or false :

   (i) If $x, y, z \in N$, N being a normed linear space. Then $d(x+z, \; y+z) = d(x, y)$

   (ii) Every convergent sequence in a normed linear space need not be a Cauchy sequence.

   (iii) Let N be a normed linear space and let $x, y \in N$. Then $\|x-y\| \leq \big| \|x\| - \|y\| \big|$

   (iv) Let $p > 1$, $\dfrac{1}{p} + \dfrac{1}{q} = 1$, $a \geq 0$, $b \geq 0$. Then $ab \leq \dfrac{a^p}{p} + \dfrac{b^q}{q}$ with equality if $a^p = p^q$

(v) Every normed space is metric space but the converse is not universally true.

(vi) Every metric on a linear space can be obtained from a norm.

1.7 Summary

In this unit, we have seen that the notion of the norm of a vector is a generalization of the concept of length. Besides discussing a fairly large number of examples of Banach spaces, we proved an interesting theorem which provides us a very useful method for constructing new normed spaces from a given normed space.

1.8 Answer to Self-Learning Exercise

1. (i) True (ii) False (iii) False

   (iv) False (v) True (Vi) False

1.9 Exercises

1. Define normed spaces, Banach spaces. Give two examples of Banach spaces.
2. Prove that the limit of a convergent sequence in a normed space is unique.
3. Show that the set X of all convergent sequences in a normed space is a normed space. Hence or otherwise show that X is also a linear space.
4. Show that every complete subspace of a normed linear space in closed.
5. Show that every normed space is metric space but the converse is not universally true.
6. Prove that a metric d induced by a norm on a normed space N satisfies

   (i) $d(x+a,\ y+a) = d(x,y)$

   (ii) $d(\alpha x,\ \alpha y) = |\alpha| d(x,y)$

   $\forall\ x,\ y,\ a \in X$ and every scalar α.

□□□

Unit - 2
Bounded Linear Transformations

Structure of the Unit

2.0 Objectives

In previous classes we have studied linear transformation from a linear space to a linear space. We now consider linear transformations from a normed linear space to a normed linear space. In particular we will be interested in questions related to the continuity of such transformations. As an illustration of the use of compactness in analysis, we shall establish basic properties of finite dimensional normed linear spaces.

2.1 Introduction

In calculus we consider the real line R and real valued functions on R (or on a subset of R). Obviously, any such function is a mapping of its domain into R. In functional analysis we consider more general spaces, such as metric spaces and normed spaces, and mappings of these spaces.

In the case of linear space and, in particular, normed spaces, a mapping is called an operator (transformation). In this unit, we consider general properties of bounded linear transformations. Weak convergence is defined in terms of bounded linear transformations.

2.2 Bounded Linear Transformations

If N and N' be two normed linear spaces with the same scalars, then a mapping $T : N \xrightarrow{\text{into}} N'$, is known as an **operator or a transformation** and the value of T at $x \in N$ is denoted by $T(x)$.

The operator T is known as **linear operator (transformation)** if it satisfies the following two conditions :

$$T(x+y) = T(x) + T(y) \quad \text{for all } x, y \in N$$

and $\quad T(\alpha x)=\alpha T(x)$ for real α and $x \in N$.

The above conditions are also equivalent to a single condition

$$T(\alpha x+\beta y)=\alpha T(x)+\beta T(y) \ \forall \alpha, \beta \in F \text{ and } \forall x, y \in N .$$

The transformation T is **bounded** if $\exists$ a real constant $K>0$ s.t.

$$\|T(x)\| \leq K\|x\| \ \forall \ x \in N .$$

The transformation T is **continuous** at a point $x_0 \in N$ if given $\in>0$, $\exists$ a $\delta(\in, x_0)>0$ s.t.

$$\|T(x)-T(x_0)\|<\in \text{ whenever } \|x-x_0\|<\delta .$$

Here T is continuous on N if it is continuous at every point of N. It is uniformly continuous if $\delta(x_0)>0$ is independent of x_0 only s.t.

$$\|T(x)-T(x_0)\|<\in \text{ with } \|x-x_0\|<\delta$$

The norm of a bounded operator (transformation) is defined as

$$\|T\|=\sup \left\{\frac{\|T(x)\|}{\|x\|}: x \neq 0\right\}$$

or equivalenty $\|T\|=\sup \{\|T(x)\|:\|x\| \leq 1\}$

and $\quad \|T\|=\sup \{\|T(x)\|:\|x\|=1 \text{ } if \text{ } N \neq\{0\}\}$

we can also express it as

$$\|T\|=\inf . \{K: K \geq 0 \text{ and } \|T(x)\| \leq K\|x\| \text{ for all } x\}$$

which follows that

$$\|T(x)\| \leq\|T\|\|x\|$$

If $N'=R$ (normed space of reals), then T is known as a **Functional** and denoted by f. A normed linear space consisting of all bounded linear functional over N, is known as a **conjugate space** (or Dual space), denoted by N^*.

Note : All continuous (or bounded) linear transformation of N into N' are denoted by $B(N, N')$, where B stands for bounded.

2.3 General Properties of Bounded (or Continuous) Linear Transformations

Our main purpose in this section is to convert the requirement of continuity into several more useful equivalent forms and to show that the set of all continuous (or bounded) linear transformation of N into N' can itself be made into a normed liear space in a natural way.

Theorem 1 : If T be a linear transformation from a normed linear space N into the normed space N', then the following statements are equivalent :

(i) T is continuous

(ii) T is continuous at the origin i.e., $x_n \to 0 \Rightarrow T(x_n) \to 0$.

(iii) T is bounded i.e., $\exists$ real $K \geq 0$ s.t. $\|T(x)\| \leq K\|x\|$ for all $x \in N$

(iv) If $S = \{x : \|x\| \leq 1\}$ is the closed unit sphere in N, then its image $T(S)$ is bounded set in N'.

Proof : (i) $\Leftrightarrow$ (ii) : Let T be continuous and $< x_n >$ is any sequence in N such that $x_n \to 0$ as $n \to \infty$. Then by continuity of T, we have $x \to 0 \Rightarrow T(x_n) \to T(0) = 0$. Hence T is continuous at the origin.

Conversely, let T be continuous at the origin and $< x_n >$ be any sequence in N such that $x_n \to x \in N$. Then

$$(x_n - x) \to 0 \Rightarrow T(x_n - x) = 0 \qquad [\because T \text{ is continuous at origin}]$$

$$\Rightarrow T(x_n) - T(x) = 0 \Rightarrow T(x_n) = T(x),$$

showing that T is continuous mapping.

(ii) $\to$ (iii) : Let T be continuous at the origin and suppose, if possible T is not bounded that is, there exists no real number K such that $\|T(x)\| \leq K\|x\|$ for every $x \in N$. Then for each positive integer n, we can find a vector x_n, such that

$$\|T(x_n)\| > n\|x_n\|$$

$$\Rightarrow \quad \frac{1}{n\|x_n\|}\|T(x_n)\| > 1$$

$$\Rightarrow \quad \left\|\frac{1}{n\|x_n\|} T(x_n)\right\| > 1 \qquad \text{by } N_4$$

$$\left[\text{Note that } \left|\frac{1}{n\|x_n\|}\right| = \frac{1}{n\|x_n\|}\right]$$

$$\Rightarrow \quad \left\|T\left(\frac{x_n}{n\|x_n\|}\right)\right\| > 1 \qquad [\because \alpha T(x) = T(\alpha x) \text{ for any scalar } \alpha]$$

Now set $y_n = \dfrac{x_n}{n\|x_n\|}$. Then $\|y_n\| = \dfrac{\|x_n\|}{n\|x_n\|} = \dfrac{1}{n} \to 0$ as $n \to \infty$ and so $y_n \to 0$ as $n \to \infty$. But

$T(y_n)$ does not tend to 0, since $\|T(y_n)\| > 1$.

Hence T is not continuous at the origin which is a contradiction. Hence T must be bounded.

Conversely, let T be bounded so that there exists a real number $K > 0$ such that

$$\|T(x)\| \leq K\|x\|, \ \forall\, x \in N \qquad ...(1)$$

Let $< x_n >$ be any sequence in N such that $x_n \to 0$. Then

$$\|x_n\| \to \|0\| = 0$$

Also from (1), $\|T(x_n)\| \leq K\|x_n\| \ \forall\, n$ ...(2)

It follows from (1) and (2) that $\|T(x_n)\| \to 0$ which implies that $T(x_n) \to 0$. We have thus shown that $x_n \to 0 \Rightarrow T(x_n) \to 0$ and consequently T is continuous at the origin.

(iii) $\Leftrightarrow$ (iv) : Assume that $\|T(x)\| \leq K\|x\|$ for every $x \in N$ and let x be any point of the closed unit sphere S so that $\|x\| \leq 1$. Then $\|T(x)\| \leq K$ for all $x \in S$. It follows that $T[S]$ is a bounded set in N'.

Conversely, let $T[S]$ be bounded so that there exists a real number $K \geq 0$ such that

$$\|T(x)\| \leq K \quad \text{for all } x \in S \qquad ...(3)$$

If $x = 0$, then $T(x) = 0$ and so clearly $\|T(x)\| \leq K\|x\|$; and if $x \neq 0$, then

$$\frac{x}{\|x\|} \in S \qquad \left[\because \left\|\frac{x}{\|x\|}\right\| = 1\right]$$

and therefore by (3)

$$\left\|T\left(\frac{x}{\|x\|}\right)\right\| \leq K \Rightarrow \left\|\frac{1}{\|x\|}\ T(x)\right\| \leq K$$

$$\Rightarrow \frac{1}{\|x\|}\ \|\ T(x)\| \leq K$$

$$\Rightarrow \|T(x)\| \leq K\|x\|$$

Thus it is shown that $\|T(x)\| \leq K\|x\|$ for all $x \in N$. Hence T is bounded.

Theorem 2 : If T be a bounded linear transformation of normed space N into normed space N', then the following norms are equivalent :

(i) $\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0 \right\}, \ x \in N$

(ii) $\|T\| = \inf. \{ K : K \geq 0, \|T(x)\| \leq K \|x\| \} \ \forall \ x \in N$

(iii) $\|T\| = \sup \{ \|T(x)\| : \|x\| \leq 1 \}, \ x \in N$

(iv) $\|T\| = \sup \{ \|T(x)\| : \|x\| = 1 \}, \ x \in N$.

Proof : (i) $\Leftrightarrow$ (ii) : Since

$$\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0 \right\}$$

$$\Rightarrow \|T\| \geq \frac{\|T(x)\|}{\|x\|} \ ; \ x \neq 0$$

$$\Rightarrow \|T(x)\| \leq \|T\| \ \|x\| \quad \text{as} \quad T(0) = 0 \qquad ...(1)$$

$\Rightarrow \|T\|$ is one K's satisfying $\|T(x)\| \leq K \|x\|$

$$\Rightarrow \|T\| \geq \inf. \{ K : K \geq 0, \|T(x)\| \leq K \|x\| \} \qquad ...(2)$$

Conversely, for $x \neq 0$, and K satisfying $\|T(x)\| \leq K \|x\|$, we have

$$\frac{\|T(x)\|}{\|x\|} \leq K \Rightarrow \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0, \right\} \leq K$$

$\Rightarrow \|T\| \leq K$ for all K and T independent of x and K

$$\Rightarrow \|T\| \leq \inf. \{ K : K \geq 0, \|T(x)\| \leq K \|x\| \} \qquad ...(3)$$

$\therefore$ (2) and (3) $\Rightarrow \|T\| = \inf. \{ K : K \geq 0, \|T(x)\| \leq K \|x\| \}$.

(ii) $\Leftrightarrow$ (iii) Since $\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0 \right\}$

$$\Rightarrow \|T(x)\| \leq \|T\| \ \|x\| \quad \text{for} \ \|x\| \leq 1$$

$$\Rightarrow \sup \{ \|T(x)\| : \|x\| \leq 1 \} \leq \|T\| \qquad ...(4)$$

Again for an $\in > 0$, $\exists\ x_1 \neq 0$ s.t.

$$\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0 \right\}$$

$$\Rightarrow \frac{\|T(x_1)\|}{\|x_1\|} > \|T\| - \in$$

so that on setting $y = \frac{x_1}{\|x_1\|}$ with $\|y\| = \frac{\|x_1\|}{\|x_1\|} = 1$,

we observe

$$\sup \left\{ \|T(x)\| : \|x\| \leq 1 \right\} \geq \|T(y)\| = \left\| T\left(\frac{x_1}{\|(x_1)\|} \right) \right\|$$

$$= \frac{1}{\|(x_1)\|} \|T(x_1)\|$$

or $\quad \sup \left\{ \|T(x)\| : \|x\| \leq 1 \right\} \geq \|T\| \quad$ ...(5)

$\therefore$ (4) and (5) $\Rightarrow \|T\| = \sup \left\{ \|T(x)\| : \|x\| \leq 1 \right\}$

(iii) $\Leftrightarrow$ (iv) Since as above, we have

$$\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0 \right\}$$

$$\Rightarrow \|T(x)\| \leq \|T\| \, \|x\|$$

$$= \|T\| \quad \text{for } \|x\| = 1$$

$$\Rightarrow \sup \left\{ \|T(x)\| : \|x\| = 1 \right\} \leq \|T\| \quad \text{...(6)}$$

Further $\frac{\|T(x_1)\|}{\|x_1\|} \geq \|T\| - \in$

and $\quad \sup \left\{ \|T(x)\| : \|x\| = 1 \right\} \geq \|T(y)\|$

where $\quad y = \frac{x_1}{\|x_1\|}$

or $\quad \sup \left\{ \|T(x)\| : \|x\| = 1 \right\} \geq \|T\| - \in$

Thus $\sup \{\|T(x)\| : \|x\| = 1\} \geq \|T\|$ ...(7)

From (6) and (7), we have

$$\|T\| = \sup \{\|T(x)\| : \|x\| = 1\}$$

Theorem 3 : If N, N' be normed linear spaces and $B(N, N')$ is the set of all bounded (or continuous) linear transformation from N into N', then $B(N, N')$ is also a normed linear space under the norm

$$\| T \| = \sup \{\|T(x)\| : \|x\| \leq 1\} \ \forall \ x \in N,$$

w.r.t. pointwise linear operations

$(T+S)(x) = T(x) + S(x)$ and $(\alpha T)(x) = \alpha T(x)$, for real α. Also $B(N, N')$ is complete if N' is complete i.e., $B(N, N')$ is a Banach space if N' is a Banach space.

Proof : Since a set S of all linear transformations from a normed linear space N into normed N' is itself a linear space w.r.t. pointwise linear operations. Therefore to show that $B(N, N')$ is a linear space, it suffices to show that $B(N, N')$ is a subspace of S.

Let $T_1, T_2 \in B(N, N')$. Then T_1, T_2 are bounded and so there exists real numbers $K_1 \geq 0$ and $K_2 \geq 0$ such that

$\|T_1(x)\| \leq K_1 \|x\|$ and $\|T_2(x)\| \leq K_2 \|x\|$ for all $x \in N$. For scalar α, β, we have

$$\begin{aligned}\|(\alpha T_1 + \beta T_2)(x)\| &= \|(\alpha T_1)(x) + (\beta T_2)(x)\| \\ &= \|\alpha T_1(x) + \beta T_2(x)\| \\ &\leq \|\alpha T_1(x)\| + \|\beta T_2(x)\| \\ &= |\alpha| \|T_1(x)\| + |\beta| \|T_2(x)\| \\ &\leq |\alpha| K_1 \|x\| + |\beta| K_2 \|x\| \\ &= (|\alpha| K_1 + |\beta| K_2) \|x\|\end{aligned}$$

Thus $\alpha T_1 + \beta T_2$ is bounded and so

$$\alpha T_1 + \beta T_2 \in B(N, N')$$

$\Rightarrow$ $B(N, N')$ is a linear subspace of S.

Now we prove that $B(N, N')$ is a normed linear space.

We verify the norm postulates one by one.

N_1 : Since $\|T\| = \sup\{\|T(x)\| : \|x\| \le 1\}$ and $\|T(x)\| \ge 0$, we conclude that $\|T\| \ge 0$

N_2 : By Theorem 2, $\|T\| = \sup\left\{\frac{\|T(x)\|}{\|x\|} : x \in N,\ x \ne 0\right\}$

$$\therefore \|T\| = 0 \Leftrightarrow \sup\left\{\frac{\|T(x)\|}{\|x\|} : x \in N,\ x \ne 0\right\} = 0$$

$$\Leftrightarrow \frac{\|T(x)\|}{\|x\|} = 0,\ x \in N,\ x \ne 0$$

$$\Leftrightarrow \|T(x)\| = 0,\ x \in N,\ x \ne 0$$

$$\Leftrightarrow T(x) = 0\ \forall\, x \in N$$

$$\Leftrightarrow T = 0 \quad \text{(zero transformation)}$$

N_3 : If $T, U \in B(N, N')$, then

$$\|T+U\| = \sup\left\{\|(T+U)(x) : x \in N,\ \|x\| \le 1\|\right\}$$

$$= \sup\left\{\|T(x)+U(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$\le \sup\left\{\|T(x)\| + \|U(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$\le \sup\left\{\|T(x)\| : x \in N,\ \|x\| \le 1\right\} + \sup\left\{\|U(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$= \|T\| + \|U\|$$

N_4 : If α is any scalar, then

$$\|\alpha T\| = \sup\left\{\|(\alpha T)(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$= \sup\left\{\|\alpha T(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$= \sup\left\{|\alpha|\ \|T(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$= |\alpha|\ \sup\left\{\|T(x)\| : x \in N,\ \|x\| \le 1\right\}$$

$$= |\alpha|\ \|T\|$$

Hence $B(N, N')$ is a normed linear space.

Again, we claim that $B(N, N')$ is complete if N' is complete. Suppose N' is complete and let $< T_n >_{n=1}^{\infty}$ be any Cauchy sequence in $B(N, N')$. Then

$$\|T_m - T_n\| \to 0 \quad \text{as} \quad m, n \to \infty \qquad ...(1)$$

For each $x \in N$, we have

$$\|T_m(x) - T_n(x)\| = \|(T_m - T_n)(x)\|$$

$$\leq \|T_m - T_n\| \, \|x\| \to 0 \qquad \text{by (1)}$$

Hence $< T_n(x) >$ is a Cauchy sequence in N' for each $x \in N$. Since N' is complete, there exists a vector in N', which we denote by $T(x)$, such that $T_n(x) \to T(x)$. This defines a mapping T of N into N'. We now show that T is linear and bounded. If $x, y \in N$ and α, β are scalars, then

$$T(\alpha x + \beta y) = \lim_{n \to \infty} T_n(\alpha x + \beta y)$$

$$= \lim_{n \to \infty} \left[\alpha T_n(x) + \beta T_n(y)\right], \; T_n \text{ being linear } \forall n.$$

$$= \alpha \lim_{n \to \infty} T_n(x) + \beta \lim_{n \to \infty} T_n(y)$$

$$= \alpha T(x) + \beta T(y)$$

This shows that T is linear. To show that T is bounded, we observe that

$$\|T(x)\| = \|\lim T_n(x)\| = \lim \|T_n(x)\| \leq \lim (\|T_n\| \, \|x\|) \quad \text{for all } n$$

$$\leq \sup (\|T_n\| \, \|x\|)$$

$$= (\sup \|T_n\|) \|x\| \qquad ...(2)$$

In view of (1), we observe that

$$\big| \|T_m\| - \|T_n\| \big| \leq \|T_m - T_n\| \to 0 \quad \text{as} \quad m, n \to \infty \quad \text{by (1)}$$

Therefore $< \|T_n\| >$ is a Cauchy sequence of real numbers and hence convergent and bounded. So there exists $K \geq 0$ such that

$$\sup \|T_n\| \leq K \qquad ...(3)$$

From (2) and (3), we have $\|T(x)\| \leq K\|x\|$,

showing that T is bounded. In other words, $T \in B(N, N')$. Finally we show that $T_n \to T$. Let $\in > 0$ be given. Since $< T_n >$ is a Cauchy sequence. There exists a positive m_0 such that

$$m, n \geq m_0 \Rightarrow \|T_m - T_n\| < \in \qquad ...(4)$$

$\Rightarrow \quad \|T_m(x)-T_n(x)\| \leq \|T_m - T_n\| \, \|x\| < \in \|x\|$ for all $m,\, n \geq m_0$ and any vector $x \in N$.

Proceeding to the limit as $m \to \infty$, we find

$$\lim_{m\to\infty} \|T_m(x)-T_n(x)\| = \|T(x)-T_n(x)\| = \|(T-T_n)(x)\| \leq \in \|x\| \qquad ...(5)$$

Since $\lim_{m\to\infty} T_m(x) = T(x)$, as norm is a continuous function

and $\lim_{n\to\infty} \|T_m(x)\| = \|T(x)\|$

(5) $\Rightarrow \|T-T_n\| = \sup \left\{ \frac{\|(T-T_n)x\|}{\|x\|} : x \neq 0 \right\} \leq \in$ for all $n \geq n_0$

$\Rightarrow \|T-T_n\| \to 0$ as $n \to \infty$

$\Rightarrow T_n - T$ as $n \to \infty$

Hence $B(N,\, N')$ is complete if N' is complete.

Theorem 4 : If T be a linear transformation of a normed linear space N into normed linear space N', then inverse of T i.e., T^{-1} exists and is continuous on its domain of definition iff $\exists$ a constant $K \geq 0$ s.t. $K\|x\| \leq \|T(x)\| \; \forall \, x \in N$.

Proof : Assuming that

$$K\|x\| \leq \|T(x)\| \; \forall \, x \in N,\; K \geq 0 \qquad ...(1)$$

is true, we claim that T^{-1} exists and is continuous.

By definition of inverse mapping T^{-1} exists $\Leftrightarrow$ T is one-one.

Taking $x_1, x_2 \in N$, we have

$$T(x_1) = T(x_2) \Rightarrow T(x_1) - T(x_2) = 0 \Rightarrow T(x_1 - x_2) = 0$$

$$\Rightarrow x_1 - x_2 = 0 \Rightarrow x_1 = x_2$$

This implies T is one-one and so T^{-1} exists

$\Rightarrow \quad \exists \; x \in N$ corresponding to each y in the domain of T^{-1}

s.t. $\quad T(x) = y \Leftrightarrow T^{-1}(y) = x \qquad ...(2)$

In view of (2), (1) can be written as

$$K\|T^{-1}(y)\| \leq \|y\| \Rightarrow \|T^{-1}(y)\| \leq \frac{1}{K}\|y\|$$

$\Rightarrow T^{-1}$ is bounded and hence continuous.

Conversely if T^{-1} exists and continuous on its domain $T[N]$, then to each $x \in N \ \exists \ y \in T[N]$ s.t. $T^{-1}(y) = x \Leftrightarrow T(x) = y$ i.e., T is one-one.

Now T^{-1} being continuous, it is bounded and so $\exists$ a positive constant M s.t.,

$$\left\|T^{-1}(y)\right\| \leq M\|y\| \Rightarrow \|x\| \leq M\|T(x)\|$$

$$\Rightarrow K\|x\| \leq \|T(x)\| \text{ for } K = \frac{1}{M} > 0$$

Theorem 5 : If T be a linear transformation from a normed linear space N into normed space N', then T is continuous either at every point or at no point of N.

Proof : Taking arbitrary $x_1, x_2 \in N$ and T continuous at x, to each $\in > 0$, $\exists \ \delta > 0$ s.t.

$$\|x - x_1\| < \delta \Rightarrow \|T(x) - T(x_1)\| < \in \qquad ...(1)$$

Then $\|x - x_2\| < \delta \Rightarrow \|(x + x_1 - x_2) - x_1\| < \delta$

$$\Rightarrow \|T(x + x_1 - x_2) - T(x_1)\| < \in \quad \text{by } (1)$$

$$\Rightarrow \|T(x) + T(x_1) - T(x_2) - T(x_1)\| < \in$$

$$\Rightarrow \|T(x) - T(x_2)\| < \in$$

$\Rightarrow T$ is continuous at x_2.

But x_1, x_2 being arbitrary, T is continuous at all points. Conclusively if T is not continuous at a particular point in N, then it is not continuous at no point of N.

Theorem 6 : If M be a closed linear subspace of a normed linear space N and T be a natural mapping (homomorphism) of N onto $\frac{N}{M}$ s.t. $T(x) = x + M$, then show that T is continuous (or bounded) linear transformation with $\|T\| \leq 1$.

Proof : Given that M is closed and $\frac{N}{M}$ is a normed linear space with the norm of a coset $x + M$ in $\frac{N}{M}$ s.t.

$$\|x + M\| = \inf \{\|x + m\| : m \in M\}$$

we claim that T is linear.

For any $x,\, y \in N$ and $\alpha,\, \beta$ being scalars, we have

$$T(\alpha\, x + \beta\, y) = (\alpha\, x + \beta\, y) + M = (\alpha\, x + M) + (\beta\, y + M)$$

$$= \alpha\,(x + M) + \beta\,(y + M)$$

or $\quad T(\alpha x+\beta y)=\alpha T(x)+\beta T(y) \Rightarrow T$ is linear.

Again, we claim that T is continuous, since

$$\|T(x)\|=\|x+M\|=\inf\{\|x+m\| \; m\in M\}$$

$$\leq \|x+m\| \;\forall\; m\in M$$

$$=\|x\| \text{ if } m=0 \text{ in particular}$$

or $\quad \|T(x)\|\leq 1\,.\,\|T(x)\| \;\forall\; x\in N$ as $0\in M$ and M is a subspace of N.

$\Rightarrow \quad T$ is bounded with bound 1

$\Rightarrow \quad T$ is continuous.

Also $\quad \|T\|=\sup\{\|T(x)\| : \|x\|\leq 1\},\; x\in N$

$$\leq \sup\{\|x\| : \|x\|\leq 1\},\; x\in N$$

$$\leq 1$$

Theorem 7 : If N, N' are two normed linear spaces and T is a continuous linear transformation of N into N' and if M is the null space (kernel) of T, then show that T induces a natural linear transformation T^1 of $\frac{N}{M}$ into N' and that $\|T^1\|=\|T\|$.

Proof : Kernel or Null space of T is defined as

$$\text{Ker } (T) \text{ or } N(T)=\{x : x\in N,\; T(x)=0\}$$

Here is given that Ker (T) or $N(T)=M$.

We first claim that M is closed, since if x be a limit point of M, then $\exists$ a sequence $<x_n>$ in M s.t. $x_n\to x$. But T is continuous, therefore $T(x_n)\to T(x)$. Now $T(x_n)=0\,\forall\, n$ $\Rightarrow T(x)=0 \Rightarrow x\in M \Rightarrow M$ is closed.

Thus M being a closed subspace of N, $\frac{N}{M}$ is a normed linear space with the norm of a cost $x+M$ in N/M s.t.

$$\|x+M\|=\inf.\{\|x+m\| \; : m\in M\}$$

Now defining $T' : N/M\to N'$ and setting $T'(x+M)=T(x)$, we claim that T' is a linear transformation s.t. $\|T\|=\|T'\|$. Taking two elements $x+M$ and $y+M$ of N/M and α, β any scalars, we have

$$T'\left[\alpha(x+M)+\beta(y+M)\right]=T'\left[(\alpha x+M)+(\beta y+M)\right]$$

$$=T'\left[(\alpha x+\beta y)+M\right] \quad \text{by property of coset}$$

$$=T(\alpha x+\beta y)$$

$$=\alpha T(x)+\beta T(y)$$

$$=\alpha T'(x+M)+\beta T'(y+M)$$

Thus T' is linear

Now $\|T'\|=\sup\left\{\|T'(x+M)\| : \|x+M\|\le 1\right\},\ x\in N$

$$=\sup\left\{\|T(x)\| : \text{inf.}\left\{\|x+m\| : m\in M\right\}\le 1\right\},\ x\in N$$

$$=\sup\left\{\|T(x)\| : \|x+m\|\le 1\right\},\ x\in N,\ m\in M$$

$$=\sup\left\{\|T(x)+T(m)\| : \|x+m\|\le 1\right\},\ x\in N,\ m\in M$$

since $m\in M\Rightarrow T(m)=0$, by det. of M

$$=\sup\left\{\|T(x+M)\| : \|x\|\le 1\right\},\ x\in N$$

$=\|T\|$ as $x\in N$, $m\in M\Rightarrow x+M\in N$ and $x\in N$

$\Rightarrow x+0\in N$ and $0\in M$

Theorem 8 : Let N and N' be normed linear spaces over the same scalar field and let T be a linear transformation of N into N'. Then T is bounded if it is continuous.

Proof : Let T be bounded so that there exists a real number $K>0$ such that

$$\|T(x)\|\le K\|x\|\ \forall\ x\in N \qquad ...(1)$$

To show that T is continuous, let $x\in N$ be arbitrary. For any $\in>0$, we choose $\delta=\frac{\in}{K}$. Then for all $y\in N$ such that $\|y-x\|<\delta$, we have

$$\|T(y)-T(x)\|=\|T(y-x)\|$$

$$\le K\|y-x\| \qquad \text{by (1)}$$

$$<K\frac{\in}{K}=\in \quad \text{as} \quad \delta=\frac{\in}{K}>\|y-x\|$$

Hence T is continuous at x. Since x is arbitrary, T is a continuous mapping.

Conversely, let T be continuuous and suppose, if possible, T is not bounded i.e., there exists no real number $\lambda > 0$ such that $\|T(x)\| \le \lambda \|x\| \;\forall\; x \in N$

Then for each positive integer n, there exists a point $x_n \in N$ such that $\|T(x_n)\| > n\|x_n\|$.

For each n, we let

$$y_n = \frac{x_n}{n\|x_n\|}$$

so that $\|y_n\| = \frac{1}{n} \to 0$ as $n \to \infty$ which implies that $y_n \to 0$ as $n \to \infty$. But for every n

$$\|T(y_n)\| = \left\| T\left(\frac{x_n}{n\|x_n\|}\right) \right\| = \left\| \frac{1}{n\|x_n\|}\; T(x_n) \right\|$$

$$= \frac{1}{n\|x_n\|}\;\|T(x_n)\|$$

$$> \frac{n\|x_n\|}{n\|x_n\|} = 1 \quad \text{as} \quad \|T(x_n)\| > n\|x_n\|$$

Which implies $T(y_n)$ does not tend to 0 (i.e., $T(0)$ as $n \to \infty$). Here $< y_n > \to 0$ but $< T_n(y_n) > \nrightarrow T(0)$, is a contradiction showing that T is bounded.

2.4 Weak Convergence

If N be a normed linear space and N^* its dual space, then a sequence $< x_n >$ of N is known as **Weakly convergent** to $x \in N$, $\forall\, f \in N^*$ s.t

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

or simply $x_n \xrightarrow{w} x$

i.e. $< x_n >$ c onverges weakly to x, and x is called as the weak limit of $< x_n >$.

Note that weak convergence means convergence of the sequence of number $a_n = f(x_n)$ for every $f \in N^*$.

Weak convergence has various applications throughout analysis (for instance, in the calculus of variation and the general theory of differential equations). The concept illustrates a basic principle of functional analysis. For applying weak convergence one needs to know certain basic properties, which we state in the following theorem.

Theorem 9 : The weak limit of a sequence is unique.

Proof : Let $< x_n >$ be any sequence. Let if possible $x_n \xrightarrow{w} x_0$ and $x_n \xrightarrow{w} x$, then for an arbitrary linear operater $T \in N^*$, N^* being dual space of normed space N, we have

$$T(x_n) \to T(x_0) \quad \text{and} \quad T(x_n) \to T(x)$$

implying that $T(x_0) = T(x)$ or $T(x_0 - x) = 0$

Choosing T s.t. $\|T\| = 1$ and $T(x_0 - x) = \|x_0 - x\|$, we have

$\|x_0 - x\| = 0$ giving $x = x_0$, i.e., the weak limit is unique.

Corollany 1 : If there are two sequences $< x_n >$ and $< y_n >$ in N s.t.

$x_n \xrightarrow{w} x$ and $y_n \xrightarrow{w} y$, then it is observed that

$$x_n + y_n \xrightarrow{w} x + y$$

and for any scalar α.

$$\alpha x_n \xrightarrow{w} \alpha x \text{ etc.}$$

Corollany 2 : Every subsequence of $< x_n >$ converges weakly to x i.e., if $< x_{nj} >$ be a subsequence of $< x_n >$ of N s.t. $x_n \xrightarrow{w} x_0$, then every subsequence $< x_{nj} >$ converges and has the same limit as the sequence.

2.5 Equivalent Norms

Let a linear space L be made into a normed linear space in two ways and let the two norms of a vector x in L be denoted by $\|x\|_1$ and $\|x\|_2$. Then these norms are said to be **equivalent,** written $\| \;\|_1 \sim \| \;\|_2$, if they generate the same topology on L.

When two norms are equivalent then if $< x_n >$ is a Cauchy sequence w.r.t $\|.\|_1$, it is essentially a Cauchy sequence w.r.t. $\|.\|_2$ and vice-versa. Moreover, in the case of equivalent norms, the class of open sets defined by one is the same as the defined by the other. In other words, in any $\in$ neighbourhood (nba) induced by $\|.\|_1$, a neighbourhood induced by $\| \;\|_2$ is wholly contained and conversely.

Remark : To understand the full implication of the above definition we remind the reader that a norm $\|.\|$ on a linear space L induces the metric $d(x, y) = \|x - y\|$ which in turns induces a topology on L called the metric topology. This is the topology generated by the norm.

Theorem 10 : If N be a normed linear space, then show that the two norms $\| \;\|_1$, $\| \;\|_2$ defined on N are equivalent iff $\exists$ positive real numbers a and b s.t.

$$a\|x\|_1 \le \|x\|_2 \le b\|x\|_1, \; \forall \; x \in N.$$

Proof : If we assume that N_1 is a normed linea space with norm $\| \;\|_1$ and N_2 is a normed linear space with norm $\| \;\|_2$ and that $T(x) = x$ is a linear transformation with domain N_1 and range N_2, then T^{-1} is a linear transformation wtih domain N_2 and range N_1 i.e.,

$$T(x) = x \Rightarrow T^{-1}(x) = x \qquad ...(1)$$

Now, we have

T is continuous $\Leftrightarrow T$ is bounded

$\Leftrightarrow \exists$ positive number b such that

$$\|T(x)\|_2 \le b\|x\|_1, \forall\ x \in N$$

$$\Leftrightarrow \|x\|_2 \le b\|x\|_1, \forall\ x \in N \quad \text{by (1)} \qquad ...(2)$$

T^{-1} is continuous $\Leftrightarrow T^{-1}$ is bounded

$\Leftrightarrow \exists$ is positive number A such that

$$\|T^{-1}(x)\| \le A\|x\|_2, \forall\ x \in N$$

$$\Leftrightarrow \|x\|_1 \le A\|x\|_2, \forall\ x \in N \quad \text{by (1)}$$

$$\Leftrightarrow \frac{1}{A}\|x\|_1 \le \|x\|_2, \forall\ x \in N$$

$$\Leftrightarrow a\|x\|_1 \le \|x\|_2, \forall\ x \in N \quad (\text{on setting } a = \frac{1}{A}) \qquad ...(3)$$

Also T and T^{-1} are continuous

$\Leftrightarrow$ inverse images of open sets in N_2 and N_1 under T, T^{-1} respectively are open in N_1 and N_2

$\Leftrightarrow$ open sets in N_1 are the same as those in N_2 ; T, T^{-1} being identity transformations

$\Leftrightarrow$ Norms $\| \ \|_1$ and $\| \ \|_2$ induces the same topology on N ...(4)

In view of (2), (3) and (4), $\|.\|_1$ and $\|.\|_2$ are equivalent

$\Leftrightarrow$ $\exists$ positive number a and b s.t.

$$a\|x\|_1 \le \|x\|_2 \le b\|x\|_1, \forall\ x \in N.$$

Theorem 11 : On a finite dimensional linear space X, all norms are equivalent.

Proof : Let $\dim X = n$ and $\{x_1, x_2, .., x_n\}$ be any basis for X. Then for each $x \in X$, there is a list of scalars $\alpha_1, \alpha_2, ..., \alpha_n$ such that $x = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$.

Let $\|.\|_1$ and $\|.\|_2$ be two norms defined on X. Then there exists a constant $c > 0$ such that

$$\|x\|_1 = \|\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n\|_1 \ge C(|\alpha_1| + ... + |\alpha_n|)$$

Also $\|x\|_2 = \|\alpha_1 x_1 + ... + \alpha_n x_n\|_2$

$$\leq \|\alpha_1 x_1\|_2 + ... + \|\alpha_n x_n\|_2 \qquad \text{using } N_3$$

$$= |\alpha_1| \|x_1\|_2 + ... + |\alpha_n| \; \|x_n\|_2 \qquad \text{using } N_4$$

$$\leq K\left(|\alpha_1| + ... + |\alpha_n|\right), \quad \text{where}$$

$$K = \max\left\{\|x_1\|_2, ..., \|x_n\|_2\right\}$$

Thus $\alpha \; \|x\|_2 \leq \|x\|_1$, where $\alpha = \dfrac{c}{K} > 0$

The reverse inequality is obtained by interchanging the roles of $\|.\|_1$ and $\|.\|_2$ in the above argument.

2.6 Compactness and Finite Dimension

Some basic properties of finite dimensional normed linear spaces and subspaces are related to the concept of compactness.

2.6.1 Compactness in Normed Spaces

If N be a normed linear space and A is a subset of N, then A is **compact** or sequentially compact if every cover of it has a finite subcover wheras a class $\{G_i\}$ of open subsets of N is known as an open cover o f N if to each point $x \in N$, there corresponds atleast one G_i i.e., $N = \underset{i}{U} G_i$ and a subclass of an open cover, which is an open cover in its own rights, is known as a subcover.

In other words, a subset $A \subset N$ is compact if every sequence in A contains a convergent subsequence whose limit point belongs to A. It should be remembered that an $x \in N$ is a limit point of $A \subset N$, if each open nbd (or open sphere with x as centre) of x contains at least one point of A other than x. In other words, an $x \in N$ is a limit point of $A \subset N$, iff $\exists$ a sequence $< x_n > \to x_0$ where $x_n \in A$, $x_n \neq x_0 \; \forall n$.

2.6.2 Related Theorems

A general property of compact sets is expressed in the following theorem.

Theorem 12 : Every compact subset of a normed linear space is complete.

Proof : Assuming that $< x_n >$ is a Cauchy sequence of a compact subset A of the normed linear space $(N, \|.\|)$. In view of compactness of A, the sequence $< x_n >$ consists of a convergent subsequence say $< x_{ni} > \to x_0 \in A$ for any i, we have

$$\|x_i - x_0\| = \|x_i - x_{ni} + x_{ni} - x_0\|$$

$$\leq \|x_i - x_{ni}\| + \|x_{ni} - x_0\| \qquad \text{by } N_3$$

$$< \frac{\in}{2} + \frac{\in}{2} = \in \to 0 \text{ as } i \to \infty$$

Since $\|x_i - x_{ni}\| < \frac{\in}{2} \to 0$ as $< x_n >$ is a Cauchy sequence and $\|x_{ni} - x_0\| < \frac{\in}{2} \to 0$ as $x_{ni} \to x_0$.

Hence $< x_n >$ is a convergent sequence so that A is complete.

Theorem 13 : Every compact subset of a normed space is bounded but the converse is not true.

Proof : Assuming that a compact subset A of normed space N is not bounded. Every open covering of A consists of unit open sphere $S_1(x_i)$ with centres at each of its points $x_i, i = 1,2,\ldots,n$ s.t.

$$A \subset \bigcup_{i=1} S_i(x_i)$$

Taking $K = \max_{1 \le i \le n} \|x_i\|$ and assuming that $\exists$ an $x \in A$ s.t.

$\|x\| > 1 + K$, Since A is not bounded, we must have an element x_i s.t. $x \in S_1(x_i)$, for $x \in A$ and $A \subset \bigcup_{i=1}^{n} S_i(x_i)$.

As such $\|x - x_i\| < 1$.

Now, $\|x\| = \|x - x_i + x_i\|$

$$\le \|x - x_i\| + \|x_i\|$$

$$< 1 + \max \|x_i\| = 1 + K$$

i.e. $\|x\| \le 1 + K$ which is a contradiction of the fact that $\|x\| > 1 + K$.

Hence A is bounded.

2.7 Reisz Lemma

Theorem 14 : If M be a closed proper subspace of a normed linear space N and a is a real number such that $0 < a < 1$, then $\exists$ a vector $x_0 \in N$ s.t. $\|x_0\| = 1$ and $\|x - x_a\| \ge a \ \forall \ x \in M$.

Proof : Select any $x_1 \in N - M$ and let

$$h = \inf_{x \in M} \left\{ \|x - x_1\| = d(x_1, M) \right\}$$

It is clear that h must be strictly greater than zero for otherwise we would have

$$h = 0 \Rightarrow d(x_1, M) = 0 \Rightarrow x_1 \in \overline{M} = M \qquad [\because \text{M is closed}]$$

Which contradicts the choice of $x_1 \in N - M$.

Now $0 < a < 1 \Rightarrow \frac{1}{a} > 1 \Rightarrow \frac{h}{a} > h$ as $h > 0$.

Hence by definition of infimum, there exists $x_0 \in M$ such that

$$h < \|x_0 - x_1\| \le \frac{h}{a} \quad ...(1)$$

because if $\|x_0 - x_1\|$ were greater than or equal to $\frac{h}{a}$ $\forall\, x_0 \in M$, then it would contradicts the fact that h is the greatest lower bound (infinum) of $\{d(x_0, x_1) : x_0 \in M\}$.

Moreover $x_1 \in N - M$ and $x_0 \in M \Rightarrow x_1 \neq x_0$.

Setting $x_a = \frac{(x_1 - x_0)}{\|x_1 - x_0\|} = K(x_1 - x_0)$ where $K = \|x_1 - x_0\|^{-1} > 0$

Then $\|x_a\| = K\|x_1 - x_0\| = K K^{-1} = 1$.

Now let $x \in M$ be arbitrary. Then $K^{-1}x + x_0 \in M$ also and so

$$\|x - x_a\| = \|x - K(x_1 - x_0)\|$$

$$= K\left\|\left(K^{-1}x + x_0\right) - x_1\right\| \ge K h \quad ...(2)$$

$\because$ $h = \inf\limits_{x \in M} \|x - x_1\|$ and $K^{-1}x + x_0 \in M$, we have

$$\left\|\left(K^{-1}x + x_0\right) - x_1\right\| \ge h$$

But $K h = \|x_1 - x_0\|^{-1} h \ge a$ by (1) ...(3)

From (2) and (3), we have

$$\|x - x_a\| \ge a \text{ for all } x \in M.$$

Theorem 15 : Let N be a normed linear space, and suppose the set $S = \{x \in N : \|x\| = 1\}$ is compact. Then N is finite dimensional.

Proof : We know that in a metric space, a subset is compact iff it is sequentially compact is iff every sequence has a convergent subsequence. Since S is given to be compact, every sequence in S must have a convergent subsequence. Suppose, if possible, N is not finite dimensional. Choose $x_1 \in S$ and let N_1 be the subspace spanned by x_1. Then N_1 is proper subspace of N. Since N_1 is finite dimensional and therefore it is closed. Hence by Reisz Lemma there exists a vector $x_2 \in S$ such that $\|x_2 - x_1\| \ge \frac{1}{2}$.

Let N_2 be closed proper subspace of N generated by x_1, x_2, then as before there must exists $x_3 \in S$ such that

$$\|x_3 - x\| \geq \frac{1}{2} \quad \text{if} \quad x \in N$$

Proceeding inductively, we obtain an infinite sequence $< x_n >$ of vectors in S such that $\|x_n - x_m\| \geq \frac{1}{2}$.

This sequence can therefore have no convergent subsequence. But this contradicts the hypothesis that S is compact. Hence N must be finite dimensional.

Self-Learning Exercise - I

1. Write whether the following statements are true or false.

   (a) We may define the norm of a bounded linear transformation T on N into N' by

   $$\|T\| = \sup \left\{ \frac{\|T(x)\|}{\|x\|} : x \neq 0 \right\}, \; x \in N.$$

   (b) The identity operator $I : N \to N$ on a normed space $N \neq \{0\}$ is not bounded.

   (c) The zero operator $0 : N \to N'$ on a normed space N is bounded and has noirm $\|0\| = 0$.

   (d) Every subsequence of $< x_n >$ converges weakly to x, where $x_n \xrightarrow{w} x$.

   (e) Two norms $\| \, \|_1, \| . \|_2$ defined on a normed space N are equivalent iff $\exists$ positive real number a and b s.t.

   $$a\|x\|_1 \leq \|x\|_2 \geq b\|x\|_1 \;\; \forall \; x \in N.$$

2. What is the zero element of the linear spac e $B(N, N')$.

2.8 Summary

In this unit, we have seen that the concept of linear transformation can be generalised from linear spaces to normed linear spaces.

We know that in calculus are defines different types of convergence (ordinary, conditional, absolute and uniform convergence). The yields greater flexibility in the theory and application of sequence and series. In functional analysis, the situation is similar.

2.9 Answers to Self-Learning Exercise

1. (a) True (b) False (c) True (d) True (e) False.

2. The zero operator $0 : N \to N'$

2.10 Exercises

1. Let N be the normed space of all polynomials on $J = [0, 1]$ with norm given $\|x\| = \max |x(t)|;\ t \in J$. A differentiation operator T is defined on N by

$$T\,x(t) = x'(t),$$

where the prime denotes differentiation with respect to t. Prove that this operator is linear but not bounded.

2. Let X, Y and Z be normed spaces and let $T : X \to Y$ and $S : Y \to Z$ be two bounded linear transformation. Then prove that $SoT : X \to Z$ is bounded linear transformation and $\|SoT\| \leq \|S\|\,\|T\|$.

3. If T be a linear transformation of normed space N into normed space N', then inverse of T i.e., T^{-1} exists and is continuous on its domain of definition iff $\exists$ a constant $K \geq 0$ s.t. $K\|x\| \leq \|T(x)\|\ \forall\ x \in N$.

4. If T is a linear transformation of a normed linear space N into a normed linear space N', then show that T is bounded iff T maps bounded sets in N into bounded sets in N'.

5. Give an example to show that a closed and bounded subset of normed linear space need not be compact.

□□□

Unit - 3
Fundamental Theorems of Functional Analysis

Structure of the Unit

3.0 Objectives

This unit contains the basis of the more advanced theory of normed and Banach spaces without which the usefulness of these spaces and their applications would be rather limited. The three import theorems included in this unit are, the open mapping theorem, the uniform boundedness theorem and the closed graph theorem.

3.1 Introduction

Banach space in a linear space which is also, in a speacial way, a complete metric space. This combination of algebraic and metric structures opened the posibility of studying of linear transformation of one Banach space into another which had the additional property of being continuous.

Most of our work in this unit centres around three fundamental theorems related to continuous linear transformation between Banach spaces. These theorems together with The Hahn-Banach theorem are often regarded as the cornerstones of functional analysis.

3.2 Multilinear Mappings

Definition : Let $X_1, X_2, \ldots, X_n$, Y be linear spaces over the same field of scalars K. A mapping

$$f : X_1 \times X_2 \times \ldots \times X_n \to Y$$

is said to be multilinear if for each $i \in \underline{n}$ the mapping

$$x_i \to f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$$

of X_i into Y is linear.

Definition : Let $X_1, X_2, \ldots, X_n$ be normed linear spaces. Then a mapping

$$\|.\| : X_1 \times X_2 \times \ldots \times X_n \to R$$

given by

$$\|(x_1, x_2, \ldots, x_n)\| = \max \{\|x_1\|, \|x_2\|, \ldots, \|x_n\|\}$$

is a norm on $X_1 \times X_2 \times \ldots \times X_n$.

The product space $X_1 \times X_2 \times X_3 \ldots \times X_n$ of normed linear spaces $X_1, X_2, \ldots, X_n$ is endowed with the norm defined above.

The following theorem is a generalization of Theorem 8-Unit-2.

Theorem 1 : Let $X_1, X_2, \ldots, X_n$, Y be normed linear spaces over the same field of scalars and let

$$f : X_1 \times \ldots \times X_n \to Y$$

be a multilinear mapping. Then f is continuous iff there exists a number $m > 0$ such that

$$\|f(x_1, x_2, \ldots, x_n)\| \le m \|x_1\| \, \|x_2\| \ldots \|x_n\|$$

for any $(x_1, x_2, \ldots, x_n) \in X_1 \times X_2 \times \ldots \times X_n$.

Proof : Let first the given condition be satisfied and let $(a_1, a_2, \ldots, a_n)$ be any point in $X_1 \times X_2 \times \ldots \times X_n$.

Since f is linear with respect to each of its variable, therefore

$$(x_1, x_2, \ldots, x_n) - f(a_1, a_2, \ldots, a_n) = f(x_1 - a_1, x_2, \ldots, x_n)$$

$$+ f(a_1, x_2 - a_2, x_3, \ldots, x_n) + \ldots + f(a_1, \ldots, a_{n-1}, x_n - a_n)$$

$$= \sum_{i=1}^{n} (a_1, \ldots, a_{i-1}, x_i - a_i, x_{i+1}, \ldots, x_n)$$

and hence using triangle inequality

$$\|f(x_1, x_2, \ldots, x_n) - f(a_1, a_2, \ldots, a_n)\|$$

$$\le \sum_{i=1}^{n} \|(a_1, \ldots, a_{i-1}, x_i - a_i, x_{i+1}, \ldots, x_n)\|$$

$$\le \sum_{i=1}^{n} \left(m \|a_1\| \ldots \|a_{i-1}\| \, \|x_i - a_i\| \, \|x_{i+1}\| \ldots \|x_n\|\right)$$

Let us assume that $\|x_i - a_i\| \le \in$ for $i \in \underline{n}$. Then $\|x_i\| \le \|a_i\| + \in$ and we can determine $\delta > 0$ such that $\|x_i\| \le \|a_i\| + \in \le \delta$ for $i \in \underline{n}$, and hence

$$\|f(x_1, \ldots, x_n) - f(a_1, \ldots, a_n)\| \le m \delta^{n-1} \sum_{i=1}^{n} \|x_i - a_i\|$$

$$\le m n \delta^{n-1} \in$$

Since for small values of $\in$ the choice of δ is independent of $\in$. We obtain that f is continuous at $(a_1, a_2, \ldots, a_n)$.

Next let f be continuous at the point $(0, 0, \ldots, 0)$. Then there exists a number $\in > 0$ such that

$$\|f(x_1, x_2, \ldots, x_n) - f(0, 0, \ldots, 0)\| \leq 1$$

for $\quad \|(x_1, x_2, \ldots, x_n) - (0, 0, \ldots, 0)\| \leq \in$

Let now $(x_1, x_2, \ldots, x_n)$, $x_1 \neq 0, \ldots, x_n \neq 0$ be any point of $X_1 \times X_2 \times \ldots \times X_n$. If

$$y_1 = \frac{\in x_1}{\|x_1\|},\ y_2 = \frac{\in \|x_2\|}{\|x_2\|}, \ldots, y_n = \frac{\in \|x_n\|}{\|x_n\|}$$

Then $\|(y_1, y_2, \ldots, y_n)\| = \in$ and $\|f(y_1, y_2, \ldots, y_n)\| \leq 1$

$$\Rightarrow \quad \left\| \frac{\in^n}{\|x_1\| \|x_2\| \ldots \|x_n\|} f(x_1, x_2, \ldots, x_n) \right\| = \leq 1$$

$$\Rightarrow \quad \|f(x_1, x_2, \ldots, x_n)\| \leq m \|x_1\| \|x_2\| \ldots \|x_n\|$$

where $\quad m = \dfrac{1}{\in^n}$.

If $x_1 = 0$ or $x_2 = 0$ or $x_n = 0$, then $f(x_1, x_2, \ldots, x_n) = 0$ and the preceeding inequality still holds.

Hence the theorem.

3.3 Open Mapping Theorem

The open mapping theorem states conditions under which a bounded linear operator is an open mapping. The present theorem exhisits reason why Banach spaces are more satisfactory than incomplete normed spaces. The proof of the open mapping theorem will be based on Baire's category theorem.

Let us begin by introducing the concept of an open mapping.

Definition : If X and Y are two topological spaces. Then a map $f : X \xrightarrow{\text{into}} Y$ is known as an open mapping if $\forall$ open set V of X, the set $f(V)$ is open in Y. In other words, f is open at a point $x \in X$ if $f(\text{U})$ contains a *nbd* of $f(x)$ whenever U is a *nbd* of x. Evidently f is open iff f is open at every point of X. Thus a linear mapping of one topological space into another is open iff it is open at the origin. It should also be noted that a one-one continuous mapping f of X onto Y is homeomorphism when f is open.

If B and B' are Banach spaces (i.e. complete normed spaces) then the open spheres with radius r and centre at x are denoted respectively by $S(x,r)$ or $S_r(x)$ and $S^1(x,r)$ or $S_r^1(x)$ whereas the open spheres in B, B' respectively are denoted by S_r, S_r^1 with radius r and centre at origin. As such the unit open spheres with centre at origin are S_1, S_1^1 respectively in B, B'. It is easy to see that

$$S(x,r) = x + S_r \qquad \text{and} \qquad S_r = r\,S_1$$

For we have

$$y \in S(x,r) \Leftrightarrow \|y - x\| < r$$

$$\Leftrightarrow \|z\| < r \quad \text{and} \quad y - x = z$$

$$\Leftrightarrow y = x + z \ \text{ and } \ \|z\| < r$$

$$\Leftrightarrow y \in x + S_r$$

and $$S_r = \{x : \|x\| < r\} = \left\{x : \frac{\|x\|}{r} < 1\right\}$$

$$= \{r\,y : \|y\| < 1\} = r\,S_1 .$$

The following lemma is the key to the proof of the open mapping theorem.

Lemma : If B and B' be Banach spaces and T a continuous linear transformation of B onto B', then the image of every open sphere centred at origin in B contains an open sphere centred at origin in B'

Proof : Taking S_r, S_r^1 as open spheres with radius r and centred at origin in B, B' respectively and S_1 an open unit sphere, we have

$$S_r = r\,S_1$$

which yields

$$T(S_r) = T(r\,S_1) = r\,T(S_1) \qquad ...(1)$$

Hence it suffies to prove that $T(S_1)$ contains some S_r^1.

We begin by proving that $\overline{T(S_1)}$ contains some S_r^1. For each positive integer n, consider open spheres S_n in B. Then it is clear that $B = \bigcup_{n=1}^{\infty} S_n$.

Since T is onto, this gives

$$B' = T(B) = T\left(\bigcup_{n=1}^{\infty} S_n\right) = \bigcup_{n=1}^{\infty} T(S_n)$$

Since B' is complete. it is of second category. Hence by Baire category theorem, $\overline{T(S_{n0})} \neq \phi$ for some n_0, that is $\overline{T(S_{n0})}$ has an interior point y_0 which may be assumed to lie in $T(S_{n0})$.

[The existence of such a point y_0 is proved as follows :

y is an interior point of $\overline{T(S_{n0})}$

$\Rightarrow$ there exists an open set G such that $y \in G \subset \overline{T(S_{n0})}$].

But $y \in \overline{T(S_{n0})} \Rightarrow y$ is an adherent point of $T(S_{n0})$

$\Rightarrow$ the nbd G of y must contain a point y_0 of $T(S_{n0})$.

Thus $y_0 \in T(S_{n0})$ is such that $y_0 \in G \subset \overline{T(S_{n0})}$ which implies that y_0 is an interior point of $\overline{T(S_{n0})}$.

The mapping of $f : B' \to B'$ s.t. $f(y) = y - y_0$ is a homomorphism. For f is evidently one-one onto and if $y_n \in B'$ is such that $y_n \to y$, then

$$f(y_n) = y_n - y_0 \to y - y_0 = f(y)$$

and $$\bar{f}^1(y_n) = y_n + y_0 \to y + y_0 = \bar{f}^1(y)$$

so that f and f^{-1} are both continuous. We use the mapping f to show that 0 is the interior point of $\overline{T(S_{n0})} - y_0$. We have y_0 is an interior point of $\overline{T(S_{n0})}$.

$\Rightarrow$ there exists an open set G such that $y_0 \in \subset \overline{T(S_{n0})}$

$\Rightarrow$ $f(y_0) \in f(G) \subset f\left[\overline{T(S_{n0})}\right]$

$\Rightarrow$ $y_0 - y_0 = 0 \in f(G) \subset \overline{T(S_{n0})} - y_0$ ...(2)

$\Rightarrow$ 0 is an interior point of $\overline{T(S_{n_0})} - y_0$

[$\because f$ is an open map (being a homeomorphism) $f(G)$ is an open set in B' and so $\overline{T(S_{n0})} - y_0$ is a nbd of 0]

we assert that $T(S_{n0}) - y_0 \subset T(S_{2n0})$

Let $y \in T(S_{n0}) - y_0$. Then there exists $x \in S_{n0}$ such that

$$y = T(x) - y_0.$$

But $y_0 \in T(S_{n0})$ implies that $y_0 = T(x_0)$ for some $x_0 \in S_{n0}$.

Thus $y = T(x) - T(x_0) = T(x - x_0)$, ...(3)

where x, $x_0 \in S_{n0}$. Also

$$x,\ x_0 \in S_{n0} \Rightarrow \|x\| < n_0, \qquad \|x_0\| < n_0$$

$$\Rightarrow \|x - x_0\| \le \|x\| + \|x_0\| < 2\,n_0$$

$$\Rightarrow x - x_0 \in S_{2n0}$$

$$\Rightarrow T(x - x_0) \in T[S_{2n0}]$$

$$\Rightarrow y \in T[S_{2n0}] \qquad \text{by (3)}$$

Thus we have shown that

$y \in T[S_{n0}] - y_0 \Rightarrow y \in T[S_{2n0}]$ and therefore

$$T[S_{n0}] - y_0 \subset T[S_{2n0}] = 2n_0 T(S_1) \qquad \text{by (1)}$$

$$\Rightarrow \qquad \overline{T(S_{n0}) - y_0} \subset \overline{2n_0 T(S_1)} \qquad [\because A \subset B \Rightarrow \overline{A} \subset \overline{B}] \qquad \text{...(4)}$$

Since f is homeomorphism,

$$f\left[\overline{T(S_{n0})}\right] \subset \overline{f[T(S_{n0})]} \qquad \text{as} \qquad f(\overline{A}) = \overline{f(A)}$$

$$\Rightarrow \qquad \overline{T(S_{n0})} - y_0 = \overline{T(S_{n0}) - y_0} \subset \overline{2n_0 T(S_1)} \qquad \text{...(5)}$$

by definition of f and (4).

The mapping

$$g : B' \to B' \text{ s.t. } g(x) = 2n_0 x$$

is easily seen to be a homeomorphism and so

$$\overline{g[T(S_1)]} = g\overline{[T(S_1)]}$$

$$\Rightarrow \qquad \overline{2n_0\, T(S_1)} = 2n_0\, \overline{T(S_1)} \qquad \text{by definition of } g,$$

which by (5) implies that

$$\overline{T(S_{n0})} - y_0 \subset 2n_0\, \overline{T(S_1)} \qquad \text{...(6)}$$

It follows from (2) and (6) that O is an interior point of $\overline{T(S_1)}$.

Hence there exists an open sphere $S^1_\in$ with radius $\in > 0$ and centered at origin in B' s.t. $S^1_\in \subset \overline{T(S_1)}$ ...(7)

We complete the proof by showing that $S^1_\in \, C\, T(S_3)$, which is clearly equivalent to $S^1_{\frac{\in}{3}} \subset T(S_1)$.

Let y be an arbitrary point of $S_{\in}^{1}$ so that $\|y\| < \in$. Then by (7), $y \in \overline{T(S_1)}$, which implies that y is an adherent point of $T(S_1)$ and hence there exists a vector $y_1 \in T(S_1)$ such that $\|y - y_1\| < \frac{\in}{2}$.

But $y_1 \in T(S_1) \Rightarrow y_1 = T(x_1)$ for some $x_1 \in S_1$ so that $\|x\| < 1$.

Again we observe from (7), we have $S_{\frac{\in}{2}}^{1} \subset \overline{T\left(S_{\frac{1}{2}}\right)}$

and since $\|y - y_1\| < \frac{\in}{2}$, we have

$$y - y_1 \in S_{\frac{\in}{2}}^{1} \subset \overline{T\left(S_{\frac{1}{2}}\right)}.$$

Therefore as before there exists a vector y_2 in $T\left(S_{\frac{1}{2}}\right)$ such that

$$\|(y - y_1) - y_2\| < \frac{\in}{2^2} \quad \text{or} \quad \|y - (y_1 + y_2)\| < \frac{\in}{2^2},$$

where $y_2 = T(x_2)$ and $\|x_2\| < \frac{1}{2}$.

Continuing in this way, we obtain a sequence $< x_n >$ in B such that $\|x_n\| < \frac{1}{2^{n-1}}$, and

$$\|y - (y_1 + y_2 + \ldots + y_n)\| < \frac{\in}{2^n} \qquad \ldots(8)$$

where $y_n = T(x_n)$. If we put,

$s_n = x_1 + x_2 \ldots + x_n$, then

$$\|s_n\| = \|x_1 + x_2 + \ldots + x_n\|$$

$$\leq \|x_1\| + \|x_2\| + \ldots + \|x_n\|$$

$$\leq 1 + \frac{1}{2} + \ldots + \frac{1}{2^{n-1}} < 2 \qquad \ldots(9)$$

Also, for $n > m$, we have

$$\|s_n - s_m\| = \|x_{m+1} + x_{m+2} + \ldots + x_n\|$$

$$\leq \|x_{m+1}\| + \|x_{m+2}\| + \ldots + \|x_n\|$$

$$< \frac{1}{2^m} + \frac{1}{2^{m+1}} + \ldots + \frac{1}{2^{n-1}}$$

$$= \frac{\frac{1}{2^m}\left(1 - \frac{1}{2^{n-m}}\right)}{1 - \frac{1}{2}} \qquad \text{(summing the G.P.)}$$

$$= \frac{1}{2^{m-1}} - \frac{1}{2^{n-m-1}} \to 0 \quad \text{as} \quad m, n \to \infty$$

Hence $< s_n >$ is a Cauchy sequence in B and since B is complete, there exists a vector x in B such that

$$\lim_{n \to \infty} s_n = x$$

and so $\|x\| = \|\lim s_n\| = \lim \|s_n\| \le 2$ by (9) <3, which implies that $x \in S_3$. Now

$$y_1 + y_2 + \ldots + y_n = T(x_1) + T(x_2) + \ldots + T(x_n)$$

$$= T(x_1 + x_2 + \ldots + x_n) = T(s_n) \qquad \ldots(10)$$

since T is continuous

$$x = \lim s_n \Rightarrow T(x) = \lim T(s_n)$$

$$= \lim (y_1 + y_2 + \ldots + y_n) \qquad \text{by (10)}$$

$$= y \qquad \text{by (8)}$$

Thus $\quad y = T(x)$, where $\|x\| < 3$, so that $y \in T(S_3)$

we have now proved that

$$y \in S'_{\epsilon} \Rightarrow y \in T(S_3)$$

and so $S^1_{\epsilon}\ C\,T(S_3)$, y being an arbitrary point in S^1_{ϵ}.

$$\frac{1}{3} S^1_{\epsilon} \subset \frac{1}{3} T(S_3)$$

$$\Rightarrow \qquad \frac{S^1_{\epsilon}}{3} \subset T(S_1) \qquad \text{by (1)}$$

Hence $T(S_1)$ contains an open sphere centred at origin in B'.

Theorem 2 [The open mapping theorem] :

Let B and B' be Banach spaces. If T is a continuous linear transformation of B onto B', then T is an open mapping.

Proof : We are given that the linear transformation

$$T : B \to B'$$

is continuous and onto. We claim that T is an open map i.e., $T(G)$ is an open set in B' for every open set G in B.

Let $y \in T(G)$ be arbitrary. Then $y = T(x)$ for some $x \in G$. Since G is open set in B, there exists an open sphere $S(x,r)$ in B centred at x such that $S(x,r) \subset G$. But as remarked earlier, we can write $S(x,r) = x + S_r$, where S_r is an open sphere in B centered at origin. Thus

$$x + S_r \subset G \qquad ...(1)$$

By our lemma, there exists an open sphere $S_{\in}^1$ in B' centered origin such that $S_{\in}^1 \subset T(S_r)$.

$\therefore \qquad y + S_{\in}^1 \subset y + T(S_r) = T(x) + T(S_r) = T(x + S_r)$

or $\qquad S^1(y, \in) \subset T(x + S_r) \subset T(G), \qquad [\because y + S_{\in}^1 = S^1(y, \in)]$

by (1).

This implies that to each $y \in T(G) \exists$ an open sphere B' centered at y and contained in $T(G)$. Consquently $T(G)$ is open.

3.4 Closed Graph Theorem

In this section, we define closed linear transformation on normed linear spaces and consider some of their properties, in particular in connection with the important closed graph theorem.

Definition : Let X and Y be any non empty sets and let $f : X \to Y$ be a mapping with domain X and range in Y. Then the **graph of** f is defined to be that subset of $X \times Y$ which consists of all ordered pairs of the form $(x, f(x))$ i.e., if D be a subset of X and $T : D \to Y$, then the graph of T is defined as

$$T_G = \{(x, T(x)) : x \in D\}.$$

In the case of two normed linear spaces N, N' with $D \subset N$ and $T : D \to N'$, then the graph of the linear transformation T is given by

$$T_G = \{(x, T(x)) : x \in D\}.$$

Remark : If N, N' are two normed linear spaces, then $N \times N'$ is also a normed linear space with co-ordinatewise linear operation under the norm

$$\|(x, y)\| = \left(\|x\|^p + \|y\|^p\right)^{\frac{1}{p}}$$

with $x \in N$, $y \in N'$ and $1 \le p < \infty$.

In our future discussion, we shall mostly use the above norm with $p=1$ i.e., $\|(x,y)\| = \|x\| + \|y\|$.

Definition (Closed Linear Transformation) :

Let N and N' be normed linear spaces and let D be a subspace of N. Then a linear transformation $T: D \to N'$ is said to be closed iff $x_n \in D$, $x_n \to x$ and $T(x_n) \to y \Rightarrow x \in D$ and $y = T(x)$.

Theorem 3 : Let N and N' be normed linear spaces and D be a subspace of N. Then a linear transformation $T: D \to N'$ is closed iff its graph T_G is closed.

Proof : Assuming that T is a closed linear transformation, we claim that its graph T_G is closed i.e., T_G contains all of its limit points. T_G is defined as

$$T_G = \{(x, T(x)) : x \in D\}$$

Taking (x,y) as a limit point of T_G, $\exists$ a sequence $< x_n, T(x_n) >$, $x_n \in D$ of points in T_G converging to (x,y) i.e.,

$$< x_n, T(x_n) > \to (x,y)$$

$$\Rightarrow \quad \|(x_n, T(x_n)) - (x,y)\| \to 0$$

$$\Rightarrow \quad \|(x_n - x), (T(x_n) - y)\| \to 0$$

$$\Rightarrow \quad \|x_n - x\| + \|T(x_n) - y\| \to 0 \qquad \text{[see remark for the norm on } N \times N' \text{]}$$

$$\Rightarrow \quad \|x_n - x\| \to 0 \text{ and } \|T(x_n) - y\| \to 0$$

$$\Rightarrow \quad x_n \to x \text{ and } T(x_n) \to y$$

$$\Rightarrow \quad x \in D \text{ and } T(x) = y, \ T \text{ being closed.}$$

$$\Rightarrow \quad (x,y) \in T_G, \text{ in view of definition of graph.}$$

$$\Rightarrow \quad T_G \text{ is closed.}$$

Conversely, let the graph T_G of T be closed. We claim that T is a closed linear transformation.

Let $\quad x_n \in D$, $x_n \to x$ and $T(x_n) \to y$.

But $\quad T_G = \overline{T_G}$, since T_G is given to be closed.

$$\Rightarrow \quad (x_n, T(x_n)) \to (x,y) \in \overline{T_G} = T_G$$

$$\Rightarrow \quad (x,y) \in T_G \Rightarrow x \in D \text{ and } y = T(x) \text{ by definition of } T_G$$

$\Rightarrow \quad T$ is a closed linear transformation.

Theorem 4 [The Closed Graph Theorem] :

If B and B' are Banach spaces and T is a linear transformation of B into B', then T is continuous $\Leftrightarrow$ its graph is closed.

Proof : Assuming that T is continuous and T_G is its graph

i.e., $\quad T_G = \{(x, T(x)) : x \in B\}$

We claim that T_G is closed i.e., $T_G = \overline{T_G}$.

Since $T_G \subset \overline{T_G}$ always. We need only prove $\overline{T_G} \subset T_G$. So let $(x, y) \in \overline{T_G}$. Then (x, y) is limit point of T_G. Hence there exists a sequence $< x_n, T(x_n) >$ in T_G such that $< x_n, T(x_n) > \to (x, y)$, which implies that $x_n \to x$ and $T(x_n) \to y$. But, since T is continuous, $x_n \to x$ implies $T(x_n) \to T(x)$ and so $y = T(x)$. This shows that $(x, y) = (x, T(x)) \in T_G$ and $\overline{T_G} \subset T_G$.

Conversely, if T_G is closed, then we claim that T is continuous. We denote by B_1 the linear space B renormed by

$$\|x\|_1 = \|x\| + \|T(x)\|, \; x \in B.$$

We first show that this is actually a norm, since

$N_1 \; : \|x\|_1 \geq 0 \quad$ as $\|x\| \geq 0$, $\|T(x)\| \geq 0$

$N_2 \; : \|x\|_1 = 0 \Leftrightarrow \|x\| + \|T(x)\| = 0 \Leftrightarrow \|x\| = 0$, $\|T(x)\| = 0 \Leftrightarrow x = 0$

$$\begin{aligned} N_3 \; : \|x+y\|_1 &= \|x+y\| + \|T(x+y)\| \\ &= \|x+y\| + \|T(x) + T(y)\| \\ &\leq \|x\| + \|y\| + \|T(x)\| + \|T(y)\| \\ &= (\|x\| + \|T(x)\|) + (\|y\| + \|T(y)\|) \\ &= \|x\|_1 + \|y\|_1 \end{aligned}$$

$$\begin{aligned} N_4 \; : \|\alpha x\|_1 &= \|\alpha x\| + \|T(\alpha x)\| = |\alpha| \; \|x\| + \|\alpha \, T(x)\| \\ &= |\alpha| \; \|x\| + |\alpha| \; \|T(x)\| \\ &= |\alpha| \; \{\|x\| + \|T(x)\|\} = |\alpha| \; \|x\|_1 \end{aligned}$$

As such B_1 is a normed linear space.

Now $\|T(x)\| \le \|x\| + \|T(x)\| = \|x\|_1 \Rightarrow \|T(x)\| \le 1\|x\|_1$, which shows that T being regarded as a mapping from B_1 to B' is bounded and therefore continuous. Consequently in order to show that T is continous from B into B', it is sufficient to show that B and B_1 have the same topology i.e., they are homomorphic.

We, now establish that the normed linear space B_1 is a Banach space, by showing that it is complete.

If $< x_n >$ be a Cauchy sequence in B_1, then

$$\|x_n - x_m\|_1 \to 0 \text{ as } m,n \to \infty$$

$$\Rightarrow \quad \|x_n - x_m\| + \|T(x_n - x_m)\| \to 0 \text{ as } m,n \to \infty$$

$$\Rightarrow \quad \|x_n - x_m\| + \|T(x_n) - T(x_m)\| \to 0 \text{ as } m,n \to \infty$$

$$\Rightarrow \quad \|x_n - x_m\| \to 0 \text{ and } \|T(x_n) - T(x_m)\| \to 0 \text{ as } m,n \to \infty$$

$\Rightarrow \quad < x_n >$ is a Cauchy sequence in B and $< T(x_n) >$ is a Cauchy sequence in B'

$\Rightarrow \quad x_n \to x \in B$ and $T(x_n) \to y \in B'$ as B, B' are complete ...(1)

Now $< x_n, T(x_n) >$ being a Cauchy sequence in T_G (which is closed)

$$(x_n, t(x_n)) \to (x, y) \in T_G \qquad \text{by (1)}$$

$$\Rightarrow \quad y = T(x)$$

$$\therefore \quad \|x_n - x\|_1 = \|x_n - x\| + \|T(x_n - x)\|$$

$$= \|x_n - x\| + \|T(x_n) - T(x)\|$$

$$= \|x_n - x\| + \|T(x_n) - y\|$$

$$\to 0 \qquad \text{by (1)}$$

It follows that the sequence $< x_n >$ in B_1 converges to $x \in B_1$.

Hence B_1 is complete. [Note that B_1 and B are the same sets so that $x \in B \Rightarrow x \in B_1$].

Lastly to show that there is a homeomorphism between B and B_1, we consider an identity map

$$I : B_1 \to B : I(x) = x \; \forall \; x \in B_1.$$

Evidently I is one-one onto mapping and

$$\|I(x)\| = \|x\| \le \|x\| + \|T(x)\| = \|x\|_1, \; \forall \; x \in B_1.$$

i.e., $\|I(x)\| \leq 1.\|x\|_1 \Rightarrow I$ is bounded and continuous.

It is also one-one onto, therefore I is a homeomorphism from a Banach space B to Banach space B_1. Also T being continuous from B_1 to B' and hence it is continuous on its homomorphic image B i.e., $T : B \to B'$ is continuous.

3.5 Uniform Boundedness Theorem

The uniform boundedness theorem is of great importance. The principle of uniform boundedness asserts that if a sequence of bounded linear transformation $T_n \in B(B, N)$, $n \in N$ where B is a Banach space and N is a normed space, is pointwise bounded, then the sequence $\{T_n\}$ is uniformly bounded. Infact, it enables us to determine whether the norms of a given family of bounded linear transformations have a finite least upper bound.

Definition : A set $F \subset B(N\ ;\ N')$ of bounded linear transformations from a normed space N into a normed space N' is said to be :

(a) Pointwise bounded if for each $x \in X$, the set $\{T(x) : T \in F\}$ is a bounded set in N'.

(b) Uniformly bounded if F is bounded set in the normed linear space $B(N : N')$.

In definition (b), the boundedness of the set F means that there is a constant $M > 0$ such that $\|T\| \leq M$, $\forall\ T \in F$.

Let $x \in X$, then

$$\|T(x)\| \leq \|T\|\|x\| \leq M\|x\| \quad \forall\ T \in F.$$

This means that F is pointwise bounded. Thus if F is uniformly bounded set in $B(N\ ;\ N')$, then it is also pointwise bounded. However, the converse of this assertion may not hold good.

Theorem 5 (Uniform Boundedness Theorem)

Let B be a Banach space, N be a normed linear space and $\{T_G\}$ a non-empty set of bounded (and so continuous) linear transformations of B into N with the property that $\{T_i(x)\}$ is a bounded subset of N for each vector x in B, the $\{\|T_i\|\}$ is a bounded set of numbers i.e., $\{T_i\}$ is bounded as a subset of $B(B, N)$.

Proof : For each positive integer n, define

$$F_n = \{x : x \in B \text{ and } \|T_i(x)\| \leq n \text{ for all } i\} \qquad ...(1)$$

Then F_n is a closed subset of B as shown below

$$x \in F_n \Leftrightarrow \|T_i(x)\| \leq n \quad \text{for all } i$$

$$\Leftrightarrow T_i(x) \in S_n^c \quad \text{for all } i$$

where S_n^c denotes the closed sphere in N with centre 0 and radius n.

$\Leftrightarrow x \in T_i^{-1}\left[S_n^c\right]$ for all i

$\Leftrightarrow x \in \bigcap_i T_i^{-1}\left[S_n^c\right]$

so that $F_n = \bigcap_i T_i^{-1}\left[S_n^c\right]$, which is closed, being an intersection of closed sets.

[Note that since each T_i is continuous and S_n^c is closed in N, each $T_i^{-1}\left[S_n^c\right]$ is closed in B]

Further, $B = \bigcup_{n=1}^{\infty} F_n$ for if $B \neq \bigcup_{n=1}^{\infty} F_n$, then there eixsts some $x \in B$ such that $x \notin F_n$ for any n.

$\Rightarrow$ $\|T_i(x)\| > n$ for all n by (1)

$\Rightarrow$ The set $\{T_i(x)\}$ is not bounded, which contradicts the hypothesis. Hence we must have

$$B = \bigcup_{n=1}^{\infty} F_n,$$

so that the complete space B is the union of sequence of its subsets. Therefore by Baire's category theorem, there exists an integer n_0 such that $\overline{F_{n0}}$ has non-empty interior. Since F_n is closed, $\overline{F_{n0}} = F_{n0}$

and so F_{n0} must have non-empty interior, that is, there exists some $x_0 \in F_{n0}$, so that F_{n0} is a nbd of x_0. Since F_{n0} is closed, there exists a closed space

$$S = \left\{x \in B : \|x - x_0\| \leq r_0\right\} \subset F_{n0} \qquad ...(2)$$

Now if $\|y\| < 1$, then for arbitrary but fixed i

$$\|T_i(y)\| = \left\|T_i\left(\frac{z}{r_0}\right)\right\|, \text{ where } z = r_0 y$$

$$= \frac{1}{r_0}\|T_i(z)\| = \frac{1}{r_0}\|T_i(z + x_0 - x_0)\|$$

$$= \frac{1}{r_0}\|T_i(z + x_0) - T_i(x_0)\|$$

$$\leq \frac{1}{r_0}\left[\|T_i(z + x_0)\| + \|-T_i(x_0)\|\right]$$

$$= \frac{1}{r_0}\left[\|T_i(z + x_0)\| + \|T_i(x_0)\|\right]$$

$$\leq \frac{1}{r_0}(n_0+n_0)=\frac{2n_0}{r_0}, \quad z+x_0 \text{ and } x_0 \in F_{n0}.$$

[Note that $\|z+x_0-x_0\|=\|y\|=\|r_0\, y\|=r_0\|y\|\leq r_0 \;(\because \|y\|\leq 1)$ so that $z+x_0 \in S \subset F_{n0}$. Ofcourse $x_0 \in S \subset F_{n0}$]

Thus $\quad \|T_i(y)\| \leq \dfrac{2n_0}{r_0}$ if $\|y\| \leq 1$.

$$\therefore \quad \|T_i\| = \sup\left\{\|T_i(y)\| : \|y\| \leq 1\right\} \leq \frac{2n_0}{r_0}$$

If follows that $\{\|T_i\|\}$ is a bounded set of numbers.

Self-Learning Exercise

1. Write whether the following statements are true or false.

   (a) The open mapping theorem states conditions under which a bounded linear operator is an open mapping.

   (b) The proof of the open mapping theorem is based on Heine-Borel theorem.

   (c) A map of $f : X \to Y$ is known as an open mapping if $\forall$ open set V of X_1 then set $f^{-1}(V)$ is open in Y, X and Y being topological spaces.

   (d) A one-one continuous mapping of f of X onto Y is homeomorphism when f is open.

   (e) The closed space theorem states conditions under which a closed linear operator will be bounded.

   (f) The closed graph theorem is usually known by the name "The Banach Steinhaus theorem".

   (g) The uniform boundedness theorem gives condition sufficient for $\{\|T_n\|\}$ to be bounded, where the T_n' s are bounded linear transformations from a Banach space into normed space.

3.6 Summary

In this unit, we have studied how a new normed space can be formed by taking the product of given normed spaces. We have seen that uniform boundedness theorem gives conditions sufficient for $\{\|T_n\|\}$ to be bounded, where the $T_n' S$ are bounded linear transformation from a Banach space into a normed space. The open mapping theorem states conditions under which a bounded linear transformation is an open mapping. We have seen that the three theorems discussed in this unit require completeness. Indeed they characterize some of the most important properties of Banach spaces which normed spaces in general may not have.

3.7 Answers to Self-Learning Exercise

1.	(a)	True	(b)	False	(c)	False		
	(d)	True	(e)	True	(f)	False	(g)	True

3.8 Exercises

1. Let $(X_1, \|.\|_1), (X_2, \|.\|_2), \ldots, (X_n, \|.\|_n)$ be n-normed spaces. Then $X = X_1 \times X_2 \times \ldots \times X_n$ is a normed spaces under the norm

$$\|x\| = \|x_1\|_1 + \|x_2\|_2 + \ldots + \|x_n\|_n$$

for $\quad x = (x_1, x_2, \ldots, x_n) \in X$.

2. Let N be a Banach spaces, N' a normed space and $T_n \in B(N, N')$ such that $(T_n x)$ is Cauchy in N' for every $x \in N$. Show that $(\|T_n\|)$ is bounded.

3. If in addition N' in Problem 2 is complete, show that $T_n x \to T_x$, where $T \in B(N, N')$.

4. Let B and B' be Banach spaces and let T be one-one continuous linear transformation of B into B'. Then T is a homeomorphism. In particular, T^{-1} is automatically continous.

□□□

Unit - 4
Continuous Linear Functionals

Structure of the Unit

4.0 Objectives

In this unit, We introduce the concept of linear functional, prove the Hahn-Banach theorem on the existence of linear functionals and derive some of its many consequences.We define the dual space of a normed space. We discuss the natural imbedding and reflexivity in normed spaces.

4.1 Introduction

It is known that R (real space) and C (complex space) are the simplest of all normed spaces. In the present unit, we study the bounded (or continuous) linear transformations from arbitrary normed space into the normed spaces R or C. Such bounded linear transformations are called bounded linear functionals. All general theorems proved in the previous unit for bounded linear transformations are also valid for bounded linear functionals. The Hahn-Banach theorem is basically an extension theorem for linear functionals.

4.2 Continuous Linear Functionals

We know that R and C are the simplest of all normed linear spaces. If we limit ourselves with the continuous linear transtomations of a normed linear space N into R or C according as N is real or complex, then the set B (N, R) or B (N,C) of all bounded (or countinuous) linear transformations is denoted by N^* and known as the **conjugate space** or **Adjoint space** or **First dual space** of N and the elements of N^* are known as **Continuous linear functionals** or simply **functionals**.

Thus a functional on a normed linear space N is a Continuous linear transformation from N into R or C. If these functionals are added and multiplied by scalars pointwise under the norm of a functional defined by

$$\|f\| = Sup\left\{ |f(x)| : \|(x)\| \le 1 \right\}$$

$$= Sup\left\{ K : K \ge 0, |f(x)| \le K\|x\| \forall x \right\}$$

then N^* constitutes a Bannach space.

4.3 Hahn-Banach Theorem and its Consequences

The Haha-Banach theorem is basically an extension theorem for linear functionals. In this theorem, we consider a bounded linear functional f defined on a subspace M of a given normed space N and then we extend this from M to the entire space N in such a way that certain basic properties of f continue to hold good for the extended functional.

Theorem I (Hahn-Banach Theorem) :

If M be a linear subspace of a normed linear space N and f is a functional defined on M, then f can be extended to a functional f_0 defined on the whole space N s.t. $\|f_0\| = \|f\|$

Proof : We first prove the following lemma.

Lemma : If f be a functional defined on a linear subspace M of a normed linear space N, $x_0 \notin M$ and

$$M_0 = \left[M \cup \{x_0\}\right] = \{x + \alpha\, x_0 : x \in M \text{ and } \alpha \text{ is real}\}$$

is the linear subspace spanned by M and x_0, then f can be extended to a functional f_0 defined on M_0 s.t. $\|f_0\| = \|f\|$

Proof of the Lemma : We prove the lemma for real and complex scalars separately.

Case I : When N is real normed space, then $x_0 \notin M \Rightarrow$ each vector m in M_0 can be uniquely expressed as $m = x + \alpha x_0 \text{ with } x \in M$

Let us define f_0 on M_0, which is extension of *of* f s.t.

$$f_0\,(m) = f_0\,(x + \alpha x_0) = f_0(x) + \alpha\, f_0\,(x_0)$$

$$= f(x) + \alpha\, r_0 \qquad ...(1)$$

with the choice of real number $r_0 = f_0(x_0)$ ...(2)

and $f_0(x) = f(x)\ \forall\ x \in M$ (by definition of extension) ...(3)

we first claim that f_0 thus defined is linear on M_0.

Taking $B, \gamma \in R$ and $x, y \in M$ we have

$$f_0(\beta(x + \alpha x_0) + \gamma(y + \alpha x_0)] = f_0[(\beta x + \gamma y) + (\beta + \gamma)\ \alpha x_0]$$

$$= f_0\ (\beta x + \gamma\ y) + (\beta + \gamma)\alpha\ f_0\ (x_0)$$

$$= f\,(\beta x + \gamma y) + (\beta + r)\ \alpha r_0 \qquad \text{by (2) and (3)}$$

$$= \beta\ f\ (x) + \gamma f\,(y) + \beta\alpha r_0 + \gamma\alpha r_0$$

$$= \beta\left[f(x) + \alpha r_0\right] + \gamma\left[f(y) + \alpha r_0\right]$$

$$= \beta\ f_0\left(x + \alpha x_0\right) + \gamma\ f_0\left(y + \alpha x_0\right) \qquad \text{by (1)}$$

which shows that f_0 is linear on M_0.

Also f_0 is an extension of f, for if $x \in M$, then $x + 0.x_0$ so that

$$f_0(x) = f_0(x + 0.x_0) = f_0(x) + 0.f_0(x_0) = f(x) + 0.r_0 = f(x)$$

i.e., $f_0(x) = f(x) \forall x \in M \Rightarrow f_0$ is an extension of f over M.

Thus f_0 extends f linearly to M_0. We now prove that $\|f_0\| = \|f\|$

We have $\|f_0\| = \sup\{|f_0(x)| : \|x\| \le 1\}, x \in M_0$

$$\ge \sup\{|f_0(x)| : \|x\| \le 1\}, x \in M \ as\ M_0 \supset M$$

$$= \sup\{|f(x)| : \|x\| \le 1\}, x \in M \quad [\because f_0 = f \ on\ M]$$

$$= \|f\|$$

Thus $\|f_0\| \ge \|f\|$ ...(4)

So our problem now is to choose r_0 such that $\|f_0\| \le \|f\|$. For this purpose, we first observe that if x_1, x_2 are any two vectors in M, then

$$f(x_2) - f(x_1) = f(x_2 - x_1) \text{ by linearity of } f$$

$$\le |f(x_2 - x_1)|$$

or $$f(x_2) - f(x_1) \le \|f\| \|x_2 - x_1\|$$

$$= \|f\| \|(x_2 + x_0) - (x_1 + x_0)\|$$

$$\le \|f\| (\|x_2 + x_0\| + \|-(x_1 + x_0)\|)$$

$$= \|f\| \, \|x_2 + x_0\| + \|f\| \, \|x_1 + x_0\|$$

or $$-f(x_1) - \|f\| \, \|x_1 + x_0\| \le -f(x_2) + \|f\| \, \|x_2 + x_0\| \quad ...(5)$$

Which holds for arbitrary $x_1, x_2 \in M$ and can be written as

$$\sup_{y \in M} \{-f(y) - \|f\| \, \|y + x_0\|\} \le \inf_{y \in M} \{-f(y) - \|f\| \, \|y + x_0\|\}$$

since between any two real numbers there always exists a real numbe r_0 s.t.

$$\sup_{y \in M} \{-f(y) - \|f\| \, \|y + x_0\|\} \le r_0 \le \inf_{y \in M} \{-f(y) - \|f\| \, \|y + x_0\|\},$$

which follows that $\forall \ y \in M$

$$-f(y)-\|f\|\,\|y+x_0\|\leq r_0\leq -f(y)+\|f\|\,\|y+x_0\| \qquad ...(6)$$

Taking arbitrary $m = x + 2\,x_0$ in M_0 and setting $y=\dfrac{x}{\alpha}$, we find

$$-f\left(\frac{x}{\alpha}\right)-\|f\|\left\|\frac{x}{\alpha}+x_0\right\|\leq r_0\leq -f\left(\frac{x}{\alpha}\right)+\|f\|\left\|\frac{x}{\alpha}+x_0\right\| \qquad ...(7)$$

For $\alpha > 0$, the last two parts of inequality (7) yields

$$r_0\leq -\frac{1}{\alpha}f(x)+\frac{1}{\alpha}\|f\|\,\|x+\alpha\,x_0\|$$

$\Rightarrow$ $f(x)+\alpha\,r_0\leq \|f\|\,\|x+\alpha\,x_0\|$

$\Rightarrow$ $f_0(x)+\alpha\,f_0(x_0)\leq \|f\|\,\|x+\alpha\,x_0\|$ by (2) and (3)

$\Rightarrow$ $f_0(x+\alpha\,x_0)\leq \|f\|\,\|x+\alpha\,x_0\|$

$\Rightarrow$ $f_0(m)\leq \|f\|\,\|m\|$ ...(8)

or $\alpha < 0$ the first two parts of inequality (7) yields

$$r_0\geq -f\left(\frac{x}{\alpha}\right)-\|f\|\left\|\frac{x}{\alpha}+x_0\right\| \qquad f_0$$

$$=-\frac{1}{\alpha}f(x)-\|f\|\left|\frac{1}{\alpha}\right|\|x+\alpha\,x_0\|$$

$$=-\frac{1}{\alpha}f(x)+\frac{1}{\alpha}\|f\|\,\|x+\alpha\,x_0\| \quad \text{as} \quad \left|\frac{1}{\alpha}\right|=-\frac{1}{\alpha}$$

where $\alpha < 0$.

On multiplying both sides by α (a negative quantity), we get

$\alpha\,r_0\leq -f(x)+\|f\|\,\|x+\alpha\,x_0\|$ (sign of inequality being reversed)

$\Rightarrow$ $f(x)+\alpha\,r_0\leq \|f\|\,\|x+\alpha\,x_0\|$

$\Rightarrow$ $f_0(x)+\alpha\,f_0(x_0)\leq \|f\|\,\|x+\alpha\,x_0\|$ by (2) and (3)

$\Rightarrow$ $f_0(x+\alpha\,x_0)\leq \|f\|\,\|x+\alpha\,x_0\|$

$\Rightarrow$ $f_0(m)\leq \|f\|\,\|m\|$ ...(9)

(8) and (9) $\Rightarrow$ $f_0(m)\leq \|f\|\,\|m\|\ \forall\ m\in M_0,\ \alpha\neq 0$ ...(10)

clearly for $\alpha = 0$, $\|f_0\| = \|f\|$

Replacing m by $-m$ in (10), we get

$$f_0(-m) \le \|f\| \, \|-m\| \Rightarrow -f_0(m) \le \|f\| \, \|m\| \qquad ...(11)$$

(10) and (11) $\Rightarrow |f_0(m)| \le \|f_0\| \, \|m\|$ ...(12)

Since $\|f_0\| = \sup \{|f_0(m)| : \|m\| \le 1\}$, $m \in M_0$, f_0 being linear functional on M_0. It follows from (12), that

$$\|f_0\| \le \|f\| \qquad ...(13)$$

(4) and (13) $\Rightarrow \|f_0\| = \|f\|$ ...(14)

Case II : When N is a complex normed linear space, over C, then f is complex valued linear functional on M as subspace of N. Suppose g and h are real and imaginary parts of f, so that

$$f(x) = g(x) + i\,h(x) \ \forall\ x \in M \qquad ...(15)$$

Now a complex linear space can be regarded as a real linear space by restricting the scalars to real numbers and g, h are real valued functionals on the real space M. We have for $x, y \in M$ and $a \in R$,

$$f(x+y) = f(x) + f(y)$$

$$\Rightarrow g(x+y) + i\,h(x+y) = g(x) + i\,h(x) + g(y) + i\,h(y)$$

$$\Rightarrow g(x+y) = g(x) + g(y) \quad \text{and} \quad h(x+y) = h(x) + h(y)$$

and $\quad f(\alpha x) = \alpha f(x) \Rightarrow g(\alpha x) + i\,h(\alpha x) = \alpha[g(x) + i\,h(x)]$

$$\Rightarrow g(\alpha x) = \alpha\, g(x) \ \text{ and } \ h(\alpha x) = \alpha\, h(x)$$

Which follows that g and h are linear on M. Also

$$|g(x)| \le |f(x)| \ \text{ as } \ w = u + i\,v \Rightarrow |u| \le |w|$$

$$\le \|f\| \, \|x\|.$$

Thus if f is bounded, then so are g and h. Consequently, g and h are real linear functionals on the real space M. Again $\forall\ x \in M$

$$g(ix) + i\,h(ix) = f(i\,x) = i(f(x)) = i[g(x) + i\,h(x)] = i\,g(x) - h(x)$$

giving $\quad g(i\,x) = -h(x)$ and $h(i\,x) = g(x)$

$$\therefore \quad f(x) = g(x) + i\,h(x) = g(x) - i\,g(i\,x) = h(i\,x) + i\,h(x) \qquad ...(16)$$

Taking $f(x) = g(x) - i\,g(i\,x)$ and g being a real valued functional on real space M, we have by Case I, that g can be extended to a real valued functional g_0 on the real space M_0 s.t.

$$\|g_0\| = \|g\| \qquad \text{...(17)}$$

If we define f_0 s.t. $f_0(x) = g_0(x) - i\,g_0(ix) \;\; \forall\, x \in M_0$, then it can be observed that f_0 is linear on the complex space M_0 such that

$f_0 = f$ on M, since

$$f_0(x+y) = g_0(x+y) - i\,g_0(i\,x + i\,y), \;\; x, y \in M_0$$

$$= g_0(x) + g_0(y) - i\,g_0(i\,x) - i\,g_0(i\,y)$$

$$= g_0(x) - i\,g_0(i\,x) + g_0(y) - i\,g_0(i\,y)$$

$$= f_0(x) + f_0(y)$$

and if $\alpha, \beta \in R$, then

$$f_0[(\alpha + i\beta)x] = g_0(\alpha\,x + i\,\beta\,x) - i\,g_0(-\beta\,x + i\,\alpha\,x)$$

$$= \alpha\,g_0(x) + \beta\,g_0(i\,x) - i(-\beta)\,g_0(x) - i\,\alpha\,g_0(i\,x)$$

$$= (\alpha + i\,\beta)[g_0(x) - i\,g_0(i\,x)]$$

$$= (\alpha + i\,\beta)\,f_0(x)$$

Thus f_0 is linear on M_0. Also $g_0 = g$ on M implies $f_0 = f$ on M. What remains to prove is that $\|f_0\| = \|f\|$.

Let $x \in M_0$ be arbitrary and write $f_0(x) = r\,e^{i\theta}$, where $r \geq 0$ and θ real. Then

$$|f_0(x)| = r = e^{-i\theta} f_0(x) = f_0(e^{-i\theta}x) = g_0(e^{-i\theta}x), \; r \text{ being real}$$

$$\leq |g_0(e^{-i\theta}x)| \leq \|g_0\|\,\|e^{-i\theta}x\|$$

$$= \|g_0\|\,\|e^{-i\theta}\|\,\|x\| = \|g_0\|\,\|x\| = \|g\|\,\|x\| \qquad \text{by (17)} \quad (\because |e^{-i\theta}| = 1)$$

$$\leq \|f\|\,\|x\|$$

This shows that f_0 is bounded (hence a functional on M_0) and that $\|f_0\| \leq \|f\|$. Also as in Case I, it is obvious that $\|f_0\| \leq \|f\|$.

Therefore $\|f_0\| = \|f\|$

Theory of the Main Theorem : In view of lemma, for any $x \in N$, but $x \notin M$, we can have an extension of f on $M \cup \{x\}$ s.t. $\|f\|$ is preserved for extension. If we consider the set of all positive extensions of f on all the subspaces $M \cup$ {element of N not in M} of N, containing M, then this set of extensions of f say G can be partially ordered as under.

Taking $g_1, g_2 \in G$ and relation $\leq$ s.t. $g_1 \leq g_2 \Rightarrow$ domain of g_1 is contained in the domain of g_2 and $g_1(x) = g_2(x)\ \forall\ x \in$ dom (g_1). We claim that $(G, \leq)$ is partially ordered, since it is reflexive, anti-symmetric and transitive.

Reflexivity : $g_1 \leq g_1\ \forall\ g_1 \in G$.

Antisymmetry : $g_1 \leq g_2$ and $g_2 \leq g_1 \Rightarrow$ dom (g_1) is contained in dom (g_2) and dom (g_2) $\subset$ dom (g_1)

$\Rightarrow \quad dom\,(g_1) = dom(g_2)$

$\Rightarrow \quad g_1(x) = g_2(x)\ \forall\ x \in dom(g_1)$ and

$g_2(x) = g_1(x)\ \forall\ x \in dom(g_2)$

$\Rightarrow \quad g_1 = g_2$, domains being same and functional values equal for all points of the domain.

Transitivity : $g_1 \leq g_2$ and $g_2 \leq g_3 \Rightarrow$ dom $(g_1) \subset$ dom (g_2) with $g_1(x) = g_2(x)\ \forall\ x \in dom(g_1)$;

dom $(g_2) \subset$ dom (g_3) with $g_2(x) = g_3(x)\ \forall\ x \in$ dom (g_2)

$\Rightarrow \quad$ dom $(g_1) \subset$ dom (g_3) with $g_1(x) = g_3(x)\ \forall\ x \in$ dom (g_1)

$\Rightarrow \quad g_1 \leq g_3$

Hence the set G is partially ordered.

Also we observe that the union of any chain of extensions is an extension and therefore there is an upper bound for the chain. Thus every chain in G has an upper bound. As such by Zorn's lemma, $\exists$ a maximal extension $f_0 \in G$, otherwise $\exists$ an $x \in N$ and $x \notin M$ s.t. f_0 can be extended to the domain of $f_0 \cup \{x\}$ i.e., $M \cup \{x\}$ by the lemma. But this violates the maximality of f_0. Hence the domain of f_0 must be the whole space N s.t. $\|f_0\| = \|f\|$.

We now derive some important consequences of theorem 1.

Theorem 2 : If N be a normed linear space and x_0 is a non zero vector in N, then $\exists$ a continuous linear functional F defined on the conjugate space N^* s.t.

$$F(x_0) = \|x_0\| \text{ and } \|F\| = 1.$$

Proof : Let $M = \{\alpha x_0\}$ be the linear subspace of N spanned by x_0. Define f_0 on M by $f_0(\alpha x_0) = \alpha \|x_0\|$. We claim that f_0 is a functional on M such that $\|f_0\| = 1$.

f_0 is linear :

Let $y_1, y_2 \in M$ so that $y_1 = \alpha x_0$, $y_2 = \beta x_0$ for some scalars α and β. If γ, δ are any scalars, then

$$f_0(\gamma y_1 + \delta y_2) = f_0(\gamma \alpha x_0 + \delta \beta x_0) = f_0\left[(\gamma \alpha + \delta \beta) x_0\right]$$

$$= (\gamma \alpha + \delta \beta) \|x_0\| \text{ by def. of } f_0$$

$$= \gamma \alpha \|x_0\| + \delta \beta \|x_0\|$$

$$= \gamma f_0(\alpha x_0) + \delta f_0(\beta x_0)$$

$$= \gamma f_0(y_1) + \delta f_0(y_2)$$

f_0 **is bounded.**

Let $y = \alpha x_0 \in M$ so that $\|y\| = \|\alpha x_0\| = |\alpha| \|x_0\|$. Now

$$|f_0(y)| = |f_0(\alpha x_0)|$$

$$= |\alpha| \|x_0\| = \|y\|$$

Hence f_0 is bounded. It follows that f_0 is a functional on M.

Further $\|f_0\| = \sup \{|f_0(y)| : y \in M, \|y\| \le 1\}$

$$= \sup \{\|y\| : \|y\| \le 1\} = 1$$

Now choosing $\alpha = 1$, $f_0(\alpha x_0) = \alpha \|x_0\| \Rightarrow f_0(x_0) = \|x_0\|$.

Hence by Hahn-Banach theorem f_0 can be extended to a norm preserving functional $F \in N^*$ so that

$$F(x_0) = f_0(x_0) = \|x_0\| \quad \text{and} \quad \|F\| = \|f_0\| = 1$$

Note : As a particular case, if $x \ne y$, $x, y \in N$, so by the above theorem, there exists an $f \in N^*$ such that

$$f(x-y) = \|x-y\| \ne 0 \Rightarrow f(x) - f(y) \ne 0 \Rightarrow f(x) \ne f(y)$$

This shows that N^* separates vectors in N.

Theorem 3 : Let N be a real normed linear space and suppose $f(x) = 0$ for all $f \in N^*$. Show that $x = 0$.

Proof : Suppose $x \ne 0$. Then by Theorem 2, there exists $f \in N^*$ such that $f(x) = \|x\| > 0$, which contradicts the hypothesis that $f(x) = 0$ for all $f \in N^*$. Hence we must have $x = 0$.

Theorem 4 : If M be a closed linear subspace of a normed linear space N and x_0 is a vector not in M, then $\exists$ a functional F in conjugate space N^* s.t. $F(M)=\{0\}$ and $F(x_0)\neq 0$.

Proof : Consider the natural mapping

$$T: N \to N/M \text{ s.t. } T(x)=x+M \ \forall \ x \in N,$$

then $\|T(x)\|=\|x+M\|=\inf\{\|x+m\| : m \in M\}$ by def.

$\leq \|x+m\| \ \forall \ m \in M$ by def. of infimum.

But M being subspace, $0 \in M$, so that above result still holds for $m=0\in M$ i.e., $\|T(x)\|\leq\|x\| \ \forall \ x \in M$, which follows that T is bounded and hence continuous.

Now $\quad T(m)=m+M=M=0$ of N/M ...(1)

and $\quad x_0 \notin M \Rightarrow T(x_0)=x_0+M\neq M$ i.e., 0 of N/M ...(2)

As such $T(x_0)$ i.e., $x_0+M\neq 0$ is a non zero vector (coset) in N/M. Therefore by Theorem 2, $\exists$ a functional f in $(N/M)^*$ s.t.

$$f(x_0+M)=\|x_0+M\|\neq 0 \qquad ...(3)$$

If we define F on N as

$$F(x)=f[T(x)],$$

then F is a linear transformation being the composition of F and T.

Also $\quad F(m)=f[T(m)]=f(0)=0 \ \forall \ m \in M \qquad$ by (1)

$\therefore \quad F(M)=0$ and $F(x_0)=f[T(x_0)]=f(x_0+M)\neq 0$

by (2) and (3).

Theorem 5 : If M be a closed linear subspace of a normed linear space N and x_0 be a vector in N, but not in M with the property that the distance from x_0 to M i.e., $d(x_0, M)=d>0$, then $\exists$ a bounded linear functional $F\in N^*$ s.t. $\|F\|=1$,

$$F(x_0)=d \text{ and } F(x)=0 \ \forall \ x \in M \text{ i.e., } F(M)=\{0\}.$$

Proof : We have by definition

$$d=\inf\{\|x_0-x\| : x \in M\},\ d>0 \qquad ...(1)$$

Now consider the subspace

$$M_0=\{x+\alpha x_0 : x \in M,\ \alpha \text{ real}\}$$

spanned by M and x_0. Since $x_0 \notin M$, the representation of each vector y in M_0 in the form $y = x + \alpha x_0$ is unique.

Define a mapping f_0 on M_0 by $f_0(y) = \alpha d$ ...(2)

Where $y = x + \alpha x_0$ and d as in the hypothesis. Because of the uniqueness of y, the mapping f_0 is well defined. It is clear that f_0 is linear on M_0.

Now $f_0(x_0) = f_0(0 + 1.x_0) = 1d = d$ by (2)

and for any $m \in M$, $f_0(m) = f_0(m + 0.x_0) = 0.d = 0 \Rightarrow f_0(M) = \{0\}$.

Now, we claim that $\|f_0\| = 1$, since

$$\|f_0\| = \sup\left\{\frac{|f_0(y)|}{\|y\|} : y \neq 0\right\}, \quad y \in M_0$$

$$= \sup\left\{\frac{|f_0(x + \alpha x_0)|}{\|x + \alpha x_0\|} : x \neq 0,\ \alpha \neq 0\right\}, \quad x \in M,\ \alpha \in R$$

$$= \sup\left\{\frac{|\alpha d|}{\|x + \alpha x_0\|} : \alpha \neq 0\right\}, \quad \alpha \in R,\ x \in M \qquad \text{by (2)}$$

$$= \sup\left\{\frac{d}{\left\|x_0 + \dfrac{x}{\alpha}\right\|} : \alpha \neq 0\right\}, \quad \alpha \in R,\ x \in M, \text{ as } d > 0 \text{ and } |\alpha d| = d|\alpha|$$

$$= d \sup\left\{\frac{1}{\|x_0 - z\|} : z = -\frac{x}{\alpha} \in M\right\}$$

$$= d\left[\inf\{\|x_0 - z\| : z \in M\}\right]^{-1} = d.\frac{1}{d} = 1 \qquad \text{by (1)}$$

so f_0 is a linear functional on M_0 such that

$f_0(M) = \{0\}$, $f_0(x_0) = d$ and $\|f_0\| = 1$ ...(3)

Hence by the Hahn-Banach theorem, there exists a functional F on the whole space N such that

$F(y) = f_0(y)\ \forall\ y \in M_0$ and $\|F\| = \|f_0\|$.

It follows from (3) that

$F(M) = \{0\}$; $F(x_0) = d$ and $\|f\| = 1$ as desired.

4.4 Natural Imbedding and Reflexivity in Normed Spaces

If N be a normed linear space, then the set of all bounded linear functionals defined on N form a Banach space, denoted by N^* and is known as the **Dual space** or the **conjugate space** or the **adjoint space** or the **first dual space** of the normed space N. The space of bounded linear functionals on N^* is known as the **second dual space** of N and denoted by N^{**}.

Taking N^* and N^{**} as the first and second conjugate spaces of a normed linear space N, so that each vector x in N gives rise to a functional f in N^* and a functional F_x in N^{**}, we defined F_x as

$$F_x(f) = f(x) \; \forall \; f \in N^*.$$

The mapping $J : x \to F_x$ of N into N^{**}, where $F_x(f) = f(x) \; \forall \; f \in N^*$, is called the **natural embedding.**

If the natural imbedding $J : x \to F_x$ of N into N^{**} is an onto mapping, then we call the normed space N as **Reflexive.**

Here F_x is also known as the functional on N^* induced by the vector x of N and we generally say it **induced functional.**

Theorem 6 : Let N be an arbitrary normed linear space. Then for each vector x in N induces a functional F_x on N^{**} defined by $F_x(f) = f(x) \; \forall \; f \in N^*$ such that $\|F_x\| = \|x\|$.

Further the mapping $J : N \to N^{**}$ defined as $J(x) = F_x \; \forall \; x \in N$ is an isometric isomophism of N into N^{**}.

Proof : We first claim that F_x is linear, since $f, g \in N^*$ and scalars α, β, we have

$$F_x(\alpha f + \beta g) = (\alpha f + \beta g)(x) \qquad \text{by def. of } F_x$$

$$= (\alpha f)(x) + (\beta g)(x)$$

$$= \alpha f(x) + \beta g(x)$$

$$= \alpha F_x(f) + \beta F_x(g) \qquad ...(1)$$

Again, we claim that F_x is bounded, since for all $f \in N^*$, we have

$$\|F_x\| = \sup \{|F_x(f)| : \|f\| \le 1\}$$

$$= \sup \{|f(x)| : \|f\| \le 1\}$$

$$= \sup \{\|f\| \, \|x\| : \|f\| \le 1\}$$

$$\le \|x\| \qquad ...(2)$$

Thus $\|F_x\| \le 1 \, . \, \|x\|$. It follows that F_x is bounded i.e., continuous. Hence F_x is a functional on N^{**}.

For a non zero vector x in N, $\exists$ a functional $f_0 \in N^*$ s.t.

$$f_0(x) = \|x\| \text{ and } \|f_0\| = 1 \qquad \text{(by Theorem 2)} \qquad ...(3)$$

For such a functional f_0, we have

$$F_x(f_0) = f_0(x) \qquad \text{by def. of } F_x$$

i.e., $F_x(f_0) = \|x\|$, where $\|f_0\| = 1$ (by (3))

$$\Rightarrow \quad \|x\| = \big|\|x\|\big| = |F_x(f_0)| \le \|F_x\|\,\|f_0\| = \|F_x\| \qquad (\because \|f_0\| = 1)$$

$$\Rightarrow \quad \|x\| \le \|F_x\| \qquad ...(4)$$

Hence (2) and (4) $\Rightarrow \|F_x\| = \|x\|$ ...(5)

When x is a zero vector, then from (1), we have

$$\|F_x\| \le \|x\| \Rightarrow \|F_x\| = \|F_0\| \le \|0\| = 0$$

and $\|F_0\| \ge 0$ as $\|F_0\| \ge 0$ always.

Hence $\|F_0\| = \|0\|$

Thus we have shown that $\|F_x\| = \|x\| \ \forall \ x \in N$.

Now we prove that J is an isometric isomophism i.e. J is a one-one linear transformation as well as an isometry.

J is linear, since for any $x, y \in N$ and scalars α, β, we have

$$F_{\alpha x + \beta y}(f) = f(\alpha x + \beta y)$$

$$= \alpha f(x) + \beta f(y) \qquad (\because f \text{ is a linear transformation } \forall \ f \in N^*)$$

$$= \alpha F_x(f) + \beta F_y(f)$$

$$= (\alpha F_x)(f) + (\beta F_y)(f)$$

$$= (\alpha F_x + \beta F_y)(f)$$

$$\Rightarrow \quad F_{\alpha x + \beta y} = \alpha F_x + \beta F_y$$

It follows that

$$J(\alpha x + \beta y) = F_{\alpha x + \beta y} = \alpha F_x + \beta F_y = \alpha J(x) + \beta J(y)$$

$$\Rightarrow \ J \text{ is linear.}$$

Lastly we claim that J is an isometry, since by (5),

$$\|J(x) - J(y)\| = \|F_x - F_y\| = \|F_{x-y}\| = \|x - y\|$$

Thus J preserves norm, so it is an isometry. Also

$$J(x) = J(y) \Rightarrow J(x) - J(y) = 0$$

$\Rightarrow \quad \|J(x)-J(y)\|=0$

$\Rightarrow \quad \|x-y\|=0 \qquad (\because J$ pressure norm)

$\Rightarrow \quad x-y=0$

$\Rightarrow \quad x=y$

i.e. J is one-one.

Hence $J: N \to N^{**}$ is an isometric isomorphism.

Self-Learning Exercise

1. Write whether the following statements are true or false :

   (a) The norm $\|.\|: N \to R$ on a normed space $(X, \|.\|)$ is functional on N which is not linear.

   (b) If f is a bounded linear functional on a complex normed space. Then $\bar{f}$ is linear.

   (c) The Hahn-Banach theorem is an extension theorem for linear functional.

   (d) If N be a real normed linear space and $f(x)=0 \ \forall \ f \in N^*$ (conjugate space). Then $x \neq 0$.

2. If f is a linear functional on an n-dimensional vector space X. What dimension can the null space $N(f)$ have?

4.5 Summary

In this unit, we have seen that Hahn-Banach theorem is an extension theorem for linear functionals on linear spaces. We defined the dual space of a normed space to be the set of all bounded linear functionals on the space. We have seen that in some cases, the second dual space of a normed space, under a specific mapping called natural embedding is isometrically isomorphic to the origional space.

4.6 Answers to Self-Learning Exercise

1. (a) True (b) False (c) True (d) False

2. n or $n-1$

4.7 Exercises

1. If M be aclosed linear subspace of a normed linear space N, x_0 be a point in N but not in M and d be the distance from x_0 to M.

   Then show that $\exists$ a functional F in N (whole space) s.t.

$$F(M)=\{0\},\ F(x_0)=1 \text{ and } \|F\|=\frac{1}{d}.$$

2. State and prove Hahn-Banach theorem.

3. Show that dual of R^n in R^n.

4. Prove that if a normed space N is reflexive, it is complete.

5. If a normed space N is reflexive, show that N^* is reflexive.

□□□

Unit - 5
Hilbert Space and Its Basic Properties

Structure of the Unit

5.1 Objectives

The aim of this unit is to study **Inner product spaces** and Hilbert spaces and its basic properties. Here we shall prove Schwarz inequality, paralleogram law and polarisation identity in Hilbert spaces.

5.2 Introduction

We know that the norm on a vector space is the generalisation of the distance from the origin in an Euclidean space. The Euclidean space is not only provided with the distance amenable to the definition of norm, but also it is provided with the geometric concepts such as dot product. Using the dot product one can find the magnitude of vector and express the condition of orthogonality. These concepts can be illustrated very well by considering the Euclidean space of three dimensions. Such ideas like dot product and condition of orthogonality are totally missing in a normed linear space. The extension of these notions to any arbitrary infinite dimensional vector spaces leads to the definition of inner product on a vector space in such a way that the inner product gives rise to a norm. Since an inner product is used to define a norm on a vector space, the inner product spaces are special normed linear spaces. A complete inner product space is called a Hilbert space. Thus every Hilbert space is a Banach space but converse is not necessarity true. In the next four units we shall study in detail the basic theory of Hilbert spaces.

5.3 Inner Product Spaces

5.3.1 Definition :

Let X be a linear space over the complex field C. An **inner product** on X is a function $(\ ): X \times X \to C$ which satisfies the following conditions :

I. $(\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z) \ \forall \ x, y, z \in X$ and $\alpha, \beta \in C$

(Linearity in the first varible)

II. $\overline{(x,y)} = (y,x)$ (Conjugate symmetry)

where the bar denotes the complex conjugate.

III. $(x,x) \geq 0$, $(x,x) = 0$ iff $x = 0$ (Positive definiteness)

A complex inner product space X is a linear space over C with an inner product defined on it. We can also define inner product by replacing C by R in the above definition. In that case, we get a **real inner product space**. Since the theory of operators on a complex inner product space alone gives non-trivial results in some important situations.

We shall consider only complex inner product spaces.

5.3.2 Examples

Example 1 : The space l_2^n consisting of all n tuples $x = (x_1, \ldots, x_n)$ of complex numbers and the inner product on l_2^n is defined as $(x,y) = \sum_{i=1}^{n} x_i \overline{y}_i$, where $y = (y_1, \ldots, y_n)$ is an inner product space.

Solution : Let $\alpha, \beta \in C$ and $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ and $z = (z_1, \ldots, z_n)$ belong l_2^n. Then

I. $$(\alpha x + \beta y, z) = \sum_{i=1}^{n} (\alpha x_i + \beta y_i)\overline{z}_i$$

$$= \alpha \sum_{i=1}^{n} x_i \overline{z}_1 + \beta \sum_{i=1}^{n} y_i \overline{z}_1$$

$$= \alpha(x,z) + \beta(y,z)$$

II. $$\overline{(x,y)} = \overline{\left(\sum_{i=1}^{n} x_i \overline{y}_i \right)}$$

$$= \overline{(x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n)}$$

$$= y_1\overline{x}_1 + y_2\overline{x}_2 + \ldots + y_n\overline{x}_n$$

$$= \sum_{i=1}^{n} y_i \bar{x}_i = (y, x)$$

III. $$(x, x) = \sum_{i=1}^{n} x_i \bar{x}_i = \sum_{i=1}^{n} |x_i|^2$$

Hence $(x, x) \geq 0$ and $(x, x) = 0$ iff $x_i = 0$ for each i i.e., $(x, x) = 0$ iff $x = 0$.

Thus l_2^n is an linear product space.

Example 2 : The linear space l_2 consisting of all complex sequences $x = (x_n)$ such that $\sum_{n=1}^{\infty} |x_n|^2$ is converg net is an inner product space.

Solution : Define the inner product on l_2 as

$$(x, y) = \sum_{n=1}^{\infty} x_n \bar{y}_n \quad \forall \; x = (x_n) \text{ and } y = (y_n) \in l_2 \qquad ...(1)$$

First we show that the inner product

(i) is well defined. For this we have to show that

(ii) is a convergent series having the sum as a complex number.

By Cauchy's inequality, we have

$$\sum_{i=1}^{n} |x_i \, \bar{y}_i| \leq \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2 \right)^{1/2}$$

$$\leq \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2} \left(\sum_{n=1}^{\infty} |y_n|^2 \right)^{1/2}$$

Since $\sum_{n=1}^{\infty} |x_n|^2$ and $\sum_{n=1}^{\infty} |y_n|^2$ are convergent, the sequence of partial sums of the series $\sum_{n=1}^{\infty} |x_n \cdot \bar{y}_n|$ is a monotonic increasing sequence bounded above. Therefore, the series $\sum_{n=1}^{\infty} |x_n \, \bar{y}_n|$ is convergent.

Hence $\sum_{n=1}^{\infty} x_n \, \bar{y}_n$ is absolutely convergent having its sum as a complex number. Therefore (1) is convergent so that the linear product (1) is well-defined.

The three axious for inner product space can be verified as in example 1.

Hence l_2 is an inner product space.

5.3.3 Basic Properties

The basic properties of inner product space are contained in the following theorem :

Theorem 1 : Let X be a complex inner product space, then

(i) $(\alpha x-\beta y,z)=\alpha(x,z)-\beta(y,z)$

(ii) $(x,\beta y+\gamma z)=\overline{\beta}(x,y)+\overline{\gamma}(x,z)$

(iii) $(x,\beta y-\gamma z)=\overline{\beta}(x,y)-\overline{\gamma}(x,z)$

(iv) $(x,0)=0$ and $(0,x)=0 \;\forall\; x\in X$

where α,β and $\gamma\in C$.

Proof :

(i) $(\alpha x-\beta y,\, z)=(\alpha x+(-\beta)y,z)$

$$=\alpha(x,z)+(-\beta)(y,z)$$

$$=\alpha(x,z)-\beta(y,z)$$

(ii) $(x,\beta y+\gamma z)=\overline{(\beta y+\gamma z,x)}$

$$=\overline{\beta(y,x)+\gamma(z,x)}$$

$$=\overline{\beta}\,\overline{(y,x)}+\overline{\gamma}\,\overline{(z,x)}$$

$$=\overline{\beta}\,(x,y)+\overline{\gamma}\,(x,z)$$

(ii) shows that an inner product is conjugate linear in the second variable.

(iii) $(x,\beta y-\gamma z)=(x,\beta y+(-\gamma)z)$

$$=\overline{\beta}(x,y)+\overline{(-\gamma)}(x,z) \qquad \text{(using (ii))}$$

$$=\overline{\beta}(x,y)-\overline{\gamma}(x,z)$$

(iv) $(0,x)=(0\theta,x)=0(\theta,x)=0$

where θ is zero element of X

and $(x,0)=\overline{(0,x)}=0$.

With the help of the inner product, on a linear space X we can define a norm on X. Define $\|x\|=[(x,x)]^{1/2}\;\forall\; x\in X$. To prove that is a norm, we require the following

Theorem 2 : If x and y are any two vectors in an inner product space X, then

$$|(x,y)| \le \|x\|\,\|y\| \qquad \ldots(2)$$

The inequality by (2) is also known as Schwarz inequality.

Proof : If $y=0$, then $\|y\|=0$ and $|(x,y)|=0$ so that both sides (2) vanish and the inequality is true. Therefore let us assume that $y \ne 0$ and $\lambda \in C$. Then

$$0 \le \|x-\lambda y\|^2 = (x-\lambda y,\ x-\lambda y)$$

Since $(x-\lambda y,\ x-\lambda y) = (x,x)-(x,\lambda y)-(\lambda y,x)+(\lambda y,\lambda y)$

$$= (x,x)-\bar{\lambda}(x,y)-\lambda(y,x)+\lambda\bar{\lambda}(y,y)$$

$$= \|x\|^2-\lambda(y,x)-\bar{\lambda}(x,y)+|\lambda|^2\|y\|^2$$

Therefore $\|x\|^2-\lambda(y,x)-\bar{\lambda}(x,y)+|\lambda|^2\|y\|^2 \ge 0 \qquad \ldots(3)$

Now $y \ne 0$, $\|y\| \ne 0$. So choosing

$$\lambda = \frac{(x,y)}{\|y\|^2} \quad \text{and taking } (y,x)=\overline{(x,y)}$$

From (3) we have

$$\|x\|^2 - \frac{(x,y)\,\overline{(x,y)}}{\|y\|^2} - \frac{\overline{(x,y)}(x,y)}{\|y\|^2} + \frac{|(x,y)|^2}{\|y\|^4}\|y\|^2 \ge 0$$

or $$\|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} - \frac{|(x,y)|^2}{\|y\|^2} + \frac{|(x,y)|^2}{\|y\|^2} \ge 0$$

or $$\|x\|^2 - \frac{|(x,y)|^2}{\|y\|^2} \ge 0$$

or $$|(x,y)| \le \|x\|\|y\|$$

Remark : In Schwaz inequality, equality holds good iff x and y are linearly dependent.

Theorem 3 : If X is an inner product space, then $\|x\| = (x,x)^{1/2}$ is a norm on X.

Proof : (i) we have $\|x\| = (x,x)^{1/2} \Rightarrow \|x\|^2 = (x,x)$

Now $\|x\| \ge 0$ and $\|x\|=0$ iff $(x,x)=0$ i.e. $x=0$

(ii) Let $x, y \in X$, then

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

$$= (x,x)+(x,y)+(y,x)+(y,y)$$

$$= (x,x)+(x,y)+\overline{(x,y)}+(y,y)$$

$$= \|x\|^2 + 2\operatorname{Re}(x,y)+\|y\|^2$$

$$\le \|x\|^2 + 2|(x,y)|+\|y\|^2 \qquad (\because \operatorname{Re}(z)\le |z|\ \forall z\in C)$$

$$\le \|x\|^2 + 2\|x\|\|y\|+\|y\|^2 \qquad \text{(using Schwarz inequality)}$$

Thus $\|x+y\|^2 \le (\|x\|+\|y\|)^2$

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

(iii) For any scalar $\alpha \in C$ and $x \in X$, we have

$$\|\alpha x\|^2 = (\alpha x,\ \alpha x) = \alpha\overline{\alpha}(x,x)$$

$$= |\alpha|^2\|x\|^2$$

$\therefore$ $\|\alpha x\| = |\alpha|\,\|x\|$

Hence $\|.\|$ satisfies all condition of the norm.

Since we are able to define a norm on X with the help of the inner product, the inner product space X becomes a **normed linear space.**

5.4 Hilbert Space

5.4.1 Definition : A complete inner product space is called a Hilbert space

or

Let H be a complex Banach space with a linear product defined on it. Then H said to be a **Hilbert space** if a complex number (x,y) called the inner product of x and y satisfy the following properties :

(i) $(x,x) = \|x\|^2$

(ii) $\overline{(x,y)} = (y,x)$

(H_3) $(\alpha x+\beta y,z) = \alpha(x,z)+\beta(y,z)$

$\forall\ x,y,z \in H$ and $\alpha,\beta \in C$

Remark 1 : Examples (1) and (2) of § 5.3 are complete inner product spaces, since l_2^n and l_2 are Banach spaces with norm defined as

$$\|x\| = \left\{ \sum_{i=1}^{n} |x_i|^2 \right\}^{1/2}$$

Example 1 is a finite dimensional space beacuse underlying vector space l_2^n is finite while Example 2 is an infinite dimensional space.

Remark 2 : Note that the set of all sequences $x = \{x_n\}$ such that x_n is ultimately zero is an incomplete inner product space, the inner product being induced by l_2, since we can find a sequence

$(x_n) = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, 0, 0, \ldots\right)$ converges in l_2 but its limit has no zero terms.

Hence we conclude that every Hilbert space is an inner product space but converse is not necessarity true.

5.4.2 Basic Properties

Theorem 3 : The inner product in a Hilbert space is jointly continuous i.e. if $x_n \to x$ and $y_n \to y$, then $(x_n, y_n) \to (x, y)$ as $n \to \infty$.

Proof : We have

$$|(x_n, y_n) - (x, y)| = |(x_n, y_n) - (x_n, y) + (x_n, y) - (x, y)|$$

$$= |(x_n, y_n - y) + (x_n - x, y)|$$

$$\leq |(x_n, y_n - y)| + |(x_n - x, y)| \qquad \text{...(4)}$$

By Schwarz inequality, we have

$$|(x_n, y_n - y)| \leq \|x_n\| \, \|y_n - y\| \qquad \text{...(5)}$$

and $$|(x_n - x, y)| \leq \|x_n - x\| \, \|y\| \qquad \text{...(6)}$$

Using (5) and (6) in (4) we get

$$|(x_n, y_n) - (x, y)| \leq \|x_n\| \, \|y_n - y\| + \|x_n - x\| \, \|y\| \qquad \text{...(7)}$$

Since $x_n \to x$ and $y_n \to y$, therefore

$$\|x_n - x\| \to 0 \quad \text{and} \quad \|y_n - y\| \to 0$$

Further since (x_n) is convergent sequence therefore it is bounded so that $\|x_n\| \leq M \;\; \forall \, n$

Using above in (7), we find that

$$(x_n, y_n) \to (x, y) \quad \text{as} \quad n \to \infty$$

Hence inner product in a Hilbert space is continuous.

Theorem 4 (Parallelogram Law) :

If x and y are any two vectors in a Hilbert space H, then $\|(x+y)\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$.

Proof : For any $x, y \in H$, we have

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

$$= (x,x) + (x,y) + (y,x) + (y,y)$$

$$= \|x\|^2 + (x,y) + (y,x) + \|y\|^2 \qquad ...(8)$$

Again $\|x-y\|^2 = (x-y, x-y)$

$$= (x,x) - (x,y) - (y,x) + (y,y)$$

$$= \|x\|^2 - (x,y) - (y,x) + \|y\|^2 \qquad ...(9)$$

Adding (8) and (9), we get

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$$

Remark : In a Hilbert space, the norm induced by the inner product satisfies the parallelogram law. However this is not true in general in Banach space i.e., the norm in a Banach space need not necessarily satisfies the parallelogram law.

Theorem 5 (Polarisation Identity) :

If x, y are any two vectors in a Hilbert space H, then

$$4(x,y) = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2$$

Proof : Subtracting (9) from (8), we get

$$\|x+y\|^2 - \|x-y\|^2 = 2(x,y) + 2(y,x) \qquad ...(10)$$

Replacing y by $i\,y$ in (10), we get

$$\|x+iy\|^2 - \|x-iy\|^2 = 2(x, iy) + 2(iy, x)$$

$$= 2\bar{i}\,(x,y) + 2i(y,x)$$

$$= -2\bar{i}\,(x,y) + 2i(y,x) \qquad ...(11)$$

Multiplying both sides of (11) by i, we get

$$i\|x+iy\|^2 - i\|x-iy\|^2 = 2(x,y) - 2(y,x) \qquad ...(12)$$

Adding (10) and (12) we get the required polarisation identity.

5.5 Some Important Theorems on Hilbert Spaces

Theorem 5 : If B is a complex Banach space whose norm obeys the parallelogram law, and if an inner product is defined on B by

$$4(x,y) = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2 \qquad ...(13)$$

then B is a Hilbert space.

Proof : For all $x, y \in B$, the parallelogram law is

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \qquad ...(14)$$

Now we show that the inner product on B satisfies the properties of Hilbert space.

(H_1) for $y = x$,

$$(13) \Rightarrow 4(x,x) = \|2x\|^2 - \|0\|^2 + i\|x(1+i)\|^2 - i\|x(1-i)\|^2$$

$$= 4\|x\|^2 - 0 + i|(1+i)|^2\|x\|^2 - i|(1-i)|^2\|x\|^2$$

$$= 4\|x\|^2 + 2i\|x\|^2 - 2i\|x\|^2$$

$$= 4\|x\|^2$$

$$\Rightarrow (x,x) = \|x\|^2$$

(H_2) Taking complex conjugates of both sides of (13), we get

$$4\overline{(x,y)} = \|x+y\|^2 - \|x-y\|^2 - i\|x+iy\|^2 + i\|x-iy\|^2$$

$$\left(\because \|x+y\|^2, \|x+y\|^2 \text{ each are real}\right)$$

$$= \|y+x\|^2 - \|-(y-x)\|^2 - i\|i(y-ix)\|^2 + i\|-i(y+ix)\|^2$$

$$= \|y+x\|^2 - \|y-x\|^2 - i|i|^2\|y-ix\|^2 + i|-i|^2\|y+ix\|^2$$

$$= \|y+x\|^2 - \|y-x\|^2 - i\|y-ix\|^2 + i\|y+ix\|$$

$$= 4(x,y) \qquad \text{(by (13))}$$

$$\therefore \quad \overline{(x,y)} = (y,x)$$

The property

$$\left(\alpha x+\beta y,z\right)=\alpha\left(x,z\right)+\beta\left(y,z\right)$$

is equivalent to

(H_{31}) $\quad (x+y,z)=(x,z)+(y,z)$

and (H_{32}) $\quad (\alpha x,y)=\alpha(x,y)$

so instead of proving H_3 we proove H_{31} and H_{32}.

(H_{31}) Replacing x by $(x+y)$ and y by z in (13), we get

$$4(x+y,z)=\left\|(x+y)+z\right\|^2-\left\|(x+y)-z\right\|^2+\left\|(x+y)+iz\right\|^2-i\left\|(x+y)-iz\right\|^2 \quad ...(15)$$

On replacing x by $(x+z)$ and using

$\left\|(x+y)+z\right\|^2=\left\|(x+z)+y\right\|^2$, (14) gives

$$\left\|(x+z)+y\right\|^2+\left\|(x+z)-y\right\|^2=2\left\|x+z\right\|^2+2\left\|y\right\|^2$$

or $$\left\|(x+y)+z\right\|^2=2\left\|x+z\right\|^2+2\left\|y\right\|^2-\left\|(x+z)-y\right\|^2 \quad ...(16)$$

Also $\left\|(x+z)-y\right\|^2=\left\|(z-y)+x\right\|^2$ (by (16))

$$=2\left\|z-y\right\|^2+2\left\|x\right\|^2-\left\|(z-y)-x\right\|^2$$

$$=2\left\|-(y-z)\right\|^2+2\left\|x\right\|^2-\left\|-\{(x+y)-z\}\right\|^2$$

$$=2\left\|y-z\right\|^2+2\left\|x\right\|^2-\left\|(x+y)-z\right\|^2 \quad ...(17)$$

Using (17) in (16) we get

$$\left\|(x+y)+z\right\|^2=2\left\|x+z\right\|^2+2\left\|y\right\|^2-2\left\|y-z\right\|^2-2\left\|x\right\|^2+\left\|(x+y)-z\right\|^2$$

$$\therefore \quad \left\|(x+y)+z\right\|^2-\left\|(x+y)-z\right\|^2=2\left\|x+z\right\|^2+2\left\|y\right\|^2-2\left\|y-z\right\|^2-2\left\|x\right\|^2 \quad ...(18)$$

Interchanging x and y in (18) we get

$$\left\|(x+y)+z\right\|^2-\left\|(x+y)-z\right\|^2=2\left\|y+z\right\|^2+2\left\|x\right\|^2-2\left\|x-z\right\|^2-2\left\|y\right\|^2 \quad ...(19)$$

Adding (18) and (19) we get

$$\left\|(x+y)+z\right\|^2-\left\|(x+y)-z\right\|^2$$

$$= \|x+z\|^2 - \|x-z\|^2 + \|y+z\|^2 - \|y-z\|^2 \qquad ...(20)$$

Now replacing z by iz and then multiplying throughout by i in (20), we get

$$i\|(x+y)+iz\|^2 - i\|(x+y)-iz\|^2$$

$$= i\|x+iz\|^2 - i\|x-iz\|^2 + i\|y+iz\|^2 - i\|y-iz\|^2 \qquad ...(21)$$

Adding (20) and (21) we find that

$$\|(x+y)+z\|^2 - \|(x+y)-z\|^2 + i\|(x+y)+iz\|^2 - i\|(x+y)+iz\|^2$$

$$= \left\{\|x+z\|^2 - \|x-z\|^2 + i\|x+iz\|^2 - i\|x-iz\|^2\right\}$$

$$+ \left\{\|y+z\|^2 - \|y-z\|^2 + i\|y+iz\|^2 - i\|y-iz\|^2\right\}$$

or $\quad 4(x+y,z) = 4(x,z) + u(y,z)$ (using polarisation identity)

or $\quad (x+y,z) = (x,z) + (y,z)$ ...(22)

(H_{32}) Let $\alpha \in C$. Then we prove H_{32} for following cases :

Case I : Let α is a positive integer

by (22) we have

$$(x+z,y) = (x,y) + (z,y)$$

Taking $z = x$, we get

$$(2x,y) = 2(x,y)$$

Hence (H_{32}) is true for $\alpha = 2$

Now assume that (H_{32}) is true for a fixed positive integer k i.e.

$$(k\,x,y) = k\,(x,y) \qquad ...(23)$$

Then $\quad ((k+1)\,x,y) = (k\,x+x,y)$

$= (k\,x,y) + (x,y)$ (by H_{31})

$= k\,(x,y) + (x,y)$ (by (23))

$= (k+1)\,(x,y)$

Thus H_{32} is true for $k+1$. Hence H_{32} is true for all positive integers k.

Case II : Let α be a negative integer.

Here first we prove that

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

For this replacing x by $-x$ in (13), we get

$$4(x, y) = \|-x + y\|^2 - \|-x - y\|^2 + i\|-x + i\,y\|^2 - i\|-x - i\,y\|^2$$

$$= \|-(x - y)\|^2 - \|-(x + y)\|^2 + i\|-(x - i\,y)\|^2 - i\|-(x + i\,y)\|^2$$

$$= \|x - y\|^2 - \|x + y\|^2 + i\|x - i\,y\|^2 - i\|x + i\,y\|^2$$

$$= -\,4(x, y) \qquad ...(24)$$

$\therefore \qquad (-x, y) = -(x, y)$

Now let $\alpha = -\beta$, where β is positive integer. Then

$$(\alpha\,x, y) = ((-\beta)\,x, y) = (-(\beta\,x),\ y)$$

$$= -(\beta\,x, y) = \alpha\,(x, y)$$

Case III : Let α be a rational number i.e., $\alpha = \dfrac{p}{q}$

where p and q are integers and $q \neq 0$. We have

$$(\alpha\,x, y) = \left(\frac{p}{q}x, y\right) = (p\,z, y) \qquad \text{(assume that } \frac{x}{q} = z\text{)}$$

$$= p(z, y)$$

Also $\quad (q\,z, y) = q\,(z, y) \Rightarrow (z, y) = \dfrac{1}{q}(q\,z, y)$

Hence $\quad (\alpha\,x, y) = \dfrac{p}{q}(q\,z, y)$

$$= \alpha\,(x, y)$$

Case IV : Let α be a complex number.

Here first we proove that

$$(i\,x, y) = i\,(x, y)$$

Replacing x by $i\,x$ in (13), we get

$$4(x,y)=\|ix+y\|^2-\|ix-y\|^2+i\|ix+iy\|^2-i\|ix-iy\|^2$$

$$=\|i(x-iy)\|^2-\|i(x+iy)\|^2+i\|i(x+y)\|^2-i\|i(x-y)\|^2$$

$$=|i|^2\ \|x-iy\|^2+|i|^2\ \|x+iy\|^2+i|i|^2\ \|x+y\|^2-i|i|^2\ \|x-y\|^2$$

$$=\|x-iy\|^2-\|x+iy\|^2+i\|x+y\|^2-i\|x-y\|^2$$

$$=-i^2\|x-iy\|^2+i^2\|x+iy\|^2+i\|x+y\|^2-i\|x-y\|^2$$

$$=\ i\left\{\|x+y\|^2-\|x-y\|^2+i\|x+iy\|^2-i\|x-iy\|^2\right\}$$

$$=i\ (x,y)$$

Now suppose that $\alpha=\alpha_1+i\alpha_2$, where $\alpha_1,\ \alpha_2\in R$.

We have

$$(\alpha x,y)=((\alpha_1+i\alpha_2)\ x,y)=(\alpha_1 x+i\alpha_2 x,y)$$

$$=(\alpha_1,x,y)+(i\alpha_2 x,y)=\alpha_1(x,y)+i(\alpha_2 x,y)$$

$$=\alpha_1(x,y)+i\alpha_2(x,y)$$

$$=(\alpha_1+i\alpha_2)(x,y)=\alpha(x,y)$$

Thus we have proved that

$(\alpha x,y)=\alpha(x,y)$ for each scalar α.

Hence B is Hilbert space.

Theorem 6 : A closed convex shubset K of aHilbert Space H contains a unique vectors of smallest norm.

Proof : Here first we define a convex set

Let X be a linear space real or complex. A norm empty subset K of X is said to be **convex** if $x,y\in K\Rightarrow(1-\lambda)x+\lambda y\in K$ where λ is any real number s.t. $0\le\lambda\le 1$.

Taking $\lambda=\frac{1}{2}$, we see that if K is convex subset of a linear space X, then $x,y\in K\Rightarrow\frac{x+y}{2}\in K$.

Now suppose that $d=\inf\{\|x\|:x\in K\}$. Then there exists a sequence $\{x_n\}$ in K s.t. $\|x_n\|\to d$.

Since K is convex, therefore $\frac{x_n+x_m}{2}\in K$ for $m,\ n\in N$.

Hence using the definition of d, we have

$$\left\|\frac{1}{2}(x_n + x_m)\right\| \geq d \Rightarrow \|x_n + x_m\| \geq 2d \qquad ...(25)$$

By parallelogram law

$$\|x_n - x_m\|^2 = 2\|x_n\|^2 + 2\|x_m\|^2 - \|x_n + x_m\|^2$$

$$\leq 2\|x_n\|^2 + 2\|x_m\|^2 - 4d^2 \qquad ...(26)$$

Since $\|x_n\|$, $\|x_m\| \to d$ as $n, m \to \infty$, we get from (26) that

$$\|x_n - x_m\|^2 \to 0 \quad \text{as} \quad m, n \to \infty$$

Hence $\{x_n\}$ is a Cauchy sequence. Since K is a closed subspace of a complete space, therefore K is complete. Hence the Cauchy sequence $\{x_n\}$ in K converges to a point x in K. Since the inner product is continuous and consequent6ly norm is also continuous.

Thus $\|x\| = \|\lim x_n\| = \lim \|x_n\| = d$

so x is a vector of smallest norm.

Uniqueness of x : Let $y \in K$ be another point with $\|y\| = d$.

Then $\frac{1}{2}(x+y) \in K$. Hence by parallelogram law, we get

$$\left\|\frac{1}{2}(x+y)\right\|^2 = 2\left\|\frac{x}{2}\right\|^2 + 2\left\|\frac{y}{2}\right\|^2 - \left\|\frac{1}{2}(x-y)\right\|^2$$

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

$$= d^2 - \left\|\frac{1}{2}(x-y)\right\|^2$$

$$< d^2$$

Which contradicts the definition of d, since $\frac{1}{2}(x+y) \in K$. Hence $x \in K$ is unique.

Theorem 7 : Let M be a closed linear subspace of a Hilbert space H, and x be a vecotr not in M. Suppose that $d = d(x, M)$. Then these exists a unique vector y_0 in M s.t. $\|x - y_0\| = d$.

Proof : We have $d = d(x, M) = \inf\{\|x - y\| : y \in M\}$

Then there exists a sequence $\{y_n\}$ in M s.t.

$\lim \|x - y_n\| = d$ or $\|x - y_n\| \to d$

Let $y_m, y_n \in \{y_n\}$ i.e., $y_m, y_n \in M$

$\Rightarrow \quad \dfrac{y_m + y_n}{2} \in M \qquad$ [$\because$ M is a subspace of H]

$\Rightarrow \quad \left\| x - \dfrac{y_m + y_n}{2} \right\| \geq d$

$\Rightarrow \quad \|2x - (y_m + y_n)\| \geq 2d$ ...(27)

By parallelogram law, we have

$$\|y_m - y_n\|^2 = \|(x - y_n) - (x - y_m)\|^2$$

$$= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - \|(x - y_n) + (x - y_m)\|^2$$

$$\leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4d^2$$

$$\to 2d^2 + 2d^2 - 4d^2 = 0 \quad \text{as} \quad m, n \to \infty$$

$\therefore$ $\{y_n\}$ is a Cauchy sequence in M which is complete. being a closed subspace of a complete space H

$\Rightarrow \quad \exists\, y_0 \in M$ s.t. $\{y_n\} \to y_0$

Now $\|x - y_0\| = \|x - \lim y_n\|$

$$= \|\lim(x - y_n)\| = \lim\|x - y_n\| = d$$

Hence y_0 is the required vector in M s.t. $\|x - y_0\| = d$

Uniqueness of y_0 : Let y_1, y_2 $(y_1 \neq y_2)$ be two vectors in M s.t.

$\|x - y_1\| = d = \|x - y_2\|$.

Now $y_1, y_2 \in M \Rightarrow \dfrac{y_1 + y_2}{2} \in M$

$$\Rightarrow \left\| x - \frac{y_1 + y_2}{2} \right\| \geq d$$

$$\Rightarrow \|2x - (y_1 + y_2)\| \geq 2d$$

By parallelogram law we have

$$\|(x-y_1)-(x-y_2)\|^2 = 2\,\|x-y_1\|^2 + 2\,\|x-y_2\|^2 - \|2x-(y_1+y_2)\|^2$$

$$\le 2d^2 + 2d^2 - 4d^2 = 0$$

$\therefore \qquad \|y_1 - y_2\|^2 \le 0 \Rightarrow \|y_1 - y_2\| = 0 \qquad \left(\because \|y_1 - y_2\|^2 \ge 0\right)$

$$\Rightarrow y_1 - y_2 = 0$$

$$\Rightarrow y_1 = y_2$$

Hence y_0 is unique.

Self-Learning Exercise

1. State linearity in the first variable for inner product.
2. If $x, y, z \in H$ (a Hilbert space) and $\alpha, \beta, \gamma \in C$, then fill up the blanks
   (i) $(\alpha x + \beta y, z) =$.....
   (ii) $(x, \beta y - \gamma z) =$.....
   (iii) $\overline{(x, y)} =$.....
3. Fill up the blanks
   (i) A .... inner product space is called a Hilbert space.
   (ii) The inner product in a Hilbert space is ....
4. State parallelogram law in a Hilbert space.
5. State polarisation identity in a Hilbert space.

5.6 Summary

In this unit you studied inner product space and Hilbert spaces and some basic properties associated with these spaces.

5.7 Answers to Self-Learning Exercise

1. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$
2. (i) $\alpha(x, z) + \beta(y, x)$ (ii) $\overline{\beta}(x, y) - \overline{\gamma}(x, z)$
3. (i) complete (ii) jointly continuous
4. $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \quad \forall\, x, y \in H$
5. $4(x, y) = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2$

5.8 Exercises

1. Let $L_2[0, 1]$ be the set of all square integrable functions on $[0, 1]$. Define the inner product on $L_2[0, 1]$ as

$$(f, g) = \int_0^1 g(t)\overline{g(t)}\, dt \qquad \forall f, g \in L_2[0, 1]$$

Prove that $L_2[0, 1]$ is an inner product space.

2. Give an example of an inner product space which is not a Hilbert space.

3. If X is an inner product space, show that $\sqrt{(x, x)}$ satisfies the properties of a norm.

4. If x and y are any two vectors in a Hilbert space H then show that

(i) $\|x+y\|^2 - \|x-y\|^2 = 4\,\text{Re}(x, y)$

(ii) $(x, y) = \text{Re}(x, y) + i\,\text{Re}(x, i\,y)$

5. For the special Hilbert space l_2^n, use Cauchy's inequality to prove the Schwarz inequality.

6. Define (i) Inner product space (ii) Hilbert space and give an example.

7. Let K be a non-empty conver subset of a Hilbert space H and $x_0 \in H$. Prove that $\exists$ a unique point $k_0 \in K$ s.t. $d(x_0, k) = \|x_0 - k_0\|$.

□□□

Unit - 6
Orthogonality and Functionals in Hilbert Spaces

Structure of the Unit

6.1 Objectives

Our objective of this unit is to study orthogenality and functionals in Hilbert spaces. We shall also study the orthonormal sets, complete orthonormal sets and reflexivity of Hilbert spaces.

6.2 Introduction

In the last unit we defined the inner product spaces and Hilbert spaces. In this unit first we define orthogonality in Hilbert spaces, and prove Pythagorean theorem, Projection theorem and some other important results connected with orthogonal complements. After that the definition of orthonormal sets and complete orthonormal sets are given and important theorems such as Bessel's inequality, Parseval's identity are proved. We also discuss functionals in Hilbert spaces and prove an important theorem viz Riesz representation theorem. Lastly we prove that every Hilbert space is Reflexive.

6.3 Orthogonal Complements

6.3.1 Definition 1 (Orthogonality) :

Let x and y be any two vectors in a Hilbert space H. Then x is said to be **orthogonal** to y written as $x \perp y$ if $(x, y) = 0$

From the definition we have the following easy consequences :

(i) The relation of orthognality is symmetric i.e.

$x \perp y \Rightarrow y \perp x$. Since $x \perp y$ gives

$(x, y) = 0 \Rightarrow \overline{(x, y)} = 0$ or $(y, x) = 0 \Rightarrow y \perp x$

(ii) If $x \perp y$, then $\alpha x \perp y \;\forall\; \alpha \in C$.

Since $(\alpha x, y) = \alpha (x, y) = 0$, therefore $x \perp y \Rightarrow \alpha x \perp y$

(iii) Since $(0, x) = 0$ for any $x \in H$, therefore $0 \perp x \;\forall\; x \in H$

(iv) If $x \perp x$, then x must be zero. For $x \perp x$, then $(x, x) = 0 \Rightarrow \|x\|^2 = 0$ i.e., $x = 0$

6.3.2 Definition 2 (Orthogonal Sets) :

Two non empty subsets S_1 and S_2 of a Hilbert space H are said to be orthogonal denoted by $S_1 \perp S_2$, if $x \perp y \;\forall\; x \in S_1$ and $y \in S_2$.

6.3.3 Definition 3 (Orthogonal Complement) :

Let S be a non empty subset of a Hilbert space H. The orthogonal complement of S denoted by $S^\perp$ and read as S perpendicular, is defined as

$$S^\perp = \{x \in H : x \perp y, \;\forall\; y \in S\}$$

Thus $S^\perp$ is the set of all those vectors in H which are orthogonal to every vector in S.

6.3.4 Pythagorean Theorem :

Statement If x and y are any two orthogonal vectors in a Hilbert space $\boldsymbol{H}$, then

$$\|x + y\|^2 = \|x - y\|^2 = \|x\|^2 + \|y\|^2$$

Proof : Since $x \perp y$ therefore

$(x, y) = 0 \Rightarrow \overline{(x + y)} = 0$

$\Rightarrow (y, x) = 0$ ...(1)

Now $\|x + y\|^2 = (x + y,\ x + y)$

$= (x, x) + (x, y) + (y, x) + (y, y)$

$= \|x\|^2 + \|y\|^2$ (using (1)) ...(2)

Similarly $\|x - y\|^2 = \|x\|^2 + \|y\|^2$ ...(3)

Combining (2) and (3) we get the Pythagorean theorem.

6.3.5 Elementary Properties :

From the definition we have the following

Theorem 1 : Let S, S_1 and S_2 be non empty subsets of a Hilbert space H. Then

(i) $\{0\}^{\perp} = H$ (ii) $H^{\perp} = \{0\}$

(iii) $S \cap S^{\perp} \subset \{0\}$ (iv) $S_1 \subset S_2 \Rightarrow S_2^{\perp} \subset S_1^{\perp}$ and $S_1^{\perp\perp} \subset S_2^{\perp\perp}$

(v) $S \subset S^{\perp\perp}$ (vi) $S_1 \perp S_2 \Rightarrow S_1 \cap S_2 = \{0\}$

Proof : (i) By definition we have $\{0\}^{\perp} \subset H$ ...(4)

Now let $x \in H$. Since $(x, 0) = 0$ $\therefore$ $x \in \{0\}^{\perp}$. Hence $H \subset \{0\}^{\perp}$ ...(5)

Combining (4) and (5) we get $\{0\}^{\perp} = H$

(ii) Let $x \in H^{\perp} \Rightarrow (x, y) = 0 \ \forall\, y \in H$

Choose $y = x$, then $(x, x) = 0 \Rightarrow \|x\|^2$ or $x = 0$

Thus $x \in H^{\perp} \Rightarrow x = 0$. Hence $H^{\perp} = \{0\}$

(iii) Let $x \in S \cap S^{\perp} \Rightarrow x \in S$ and $x \in S^{\perp}$

$\Rightarrow x \perp x$ or $(x, x) = 0$

$\Rightarrow \|x\|^2 = 0 \Rightarrow x = 0 \in \{0\}$

Thus $S \cap S^{\perp} \subset \{0\}$

(**Remark :** If S is subspace of H, then $S^{\perp}$ is also subspace of H. So both S and $S^{\perp}$ contain zero vector. Thus is S is subspace of H, then $0 \in S \cap S^{\perp} \Rightarrow S \cap S^{\perp} = \{0\}$.)

(iv) Let $x \in S_2^{+}$. Then x is orthogonal to every vector in S_2.

Since $S_1 \subset S_2$, therefore x is orthogonal to every vector in S_1

which implies $x \in S_1^{\perp}$. Thus $S_2^{\perp} \subset S_1^{\perp}$.

In a similar manner we can prove that $S_1^{\perp\perp} \subset S_2^{\perp\perp}$.

(v) Let $x \in S$. Then $(x, y) = 0 \ \forall \ y \in S^{\perp}$

so if $y \in S^{\perp}$, then from the definition of $S^{\perp\perp}$, $x \in S^{\perp\perp}$.

Thus $x \in S \Rightarrow x \in S^{\perp\perp}$. Hence $S \subset S^{\perp\perp}$

(vi) If $S_1 \cap S_2 \neq \{0\}$, then suppose that $x \in S_1 \cap S_2$.

Since $S_1 \perp S_2$, therefore $(x, x) = 0 \Rightarrow \|x\|^2 = 0 \Rightarrow x = 0$,

therefore $S_1 \cap S_2 = \{0\}$.

Theorem 2 : If S is a non empty subset of a Hilbert space H, then $S^{\perp}$ is a closed linear subspace of H and hence a Hilbert space.

Proof : By definition of $S^{\perp}$, we have

$$S^{\perp} = \{x \in H : (x, y) = 0 \ \forall \ y \in S\}$$

Since $(0, y) = 0 \ \forall \ y \in S$, therefore $0 \in S^{\perp}$ and so $S^{\perp}$ is non empty.

Let $x_1, x_2 \in S^{\perp}$ and α, β are scalars. Then

$$(x_1, y) = 0 \text{ and } (x_2, y) = 0 \quad \forall \ y \in S.$$

Hence for every $y \in S$, we get

$$(\alpha x_1 + \beta x_2, y) = \alpha (x_1, y) + \beta (x_2, y) = \alpha 0 + \beta 0 = 0$$

$\Rightarrow$ $\alpha x_1 + \beta x_2 \in S^{\perp}$

$\Rightarrow$ $S^{\perp}$ is a subspace of H.

Now we prove that $S^{\perp}$ is a closed subset of H. For this let $\{x_n\}$ be a sequence is $S^{\perp}$ converging to x in H.

Then we have to show that $x \in S^{\perp}$. For this we should prove that $(x, y) = 0 \ \forall \ y \in S$.

Since $x_n \in S^{\perp}$, therefore $(x_n, y) = 0 \ \forall y \in S$ and $n \in N$. Since inner product is a continuous function, therefore $(x_n, y) \to (x, y)$ as $n \to \infty$

Since $(x_n, y) = 0 \ \forall n$, therefore $(x, y) = 0$. Thus $x \in S^{\perp}$.

Hence $S^{\perp}$ is a closed subset of H.

Now $S^{\perp}$ is a closed subspace of Hilbert space H. So, $S^{\perp}$ is complete and hence it is a Hilbert space.

6.4 Projection Theorem

In this section, we shall first develop some preliminary results for the proof of Projection Theorem.

Theorem 3 : Let M be a proper closed linear subspace of a Hilbert space H. Then there exists a non-zero vector z_0 in H s.t. $z_0 \perp M$.

Proof : Since M is a proper closed subspace of H, therefore there eixsts a vector x in H which is not in M.

Let $d = d(x, M) = \inf \{\|x - y\| : y \in M\}$

As $x \notin M$, so $d > 0$. Again M is a closed subspace of H, so by Theorem 7 of unit 5, there exists a unique vector y_0 in M s.t. $\|x - y_0\| = d$. Suppose that $z_0 = x - y_0$.

Now $\|z_0\| = \|x - y_0\| = d > 0$

Hence z_0 is a non zero vector. We prove that $z_0 \perp M$. For this we must show that $(z_0, y) = 0$ $\forall \ y \in M$.

For any scalar α, consider

$$z_0 - \alpha\, y = x - y_0 - \alpha\, y = x - (y_0 + \alpha\, y)$$

Since M is a subspace of H and $y,\ y_0 \in M$, therefore $y_0 + \alpha\, y \in M$. Hence using the definition of d, we get

$$\|x - (y_0 + \alpha\, y)\| \geq d = \|z_0\|$$

Therefore $\|z_0 - \alpha\, y\| \geq \|z_0\|^2$

Now $\quad \|z_0 - \alpha\, y\|^2 - \|z_0\|^2 = (z_0 - \alpha\, y,\ z_0 - \alpha\, y) - (z_0, z_0) \geq 0$

or $\quad (z_0, z_0) - \overline{\alpha}(z_0, y) - \alpha(y, z_0) + \alpha\,\overline{\alpha}(y, y) - (z_0, z_0) \geq 0$

or $\quad -\overline{\alpha}(z_0, y) - \alpha(y, z_0) + \alpha\,\overline{\alpha}(y, y) \geq 0 \qquad \qquad ...(6)$

The result (6) is true for all scalars α. Let $\alpha = \beta(z_0, y)$ where β is any arbitrary real number. Then $\overline{\alpha} = \beta\overline{(z_0, y)}$. Using α and $\overline{\alpha}$ in (6) we get

$$-\beta\overline{(z_0, y)}(z_0, y) - \beta(z_0, y)\overline{(z_0, y)} + \beta^2 (z_0, y)\overline{(z_0, y)}\ \|y\|^2 \geq 0$$

or $\quad -2\beta|(z_0,y)|^2+\beta^2|(z_0,y)|^2\,\|y\|^2\geq 0$

or $\quad \beta|(z_0,y)|^2+\left[\beta\|y\|^2-2\right]\geq 0$ ...(7)

The relation (7) is true for all real β. Suppose that $(z_0,y)\neq 0$. Choosing β to be positive s.t. $\beta\|y\|^2<2$, then from (7). We have

$$\beta|(\beta_0,y)|^2\left[\beta\|y\|^2-2\right]<0$$

which contradicts (7). Hence $(z_0,y)=0$ showing that $z_0\perp y$.

Thus $z_0\perp y\ \forall\ y\in M\Rightarrow z_0\perp y$ which completes the proof of the theorem.

Theorem 4 : Let M be a linear subspace of Hilbert space H. Then M is closed if and only if $M=M^{\perp\perp}$.

Proof : Let $M=\left(M^{\perp}\right)^{\perp}=M^{\perp\perp}$ where M is a subspace of H.

Using Theorem 2, $M^{\perp\perp}$ is closed. Therefore M is closed conversly let M be a closed subspace of H.

We know that $M\subset M^{\perp\perp}$ (by Theorem 1, (v)).

Now let $M\neq M^{\perp\perp}$. Then M is a proper closed subspace of Hilbert space $M^{\perp\perp}$. Hence by Theorem 3, there exists a non-zero vector z_0 in $M^{\perp\perp}$ s.t. $z_0\perp M$ or $z_0\in M^{\perp}$.

Now $z_0\in M^{\perp}$ and $M^{\perp\perp}\Rightarrow z_0\in M^{\perp}\cap M^{\perp\perp}$ ...(8)

Since $M^{\perp}$ is a subspace of H, therefore

$M^{\perp}\bigcap M^{\perp\perp}=\{0\}$ (by Theorem 1, Remark (iii)) ...(9)

From (8) and (9) we have $z_0=0$ contradicting z_0 is a non-zero vector. Hence $M\subset M^{\perp\perp}$ can not be a proper inclusion. Hence we have $M=M^{\perp\perp}$.

Remark : By the Theorem 2, $M^{\perp}$ is closed subspace of H. So $M^{\perp}=\left(M^{\perp}\right)^{\perp\perp}=M^{\perp\perp\perp}$.

Theorem 5 : M and N are closed linear subspaces of Hilbert space H s.t. $M\perp N$, then the linear subspace $M+N$ is closed.

Proof : To show that $M+N$ is closed. We have prove that it contains all its limits points. Let z is a limit point of $M+N$. Then there exists a sequence $\{z_n\}$ in $M+N$ s.t. $z_n\to z$ in H. Now $M\perp N$, $M\cap N=\{0\}$ and $M+N$ is a direct sum of the subspaces M and N, therefore z_n can be written uniquely as $z_n=x_n+y_n$ where $x_n\in M$ and $y_n\in N$.

Taking $z_m=x_m+y_m$ and $z_n=x_n+y_n$, we have

$$z_m - z_n = (x_m - x_n) + (y_m - y_n).$$

Since $(x_m - x_n) \in M$ and $(y_m - y_n) \in N$, therefore $(x_m - x_n) \perp (y_m - y_n)$

Hence by Pythagorean theorem, we get

$$\|(x_m - x_n) + (y_m - y_n)\|^2 = \|x_m - x_n\|^2 + \|y_m - y_n\|^2$$

$$\Rightarrow \quad \|z_m - z_n\|^2 = \|x_m - x_n\|^2 + \|y_m - y_n\|^2 \qquad ...(10)$$

Since $\{z_n\}$ is convergent sequence in H, it is a Cauchy's sequence in H, therefore $\|z_m - z_n\|^2 \to 0$ as $m, n \to \infty$. Using it in (10) we get $\|x_m - x_n\|^2 \to 0$ and $\|y_m - y_n\|^2 \to 0$ as $m, n \to \infty$. Hence $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in M and N. Since H is complete and M and N are closed subspaces of a complete space H, therefore M and N are complete. Hence the Cauchy sequence $\{x_n\}$ converges to x in M and $\{y_n\}$ converges to y in N.

Now $z = \lim z_n = \lim x_n + \lim y_n = x + y \in M + N$.

Therefore $M + N$ is closed

Now we state projection theorem.

Theorem 6 : If M is a closed linear subspace of a Hilbert space H, then $H = M \oplus M^{\perp}$.

Proof : Since M is a subspace of H, therefore by Theorem 2, $M^{\perp}$ is a closed and $M \cap M^{\perp} = \{0\}$. Thus in order to prove the theorem it is sufficient to verify that $H = M + M^{\perp}$.

Now M and $M^{\perp}$ are closed subspaces of H, therefore by Theorem 5, $M + M^{\perp}$ is also a closed subspace of H.

Suppose that $N = M + M^{\perp}$, then we prove that $N = H$.

From the definition of N, we have $M \subset N$ and $M^{\perp} \subset N$.

Thus $N^{\perp} \subset M^{\perp}$ and $N^{\perp} \subset M^{\perp\perp}$. Hence $N^{\perp} \subset M^{\perp} \cap M^{\perp\perp} = \{0\}$.

$$\text{Now } N^{\perp} = \{0\} \Rightarrow N^{\perp\perp} = \{0\}^{\perp} = H \qquad ...(11)$$

Since $N = M + M^{\perp\perp}$ is a closed subspace of H, therefore

$$N^{\perp\perp} = N \qquad ...(12)$$

From (11) and (12) we get

$$N = M + M^{\perp\perp} = H$$

Self-Learning Exercise - I

1. Define orthogonal sets.
2. State Pythagorean theorem.

3. State Projection theorem.

4. Define orthogonal complement of a set.

Fill up the blanks

5. $\{0\}^{\perp}$ =...........

6. $H^{\perp}$ =........... where H is a Hilbert space.

7. If M and N are subspaces of a Hilbert space H and $M \perp N$, then $M \cap N$ =...........

8. If S is a non empty subset of a Hilbert space, then $S^{\perp}$ is a .... space.

9. Let M be a linear subspace of a Hilbert space H. Then M is closed iff...........

10. For any non empty subset M of a Hilbert space H, $M^{\perp}$ =.....$^{\perp\perp\perp}$

6.5 Orthonormal Sets

6.5.1 Definition 1 : Let H be a Hilbert space. If $x \in H$ s.t. $\|x\| = 1$ i.e., $(x, x) = 1$, then x is said to be a **unit or normal vector.**

6.5.2 Definition 2 : A non empty subset $\{e_i\}$ of the Hilbert space H is said to be an **orthonormal set** if

(a) $e_i \perp e_j$ or $(e_i, e_j) = 0 \ \forall \ i \neq j$

(b) $\|e_i\| = 1$ or $(e_i, e_j) = 1$ for every i.

or

A non-empty subset of Hilbert space is said to be an orthonormal set if it contains mutually orthogonal unit vector.

Remarks :

1. An orthonormal set cannot contain zero vector as $\|0\| = 0$.

2. If H contains only the zero vector, then it has no orthonormal sets.

3. Every Hilbert space $H \neq \{0\}$ posesses an orthonormal set

4. If $\{x_i\}$ is a non-empty set of mutually orthogonal vectors in H, then $\{e_i\} = \left\{\frac{x_i}{\|x_i\|}\right\}$ is an orthonormal set.

6.5.3 Example : In the Hilbert space l_2^n, the subset $\{e_1, e_2, \ldots, e_n\}$ where e_i is the i^{th} tuple with 1 in the i^{th} place and 0 elsewhere is an orthonormal set.

6.6 Important Theorems on Orthonormal sets

Theorem 7 : If $\{e_1, e_2, ..., e_n\}$ be finite orthonormal set in a Hilbert space H, and x be any vector in H, then

(i) $\sum_{i=1}^{n} |(x, e_i)|^2 \le |x|^2$ and (ii) $x - \sum_{i=1}^{n} (x, e_i) e_i \perp e_j \;\; \forall j$

Proof : Let $y = x - \sum_{i=1}^{n} (x, e_i) e_i$. Then

$$\|y\|^2 = (y, y)$$

$$= \left(x - \sum_{i=1}^{n} (x, e_i) e_i ,\; x - \sum_{j=1}^{n} (x, e_j) e_j \right)$$

$$= (x, x) - \sum_{i=1}^{n} (x, e_i)(e_i, x) - \sum_{j=1}^{n} \overline{(x, e_j)}(x, e_j)$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)\overline{(x, e_j)}(e_i, e_j) \qquad ...(12)$$

Now $(e_i, e_j) = 0$, $i \ne j$ and $(e_i, e_{\hat{i}}) = 1$ ...(13)

Therefore

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (x, e_i)\overline{(x, e_j)}(e_i, e_j) = \sum_{i=1}^{n} (x, e_i)\overline{(x, e_i)} \qquad ...(14)$$

Using (14) in (12) we get

$$\|y\|^2 = \|x\|^2 - \sum_{i=1}^{n} (x, e_i)\overline{(x, e_i)} - \sum_{i=1}^{n} (x, e_i)\overline{(x, e_i)} + \sum_{i=1}^{n} (x, e_i)\overline{(x, e_i)}$$

$$= \|x\|^2 - \sum_{i=1}^{n} |(x, e_i)|^2 \ge 0 \qquad \left(\because \|y\|^2 \ge 0\right)$$

which gives (i)

Again consider

$$\left(x - \sum_{i=1}^{n} (x, e_i)\, e_i , e_j \right) = (x, e_j) - \sum_{i=1}^{n} (x, e_i)\,(e_i, e_j)$$

$$= (x, e_j) - (x, e_j) = 0 \qquad \text{(by (13))}$$

This proves (ii).

The inequality given in (i) is also known as **Bessel's inequality for finite orthonormal sets.**

Theorem 8 : If $\{e_i\}$ is an orthonormal set in a Hilbert space H and if x is any vector in H, then the set $S = \{e_i : (x, e_i) \neq 0\}$ is either empty or countable.

Proof : For each positive integer n and fixed x, consider the set

$$S_n = \left\{ e_i : \left| (x, e_i)^2 > \left(\frac{\|x\|^2}{n} \right) \right| \right\}$$

Hence S_n contains atmost $(n-1)$ vectors, otherwise if S_n contains n or more vectors than n, then we have for $e_i \in S_n$

$$\sum_{i=1}^{n} |(x, e_i)|^2 > n \frac{\|x\|^2}{n} = \|x\|^2 \qquad \text{...(15)}$$

But by Theorem 1, we have

$$\sum_{i=1}^{n} |(x, e_i)|^2 \leq \|x\|^2, \; e_i \in S_n \qquad \text{...(16)}$$

Which contradicts (15). Hence S_n contains atmost $(n-1)$ vectors. Hence for each positive integer n, the set S_n is finite or countably infinite, since if

$x \perp e_i \;\; \forall i \Rightarrow (x, e_i) = 0 \; \forall i$ then $S = \phi$,

if S is non-empty then it is either finite or infinite. When S is finite, it is clearly countable but if it is infinite, it can be written as $S = \bigcup_{n=1}^{\infty} S_n$ with S_n not containing more than $(n-1)$ elements, because if $e_i \in S \Rightarrow (x, e_i) \neq 0$, then however small be the value of $|(x, e_i)|^2$, n can be choosen so large that

$$|(x, e_i)|^2 > \frac{\|x\|^2}{n} \text{ so that } e_i \in S \Rightarrow e_i \in S_n.$$

Now $S = \bigcup_{n=1}^{\infty} S_n \Rightarrow S$ is expressible as countable union of finite sets

$\Rightarrow S$ is countable.

Theorem 9 (Bessel's Inequality) : If $\{e_i\}$ is an orthonormal set in a Hilbert space H, then

$$\sum |(x, e_i)|^2 \leq \|x\|^2 \;\; \forall \; x \in H.$$

Proof : Let $S = \{e_i : (x, e_i) \neq 0\}$, then by Theorem 8, S is either empty or countable. If S is empty, then $(x,e_i)=0 \ \forall\ i \Rightarrow \sum |(x,e_i)|^2 = 0$

Hence $\sum |(x,e_i)|^2 = 0 \le \|x\|^2$.

So the inequality is satisfied when S is empty.

Let $S \neq \phi$, then S is finite or countably infinite. If S is finite, then suppose that $S = \{e_1, e_2, \ldots, e_3\}$ for some positive integer n. In this case we have by Theorem 7 that

$$\sum_{i=1} |(x,e_i)|^2 = \sum_{i=1}^{n} |(x,e_i)|^2 \le \|x\|^2$$

Secondly taking S as contally infinite, then the vectors in S can be arranged in a definite order s.t. $S = \{e_1, e_2, \ldots, e_n, \ldots.\}$. In this case

$$\sum |(x,e_i)|^2 = \sum_{n=1}^{\infty} |(x,e_n)|^2$$

This sum is well defined if the series $\sum_{n=1}^{\infty} |(x,e_n)|^2$ is convergent irrespective of any arrangement of its terms i.e., irrespective of the arrangements of vectors in S.

By the Bessel's inequality for finite case, $\sum_{i=1}^{n} |(x,e_i)|^2 \le \|x\|^2$ is true for every positivie integer n, and so it must be true in limit also i.e.

$$\lim_{n\to\infty} \sum_{i=1}^{n} |(x,e_i)|^2 = \sum_{ni=1}^{\infty} |(x,e_n)|^2 \le \|x\|^2$$

which follows that the series $\sum_{n=1}^{\infty} |(x,e_n)|^2$ is convergent. Moreover by the theory of absolute convergence, this convergent series having all its terms positive is absolutely convergent. Consequently its sum will not alter by arrangement of its terms, which completes the proof of the theorem.

Theorem 10 : If $\{e_i\}$ be an orthonormal set in a Hilbert space H and x be an arbitrary vector in H, then

$$x - \sum (x,e_i)\, e_i \perp e_i \ \text{ for } \forall\ j$$

Proof : Taking $S = \{e_i : (x,e_i) \neq 0\}$. There arise three cases :

Case I : If S is empty i.e., $(x,e_i)=0 \ \forall\ i$, then we define $\sum (x,e_i)\, e_i$ to be the zero vector $\mathbf{0}$, so that

$$x-\sum (x,e_i)\ e_i = x-0 = x$$

Since $S=\phi \Rightarrow (x,e_j)=0\ \forall\ j \Rightarrow x \perp e_j\ \ \forall\ j$

Case II : Let $S \neq \phi$ and S is finite, Then the result follows by Theorem 7 (ii)

Case III : Let $S \neq \phi$ and S is countally infinite. Then arranging the vectors of S in a definite order as $S=\{e_1, e_2, \ldots, e_n, \ldots..\}$.

We set $s_n = \sum_{i=1}^{n} (x,e_i) e_i$

so that for $m > n$, we have

$$\|s_m - s_n\|^2 = \left\|\sum_{i=n+1}^{m} (x,e_i)\ e_i\right\|^2 = \sum_{i=n+1}^{m} \|(x,e_i) e_i\|^2$$

$$= \sum_{i=n+1}^{m} |(x,e_i)|^2\ \|e_i\|^2 = \sum_{i=n+1}^{m} |(x,e_i)|^2 \quad \text{as} \quad \|e_i\|^2 = 1. \forall\, i$$

By Bessel's inequality, the series $\sum_{i=1}^{\infty} |(x,e_i)|^2$ is convergent, so that for $m,n \to \infty$, $\sum_{i=n+1}^{\infty} |(x,e_i)|^2$ can be made to converge to zero i.e.,

$$\|s_m - s_n\|^2 \to 0 \quad \text{as} \quad m,n \to \infty$$

$\Rightarrow$ the sequence $\{s_n\}$ is a Cauchy's sequence in H and H is complete

$\Rightarrow$ a vector s in H s.t. $\lim_{n\to\infty} s_n = s$

$\Rightarrow$ $s = \sum_{n=1}^{\infty} (x,e_i)\ e_n$

Now we can define $\sum (x,e_i) e_i = \sum_{n=1}^{\infty} (x,e_n)\ e_n$

Now we shall show that the above sum is well defined and does not depend upon the rearrangement of vectors.

For this suppose that the vectors in S are arranged in a different manner as $S=\{f_1, f_2, \ldots, f_n, \ldots\}$.

Let $s'_n = \sum_{n=1}^{n} (x, f_i)\ f_i$.

As proved above, let $s'_n \to s'$ in H where $s' = \sum_{n=1}^{\infty} (x, f_n)\ f_n$

Now we prove that $s = s'$.

For a given $\in > 0$, we can find n_0 s.t.

$$\forall\ n \geq n_0,\ \sum_{n_0+1}^{\infty} \left|(x, e_i)\right|^2 < \in^2,\ \left\|s_n - s\right\| < \in \text{ and } \left\|s'_n - s'\right\| < \in$$

For some positive integer $m_0 > n_0$, we can find all the terms of s_{n_0} in s'_{m_0} so that $s'_{m_0} - s'_{n_0}$ is a finite sum of terms of the type $(x, e_i)\ e_i$ for $i = n_0 + 1,\ n_0 + 2, \ldots\ldots$

Thus $\left\|s'_{m_0} - s'_{n_0}\right\|^2 \leq \sum_{i=n+1}^{\infty} \left|(x, e_i)\right|^2 < \in^2$ with $\left\|s'_{m_0} - s'_{n_0}\right\|^2 < \in^2$

$$\text{Now}\quad \left\|s' - s\right\|^2 = \left\|s' - s'_{m_0} + s'_{m_0} - s_{n_0} + s_{n_0} - s\right\|$$

$$\leq \left\|s' - s'_{m_0}\right\| + \left\|s'_{m_0} - s_{n_0}\right\| + \left\|s_{n_0} - s\right\|$$

$$\leq \in + \in + \in = 3\in$$

$$\to 0 \text{ as } \in \text{ is arbitrary}$$

Hence $s = s'$

Now consider

$$\left(x - \sum (x, e_i) e_i, e_j\right) = \left(x - s, e_j\right)$$

$$= \left(x, e_j\right) - \left(s, e_j\right) = \left(x, e_j\right) - \left(\lim s_n, e_j\right) \qquad \text{...(17)}$$

By continuity of the inner product we have

$$\left(\lim s_n, e_j\right) = \lim\left(s_n, e_j\right) \qquad \text{...(18)}$$

Using (18) in (17), we get

$$\left(x - \sum (x, e_i)\ e_i, e_j\right) = \left(x, e_j\right) - \lim\left(s_n, e_j\right)$$

If $e_j \neq S$, then $\left(s_n, e_j\right) = \left(\sum_{i=1}^{n} (x, e_i)\ e_i, e_j\right) = 0 \Rightarrow \lim\left(s_n, e_j\right) = 0$

Hence $\left(x - \sum (x, e_i)\ e_i, e_j\right) = \left(x, e_j\right) = 0$ as $e_j \neq S$

But if $e_j \in S$, then $\left(s_n, e_j\right) = \left(\sum_{i=1}^{n} (x, e_i)\ e_i, e_j\right)$ ...(19)

Now for $n > j$, we have

$$\left(\sum_{i=1}^{n}(x,e_i)\ e_i,e_j\right)=(x,e_j) \quad ...(20)$$

Using (20) in (19), we get

$$\lim_{n\to\infty}(s_n,e_j)=(x,e_j)$$

So in this case $\left(x-\sum (x,e_i)\ e_i,e_j\right)=(x,e_j)-(x,e_j)=0$

Thus $\left(x-\sum (x,e_i)\ e_i,e_j\right)=0$ for each j.

Hence $x-\sum (x,e_i)\ e_i \perp e_j$ for each j, which completes the proof of the theorem.

6.7 Complete Orthonormal Sets

6.7.1 Definition 1 : An orthonormal set S in a Hilbert space is **complete,** if there exists no other orthonormal set containing S. This is S must be a maximal orthonormal set.

Thus an orthonormal set $\{e_i\}$ in a Hilbert space is complete if it is not possible to adjoin a vector e to $\{e_i\}$ in such a way that $\{e_i,e\}$ is an orthonormal set properly containing $\{e_i\}$.

6.7.2 Definition 2 : Let $\{e_i\}$ be a complete orthonormal set in a Hilbert space H and x be any arbitrary vector in H. Then the numbers (x,e_i) are called the **Fourier coefficients of** x**.**

6.7.3 Definition 3 : The expansion $x=\sum (x,e_i)e_i$ is called the **Fourier expansion** of x.

6.7.4 Definition 4 : The expansion $\|x\|^2=\sum|(x,e_i)|^2$ is called the **Parseval's equation** or **Parseval's identity.**

6.7.5 Criterian for Complete Orthonormal Set

Theorem 11 : An orthonormal set S in a Hilbert space H is complete iff $x \perp S \Rightarrow x=0$ $\forall\ \boldsymbol{x}\in \boldsymbol{H}$.

Proof : Let S be complete and x is any non zero vector in H s.t., $x \perp S$. Then the set $S\cup\{e\}$ where $e=\left(\frac{x}{\|x\|}\right)$ is an orthonormal set properly containing S, contradicting the maximality of S. Hence $x=0$.

Conversly let $x\perp S \Rightarrow x=0$. If S is non complete, then $\exists$ some orthonormal set S' such that $S'\supset S$ properly. In that case, let $x=S'-S$. Since $\|x\|=1$ and $x \perp S$, $x\neq 0$ contradicting the given condition. Hence S must complete.

6.7.6 Example : In the Hilbert space l_2^n, the set $\{e_1,e_2,\ldots,e_n,\ldots\}$, where e_n is a sequence with 1 in the n^{th} place and 0's elsewhere, is a complete orthonormal set.

Solution : Let $S=\{e_1,e_2,\ldots,e_n,\ldots\}$. If $x=\{x_n\}$ and $y=(y_n)\in l_2^n$, then

$$(x,y)=\sum_{n=1}^{\infty}x_n\bar{y}_n \quad \text{and} \quad \|x\|^2=\left[\sum_{n=1}^{\infty}|x_n|^2\right]^{1/2}$$

As noted before S is an orthonormal set. Let $x\perp S$.

Now $(x,e_1)=x_1.1+x_2.0+\ldots+x_n.0=0$

$$\Rightarrow x_1=0$$

Similarly $x\perp e_2,\ldots,\ x\perp e_n,\ldots$ will give $x_2=0$, $x_3=0,\ldots,\ x_n=0,\ldots$. Hence $x\perp S\Rightarrow x=0$ therefore the orthonormal set is complete.

6.6.7 Properties of Complete Orthonormal Sets

Theorem 12 : If H be a Hilbert space and $\{e_i\}$ be an orthonormal set in H, then the following statements are equivalent :

(i) $\{e_i\}$ is complete

(ii) $x\perp\{e_i\}\Rightarrow x=0$

(iii) If x is an arbitrary vector in H, then

$$x=\sum(x,e_i)e_i$$

(iv) If x is an arbitrary vector in H, then

$$\|x\|^2=\sum|(x,e_i)|^2$$

Proof : (i) $\Rightarrow$ (ii)

Let $\{e_i\}$ is complete, we claim that

$$x\perp\{e_i\}\Rightarrow x=0$$

Suppose that $x\perp\{e_i\}$ and $x\neq 0$, then we can find a unit vector $e=\left\{\frac{x}{\|x\|}\right\}$ with $\|e\|=1$, s.t.

$e\perp\{e_i\}\Rightarrow(e,e_i)=0$ for each i.

Thus (e,e_i) is an orthonormal set which properly contains $\{e_i\}$ which contradicts the completeness of $\{e_i\}$. Hence our assumption i.e., $x\neq 0$ is wrong and so $x\perp\{e_i\}\Rightarrow x=0$

(ii) $\Rightarrow$ (iii)

Let $x\perp\{e_i\}\Rightarrow x=0$

Choosing an $e_j \in \{e_i\}$, we claim that the vector

$$x - \sum (x, e_i) e_i \perp e_j$$

For this consider

$$\left(x - \sum (x, e_i) e_i,\, e_j\right) = (x, e_j) - \left(\sum (x, e_i) e_i,\, e_j\right)$$

$$= (x, e_j) - \sum (x, e_i)\,(e_i,\, e_j)$$

$$= (x, e_j) - (x, e_i)\,(e_j,\, e_j) = 0$$

$\Rightarrow \quad \left(x - \sum (x, e_i)\, e_i\right) \perp e_j$ for each j

$\Rightarrow \quad \left(x - \sum (x, e_i)\, e_i\right) \perp \{e_i\} \Rightarrow x - \sum (x, e_i)\, e_i = 0 \qquad (\because x \perp \{e_i\} \Rightarrow x = 0)$

$\Rightarrow \quad x = \sum (x, e_i)\, e_i$

(iii) $\Rightarrow$ (iv) : Given that for any vector x in H s.t. $x = \sum (x, e_i)\, e_i$.

To prove that $\|x\|^2 = \sum |(x, e_i)|^2$.

We have

$$\|x\|^2 = (x, x) = \left(\sum (x, e_i)\, e_i,\, \sum (x, e_j)\, e_j\right)$$

$$= \sum (x, e_i)\, \overline{\sum (x, e_j)}\, (e_i, e_j)$$

$$= \sum_i \sum_j (x, e_i)\overline{(x, e_j)}(e_i, e_j)$$

$$= \sum_i (x, e_i)\overline{(x, e_i)}(e_i, e_j)$$

$$= \sum_i (x, e_i)\overline{(x, e_i)} \quad \text{as} \quad (e_i, e_j) = \|e_i\|^2 = 1$$

$$= \sum |(x, e_i)|^2$$

(iv) $\Rightarrow$ (i) : Given $\|x\|^2 = \sum |(x, e_i)|^2$. To prove that $\{e_i\}$ is complete.

Let $\{e_i\}$ be not complex. Then $\{e_i\}$ is a proper subset of an orthonormal set $\{e_i, e\}$. Hence taking e for x in the hypothesis, we get

$$\|e\|^2 = \sum |(e, e_i)|^2 = 0. \text{ Since } e \perp e_i \;\; \forall\, i,$$

Thus $\|e\|^2 = 0$ which contradicts that e is a unit vector. Therefore $\{e_i\}$ is a complete orthonormal set

Self-Learning Exercise - II

1. Define an orthonormal set.
2. Define a complete orthonormal set.
3. Define a Fourier series for a vector $\boldsymbol{x}$ in Hilbert space H.
4. An orthonormal set contains a zero vector (T/F)
5. Every Hilbert space $H \neq \{0\}$ possesses on orthonormal set (T/F).
6. State Bessel's inequality in a Hilbert space.
7. Complete the following statements :
   (a) An orthonormal S in a Hilbert space H is complete iff for any x in H, $x \perp S \Rightarrow$......
   (b) If $\{e_i\}$ is an orthonormal set in H, then
      (i) $\{e_i\}$ is ....
      (ii) $x \perp \{e_i\} \Rightarrow x =$...
   (c) If $\{e_1, e_2\}$ is a orthonormal set in a Hilbert space H, then $\|e_1 - e_2\| =$....
   (d) Every non-zero Hilbert space contains a .... set.

6.8 Functional in Hilbert Sapces

If H is a Hilbert space and if we define a **continuous linear functional** or simply a **functional** on H as a continuous linear transformation from H into $\boldsymbol{C}$, then the set of all these functionals constitutes a vector space denoted by H^* are known at the **conjugate space** of H.

The elements of H^* are known as functionals and denoted by f. Thus if $f \in H^*$, then f is a functional in H^* and as mentioned above f is a continuous linear functional on H. If we define addition and scalar multiplication in H^* pointwise and the norm of $f \in H^*$ is defined as

$$\|f\| = \sup \{|f(x)| : \|x\| \leq 1\}$$

then H^* is a Banach space. By defining a suitable inner product on H^* it is seen that H^* maintains the structure of a Hilbert space. As such the conjugate space on H^* is second conjugate space $(H^*)^*$ or H^{**} of H also becomes a Hilbert space.

Theorem 13 : Let y be a fixed elements of Hilbert space H and f_y be a scalar valued functional on H defined as $f_y(x) = (x, y),\ \forall\ x \in H$.

Then the mapping f_y is a functional on H and $\|y\| = \|f_y\|$.

Proof : From the definition, we have $f_y : H \to \boldsymbol{C}$. Now we prove that f_y is linear and continuous so that it is a functional.

Let $x_1, x_2 \in H$ and $\alpha, \beta \in \boldsymbol{C}$. Then for fixed $y \in H$, we have

$$f_y(\alpha\, x_1, \beta\, x_2) = (\alpha\, x_1 + \beta\, x_2, y)$$

$$= \alpha(x_1, y) + \beta(x_2, y) = \alpha\, f_y(x_1) + \beta\, f_y(x_2)$$

$\Rightarrow f_y$ is linear.

Also for any $x \in H$,

$$|f_y(x)| = |(x, y)| \le \|x\|\,\|y\| \qquad \text{(by Schwarz inequality)} \qquad ...(21)$$

Now let $\|y\| \le M$. Then $M > 0$, we get

$|f_y(x)| \le M\,\|x\| \Rightarrow f_y$ is bounded hence continuous.

Hence f_y is a functional.

Again if $y = \mathbf{0}$, then $\|y\| = 0$ and from definition $f_y = 0$ so that $\|f_y\| \le \|y\|$.

Suppose that $y \ne 0$, then from (21), we get

$$\sup \frac{|f_y(x)|}{\|x\|} \le \|y\| \Rightarrow \|f_y\| \le \|y\| \qquad ...(22)$$

Further since $y \ne 0$, therefore $\frac{y}{\|y\|}$ is a unit vector setting $x = \frac{y}{\|y\|}$ in the definition

$\|f_y\| = \sup\{|f_y(x)| : \|x\| \le 1\}$, we get

$$\|f_y\| \ge \left|f_y\left(\frac{y}{\|y\|}\right)\right| = \left(\frac{y}{\|y\|},\, y\right) = \frac{1}{\|y\|}(y, y) = \|y\|$$

Hence $\|f_y\| \ge \|y\|$ ...(23)

Thus (22) and (23) gives $\|y\| = \|f_y\|$.

From the above theorem, we can say that $T : H \to H^*$ s.t. $T(y) = f_y$ is a norm preserving mapping.

Now we shall prove that every $f \in H^*$ arises in this manner.

Theorem 14 (Risez Representation Theorem) : Let H be a Hilbert space and f be an arbitrary functional in $H*$. Then there exists a unique vector y in H s.t. $f(x)=(x,y) \forall f\ x \in H$ and $\|f\|=\|y\|$.

Proof: Let $\exists$ a vector $y \in H$ s.t. $f(x)=(x,y)\ \forall\, x \in H$. We first prove that y is unique.

Suppose that y is not unique i.e. $\exists$ two vectors $y_1, y_2 \in H$ corresponding to a functional $f \in H*$ s.t.

$$f(x)=(x,y_1) \text{ and } f(x)=(x,y_2)\quad \forall\ x \in H$$

$\Rightarrow \quad (x,y_1)=(x,y_2)\ \forall\, x \in H$

$\Rightarrow \quad (x,y_1-y_2)=0\ \forall\, x \in H$

Taking $x=y_1-y_2$, we get $(y_1-y_2, y_1-y_2)=\|y_1-y_2\|^2=0 \Rightarrow y_1-y_2=0 \Rightarrow y_1=y_2$

Hence y is unique

Next we prove that y exists.

If f is a zero functional i.e. $f=\mathbf{0}$, then

$$f(x)=0\ \forall\ x \in H \text{ and } f(x)=(x,y)\ \forall\ x \in H$$

$\Rightarrow \quad (x,y)=0 \forall\ x \in H$

$\Rightarrow \quad y=\mathbf{0}$ which shows that $y=\mathbf{0}$ exists when $f=\mathbf{0}$.

If $f \neq \mathbf{0}$ i.e. $f(x)\neq 0$ for some $x \in H$, then consider null space say M of f s.t.

$$M=\{x : f(x)=0\},\ x \in H$$

We observe that

(a) M is non empty : Since $f(0)=0$: $\mathbf{0} \in M$

(b) M is a subspace : Since

If $x_1, x_2 \in M$ and α, β are scalars s.t. $f(x_1)=0$, $f(x_2)=0$, then

$$\begin{aligned} f(\alpha x_1+\beta x_2) &= (\alpha x_1+\beta x_2, y) \\ &= \alpha(x_1,y)+\beta(x_2,y) \\ &= \alpha f(x_1)+\beta f(x_2)=\alpha.0+\beta.0=0 \end{aligned}$$

$\therefore\ x_1, x_2 \in M$ and α and β are scalars $\Rightarrow \alpha x_1+\beta x_2 \in M$

(c) M is a proper subspace of H : Since $f(x)\neq 0$ for some

$x \in H \Rightarrow$ all such x do not belong to M

$\Rightarrow$ there are elements of H which are not in M

$\Rightarrow$ M is a proper subspace of H.

(d) M is a closed subspace of H : Since M is a subspace of a complete space H, therefore M is closed

Thus f is continuous and M is a proper closed subspace of H, therefore $\exists$ a non zero vector $y_0 \in H$ s.t. $y_0 \perp M$ or $y_0 \in M^{\perp}$ or we can say that $(y_0, x) = 0 \ \forall \ x \in M$.

Now we prove that $\exists$ a vector $y \in M$ s.t. $f(x) = (x, y) \ \forall \ x \in H$. Three cases arise.

Case I : If $x \in H$ and $x \in M \Rightarrow f(x) = 0$

Also $\quad f(x) = (x, y) = (x, \alpha\, y_0) \qquad$ (choosing $y = \alpha\, y_0$ with $y_0 \perp M$)

$$= \overline{\alpha}(x, y_0) = 0 \text{ as } x \in M \text{ and } y_0 \perp M$$

Hence $f(x) = (x, y)$ is satisfied for $x \in M$ and $y = \alpha\, y_0$

Case II : If $x \in H$ and $x = y_0$, then

$$f(x) = (x, y) \Rightarrow f(y_0) = (y_0, \alpha\, y_0) \quad \text{(choosing } y_0 = \alpha\, y_0\text{)}$$

$$= \overline{\alpha}\, \|y_0\|^2$$

giving $\quad \overline{\alpha} = \dfrac{f(y_0)}{\|y_0\|^2} \Rightarrow \alpha = \dfrac{\overline{f(y_0)}}{\|y_0\|^2}$

Then $f(x) = (x, y)$ is satisfied $\forall \ x \in M$ with $x = y_0$ and $y = \alpha\, y_0$.

Case III : $x \in H$, and $x \notin M$ with $x \neq y_0$

Since $H = M \oplus M^{\perp}$, therefore any vector $x \in H$ is uniquely expressible as the sum of the vector $m \in M$ and a vector $\beta\, y_0 \in M^{\perp}$ i.e., $x = m + \beta\, y_0$, β is a scalar.

By definition of m,

$$f(m) = 0 \Rightarrow f(x - \beta\, y_0) = 0$$

$$\Rightarrow f(x) - \beta\, f(y_0) = 0 \Rightarrow \beta = \frac{f(x)}{f(y)}$$

$\therefore \quad f(x) = f(m + \beta\, y_0) = f(m) + \beta\, f(y_0)$

$$= (m, y) + \beta(y_0, y) = (m + \beta\, y_0, y)$$

$$= (x, y) \ \forall \ x \in H$$

Lastly we show that $\|f\| = \|y\|$.

Now for each $x \in H$ $\exists$ a unique $y \in H$ s.t.

$$f(x) = (x, y) \quad \text{or} \quad |f(x)| = |(x, y)| \le \|x\| \, \|y\|$$

$$\Rightarrow \|f\| \le \|y\| \quad \text{(by def. of norm of a functional for which } \|x\| \le 1)$$

In the case $\|y\| = 0$ or $y = 0$ then $|f(x)| = |(x, 0)| = 0 \;\; \forall \; x$

and so $\|f\| = \sup \left\{ \dfrac{|f(x)|}{\|x\|} : x \ne 0 \right\} = 0$

$$\Rightarrow \|f\| = \|y\|, \text{ since } |f(x)| \le \|x\| \, \|y\|$$

$$\Rightarrow \frac{|f(x)|}{\|x\|} \le \|y\|$$

If $y \ne 0$, then $\|f\| = \sup \{ f(x) : \|x\| = 1 \}$

$$\ge \left| f\left(\frac{y}{\|y\|} \right) \right| \quad \text{on setting } x = \frac{y}{\|y\|} \quad \text{or} \quad \|x\| = 1$$

$$= \left| \left(\frac{y}{\|y\|}, y \right) \right| = \frac{1}{\|y\|} |(x, y)|$$

$$= \frac{1}{\|y\|} \|y\|^2 = \|y\|$$

$\therefore \quad \|f\| \ge \|y\|$

so $\quad \|f\| \le \|y\|$ and $\|f\| \ge \|y\| \Rightarrow \|f\| = \|y\|$

Theorem 15 : The mapping $\psi : H \to H^*$ defined by $\psi(y) = f_y$ where $f_y(x) = (x, y)$ for every $x \in H$ is an additive, one-to-one onto isometry but not linear.

Proof : (i) we have for $y_1, y_2 \in H$, $\psi(y_1 + y_2) = f_{y_1 + y_2}$

Hence for every $x \in H$, we get

$$f_{y_1 + y_2}(x) = (x, y_1 + y_2)$$

$$= (x, y_1) + (x, y_2)$$

$= f_{y_1}(x) + f_{y_2}(x)$

Hence $f_{y_1+y_2} = f_{y_1} + f_{y_2} \Rightarrow \psi(y_1 + y_2) = \psi(y_1) + \psi(y_2)$

Hence ψ is an additive.

(ii) ψ is one-one : Let $y_1, y_2 \in H$. Then $\psi(y_1) = f_{y_1}$ and $\psi(y_2) = f_{y_2}$.

Then $\psi(y_1) = \psi(y_2) \Rightarrow f_{y_1} = f_{y_2}$

$$\Rightarrow f_{y_1}(x) = f_{y_2}(x) \qquad \forall x \in H$$

$$\Rightarrow (y_1, x) = (y_2, x)$$

$$\Rightarrow (y_1 - y_2, x) = 0 \qquad \forall x \in H$$

Choose $x = y_1 - y_2$, then we get $(y_1 - y_2, y_1 - y_2) = 0 \Rightarrow \|y_1 - y_2\|^2 = 0 \Rightarrow y_1 - y_2 = \mathbf{0}$

Thus $\psi(y_1) = \psi(y_2) \Rightarrow y_1 = y_2 \Rightarrow \psi$ is one-one.

(iii) ψ is onto : Let $f \in H^*$, then $\exists\ y \in H$ s.t. $f(x) = (x, y)$ since $f_y(x) = (x, y)$, therefore we get $f = f_y$, so that $\psi(y) = f_y = f$. Hence for $f \in H^*$ $\exists$ a pre-image y in H. Thus ψ is onto.

(iv) ψ is isometry : Let $y_1, y_2 \in H$. Then

$$\|\psi(y_1) - \psi(y_2)\| = \|f_{y_1} - f_{y_2}\| = \|f_{y_1} + f_{-y_2}\|$$

$$= \|f_{y_1 - y_2}\| = \|y_1 - y_2\|$$

Hence ψ is **isometry.**

(v) ψ is not linear : Let $y \in H$ and α be any scalar. Then

$$\psi(\alpha y) = f_{\alpha y} = (x, \alpha y) = \overline{\alpha}(x, y) = \overline{\alpha} f_y(x) = \overline{\alpha}\psi(y)$$

Thus the mapping is not linear. Such a mapping is called conjugate linear.

We shall refer to the above mapping ψ as the natural mapping between H and H^*.

6.8 Reflexivity of Hilbert Space

Theorem 16 : Every Hilbert space is reflexive.

Proof : We prove that the natural inbedding on H to H^{**} is an **isometric isomorphism.**

Suppose that x be any fixed element of H and F_x be a scalar valued function defined on H^* by $F_x(f) = f(x)\ \forall f \in H^*$. Then F_x will be a functional on H^* i.e. $F_x \in H^{**}$. Thus each vector $x \in H$ gives rise to a functional F_x in H^{**}. F_x is called the functional on H^* induced by the vector x.

Let $T : H \to H^{**}$ defined by $T(x) = F_x \ \forall \ x \in H$.

From the theory of Banach spaces T is an isometric isomorphism of H into H^{**}. We shall show that T is a mapping of H onto H^{**}.

Let T_1 be a mapping from H into H^* s.t. $T_1(x) = f_x$ where $f_x(y) = (y, x) \ \forall \ y \in H$ and T_2 be a mapping from H^* into H^{**} defined by $T_2(f_x) = F_{f_x}$, where $F_{f_x}(f) = (f, f_x)$ for $f \in H^*$. Then $T_2 \ T_1$ is a composition of T_2 and T_1 from H to H^{**}. By Theorem 15, T_1 and T_2 are one-one and onto. Hence $T_2 \ T_1$ is the same as the natural imbedding T. For this we prove that $f(x) = (T_2 \ T_1)x \ \forall \ x \in H$

Now $(T_2 \ T_1)x = T_2(T_1(x)) = T_2(f_x) = F_{f_x} = T(x)$. In order to show that $T_2 \ T_1 = T$, we should prove that $F_x = F_{f_x}$ for this let $f \in H^*$. Then $f = f_y$ where f corresponds to y in the representation $F_{f_x}(y) = (f, f_x) = (f_y, f_n) = (x, y)$.

But $(x, y) = f_y(x) = f(x) = F_x(f)$. Thus we get $F_{f_x}(f) = F_x(f)$ for every $f \in H^*$. Hence the mappings F_{f_n} and F_x are equal i.e., $T_2 \ T_1 = T$ and so T is a mapping of H onto H^{**} so that H is reflexive.

From the above, we get

$$(F_x, F_y) = (F_{f_x}, F_{f_y}) = (f_y, f_x) = (x, y)$$

Hence f is an isometric isomorphism of H onto H^* so that H and H^{**} are conjugate.

Self-Learning Exercise - III

1. Define a functional on a Hilbert space.
2. State Riesz representation theorem.
3. Every Hilbert space is reflexive (T/F)
4. Riesz representation theorem is valid in an inner product space which is not complete (T/F).

6.9 Summary

In this unit you studied orthogonality and functionals in Hilbert spaces. Orthonormal sets, complete orthonormal sets and reflexivity a Hilbert spaces were defined and important results connected with them were also proved.

6.10 Answers to Self-Learning Exercises

Exercise - I

4. H 5. ϕ 6. $\{0\}$ 7. closed linear 8. $M = M^{\perp\perp}$

9. M

Exercise - II

4. F 5. T 7 (a) $x = \mathbf{0}$ (b) (i) complete (ii) $\mathbf{0}$

7. (c) $\sqrt{2}$ (d) complete orthonormal

Exercise - III

3. T 4. F

6.11 Exercises

1. If M be a non-empty subset of a Hilbert space H, then show that $M^{\perp\perp}$ is the closure of the set of all linear combinations of vectors in M i.e. $M^{\perp\perp} = \left[\overline{M}\right]$.
2. Prove that in the Hilbert space l_2, the set $\{e_1, e_2, \ldots, e_n, \ldots\}$ where e_n is a sequence with 1 in the nth place and 0's elsewhere is an orthonormal set.
3. State and prove Bessel's inequality in Hilbert spaces.
4. Prove that a Hilbert space H is a separable if every orthonormal set in H is countable.
5. Prove that an orthonormal set in a Hilbert space is linearly independent.
6. Prove that every orthonormal set in a Hilbert space is contained in some complete orthonormal set.
7. Show that every non-zero Hilbert space contains a complete orthonormal set.
8. If H is a Hilbert space, then prove that H^* is also a Hilbert space with the inner product defined by $(f_x, f_y) = (y, x)$.
9. Prove that conjugate space H^{**} of H^* is a Hilber space with some inner product defined on it.

□□□

Unit - 7
Operators on Hilbert Spaces

Structure of the Unit

7.1 Objectives

The objective of thus unit is to study various operators such as adjoint, self-adjoint, positive, normal and unitary operators on Hilbert spaces. Various properties and results on these operators will also be proved.

7.2 Introduction

In this unit, we shall introduce the adjoint of a bounded linear operator on a Hilbert space. With the help of the adjoint of a bounded linear operator, we shall define three important cases of operators called self-adjoint, normal and unitary operators. Besides this, we shall discuss in details the properties of these operators.

7.3 Adjoint Operator

7.3.1 Definition : The operator T* defined on H s.t.

$$(Tx, y) = (x, T^* y) \quad \forall \; x, y \in H$$

is called the adjoint of T.

7.3.2 Remark : Though we are using the same symbol for the conjugate and adjoint operators, one should note that the conjugate operator is defined on H^* and operates on functionals in H^*, whereas if T^* is adjoint of the operator T, then it is an operator on H and operates on vectors in H. However if we identify H and H^* under the natural correspondence, then the adjoint of T and conjugate of T coincide.

7.3.3 Important Theorems

Theorem 1 : Let T be an operator on a Hilbert space H, then $\exists$ a unique linear operator T^* on H s.t.

$$(Tx, y) = (x, T^* y) \; \forall \, x, y \in H$$

obviously T^* is the adjoint operator H.

Proof : First we prove that T^* exists.

Let y be a vector in H and f_y its corresponding functional in H^*. Define T^* on H^* into H^* by

$$T^* f_y = f_z$$

Under the natural correspondence between H and H^*, let $z \in H$ corresponding to $f_z \in H^*$. Thus starting with a vector y in H, we arrive at the vector z in H in the following manner

$$y \to f_y \to T^* f_y = f_z \to z$$

where $T^* : H^* \to H^*$ and $y \to f_y$ and $z \to f_z$ are on H to H^* under the natural correspondence. The product of the above there mappings exists and it is denoted by T^*.

Thus T^* is a mapping on H into H s.t. $T^* y = z$.

We define this T^* to be adjoint of T.

Now we prove (1). for $x \in H$ and from the definition of the conjugate T^* of an operater T,

$$(T^* f_y)\, x = f_y(Tx) \qquad ...(3)$$

By Riesz representation theorem, $y \to f_y$ so that we get

$$f_y(Tx) = (Tx, y) \qquad ...(4)$$

Since T^* is defined on H^*, we have

$$(T^* f_y) x = f_z(x) = (x, z) \qquad ...(5)$$

But according to our definition

$$T^* y = z \qquad \text{...(6)}$$

From (3) and (4), we get

$$\left(T^* f_y\right) x = \left(Tx, y\right) \qquad \text{...(7)}$$

and from (5) and (6), we get

$$\left(T^* f_y\right) x = \left(x, T^* y\right) \qquad \text{...(8)}$$

From (7) and (8), we get

$$\left(Tx, y\right) = \left(x, T^* y\right) \;\forall\; x, y \in H$$

Remark : The relation $\left(Tx, y\right) = \left(x, T^* y\right)$ can also be written as $\left(T^* x, y\right) = \left(x, Ty\right)$

$$\left(T^* x, y\right) = \overline{\left(y, T^* x\right)} = \overline{\left(Ty, x\right)} = \overline{\left(x, Ty\right)}$$

Hence $\left(T^* x, y\right) = \left(x, Ty\right) \quad \forall\; x, y \in H$

Theorem 2 : Let H be a given Hilbert space and T^* be adjoint of the operater T. Then T^* is a bounded linear transformation and T determines T^* uniquely.

Proof : First we prove that T^* is linear. Let vectors $y_1, y_2 \in H$ and α, β are scalars. Then for any vector $x \in H$, we have

$$\begin{aligned}
\left(x, T^*(\alpha y_1 + \beta y_2)\right) &= \left(Tx, \alpha y_1 + \beta y_2\right) \\
&= \left(Tx, \alpha y_1\right) + \left(Tx, \beta y_2\right) \\
&= \overline{\alpha}\left(Tx, y_1\right) + \overline{\beta}\left(Tx, y_2\right) \\
&= \overline{\alpha}\left(x, T^* y_1\right) + \overline{\beta}\left(x, T^* y_2\right) \\
&= \left(x, \alpha T^* y_1\right) + \left(x, \beta T^* y_2\right) \\
&= \left(x, \alpha T^* y_1 + \beta T^* y_2\right) \;\forall\; x \in H
\end{aligned}$$

$$\Rightarrow \qquad T^*\left(\alpha y_1 + \beta y_2\right) = \alpha T^* y_1 + \beta T^* y_2 \Rightarrow T^* \text{ is linear.}$$

Next we prove that T^* is bounded.

For any $x \in H$, let us consider,

$$\begin{aligned}
\|T^* x\|^2 &= \left(T^* x,\, T^* x\right) \\
&= \left|\left(T^* x,\, T^* x\right)\right| \qquad (\because \|T^*\|^2 \geq 0) \\
&= \left|\left(T\,T^* x,\, x\right)\right| \qquad \text{(by (1))}
\end{aligned}$$

$$\leq \|T\,T^*x\|\,\|x\| \qquad \text{(by Schwarz inequality)}$$

$$\leq \|T\|\,\|T^*x\|\,\|x\|$$

or $\quad \|T^*x\| \leq \|T\|\,\|x\| \qquad$ as $\qquad \|T^*x\| \leq \|T\|\,\|x\| \qquad \forall\ x \in H$

or $\quad \sup\left\{\dfrac{\|T^*x\|}{\|x\|} : x \neq 0\right\} \leq \|T\|$

$\Rightarrow \quad T^*$ is bounded since T is bounded

Lastly we show that T^* is unique. Let us assume that T^* is not unique. Let T_1 be another mapping of H into H with the property (1). Then $\forall\ x, y \in H$

$$(Tx, y) = (x, T_1 y)$$

and $\quad (Tx, y) = (x, T^*y)$

$\Rightarrow \quad (x, T_1\, y) = (x, T^*y) \qquad \forall\, x, y \in H$

$\Rightarrow \quad (x, T_1 y - T^*y) = 0 \qquad$ or $\qquad (x, (T_1 - T^*)y) = 0 \qquad \forall\, x \in H$

$\Rightarrow \quad (T_1 - T^*)\,y = 0 \qquad \forall\, y \in H$

$\Rightarrow \quad T_1 y = T^*y \qquad \forall\, y \in H$

$\Rightarrow \quad T_1 = T^*$

Remark : Using (1) we note that zero and identity operators are adjoint operators since $\forall\, x, y \in H$, we have

(i) $\quad (x, 0^*y) = (0x, y) = (0, y) = 0 = (x, 0) = (x, 0y)$

So from the uniqueness of the adjoint we get $0^* = 0$

(ii) $\quad (x, I^*y) = (I\,x, y) = (x, y) = (x, I\,y)$

So from uniqueness of the adjoint operator, $I^* = I$

7.3.4 Properties of Adjoint Operator :

Theorem 3 : Let H be a Hilbert space and $\beta(H)$ be the complex Banach space of all bounded linear transformations on H into H. Then the adjoint operation $T \to T^*$ on $\beta(H)$, where T is a bounded linear operator on H, has the following properties :

(a) $\quad (T + S)^* = T^* + S^* \qquad S$ be another bounded linear operator on H

(b) $\quad (\alpha\, T)^* = \overline{\alpha}\, T^*, \qquad \alpha$ being a scalar

(c) $(TS)^* = S^*T^*$

(d) $T^{***} = T$

(e) $\|T^*\| = \|T\|$

(f) $\|T^*T\| = \|T\|^2$

(g) $(T^*)^{-1} = (T^{-1})^*$ if T is invertible i.e. T is a non-singular operator on T.

Proof : (a) We have $\forall\ x, y \in H$

$$(x,(T+S)^*y) = ((T+S)x, y)$$

$$= (Tx + Sx, y) = (Tx, y) + (Sx, y)$$

$$= (x, T^*y) + (x, S^*y)$$

$$= (x, T^*y + S^*y) = (x, (T^* + S^*)y)$$

$\Rightarrow$ $(T+S)^* = T^* + S^*$ (by uniqueness of adjoint operator).

(b) $\forall\, x, y \in H$, we have

$$(x,(\alpha T)^*y) = ((\alpha T)x, y)$$

$$= \alpha(Tx, y)$$

$$= \alpha(x, T^*y)$$

$$= (x, \overline{\alpha}(T^*y))$$

$\Rightarrow$ $(\alpha T)^* = \overline{\alpha}\, T^*$ (by uniqueness property).

(c) $\forall\, x, y \in H$, we have

$$(x,(TS)^*y) = ((TS)x, y) = (T(Sx), y)$$

$$= (Sx, T^*y) = (x, S^*T^*y)$$

$\Rightarrow$ $(TS)^* = S^*T^*$

(d) $\forall\, x, y \in H$, we have

$$(x, T^{**}y) = (x, (T^*)^*y) = (T^*x, y)$$

$$= \overline{(y, T^*x)} = \overline{(Ty, x)}$$

$$= (x, Ty)$$

$$\Rightarrow \quad T^{**} = T$$

(e) $\forall\, x \in H$, we have

$\|T^*x\|^2 = (T^*x, T^*x) = (TT^*x, x)$ which is a real number and ≥ 0

$$= |(TT^*x, x)|$$

$$\leq \|T\,T^*x\|\,\|x\| \qquad \text{(by Schwarz inequality)}$$

$$\leq \|T\|\,\|T^*x\|\,\|x\|$$

or $\quad \|T^*x\| \leq \|T\|\,\|x\| \quad$ as $\quad \|T^*x\| \neq 0$

$$\therefore \quad \sup\left\{\frac{\|T^*x\|}{\|x\|} : x \neq 0\right\} \leq \|T\|$$

or $\quad \|T^*\| \leq \|T\| \qquad$ ...(9)

On replacing T by T^*, (9) gives

$$\|T^{**}\| \leq \|T^*\|$$

or $\quad \|T\| \leq \|T^*\| \quad$ by (d) $\qquad$ ...(10)

Hence (9) and (10) $\Rightarrow \|T^*\| = \|T\|$

(f) we have

$$\|T^*T\| = \sup\{\|T^*Tx\| : \|x\| \leq 1\}$$

$$= \sup\{\|T^*(Tx)\| : \|x\| \leq 1\}$$

$$\leq \sup\{\|T^*\|\,\|Tx\| : \|x\| \leq 1\}$$

$$\leq \|T^*\| \sup\{\|Tx\| : \|x\| \leq 1\}$$

$$\leq \|T^*\|\,\|T\| = \|T\|^2 \qquad \text{(by (e))}$$

$\therefore \quad \|T^*T\| \leq \|T\|^2 \qquad$ ...(11)

Also $\quad \|Tx\|^2 = (Tx, Tx) = (T^*Tx, x) = |T^*Tx, x| \quad \forall\, x \in H$

$$\leq \|T^*Tx\|\,\|x\| \qquad \text{(by Schwarz inequality)}$$

$$\leq \|T^*T\|\,\|x\|\,\|x\| \qquad \text{as} \qquad \|Tx\| \leq \|T\|\,\|x\|$$

or $\quad \sup\left\{\frac{\|Tx\|^2}{\|x\|^2} : x \neq 0\right\} \leq \|T^*T\|$

or $\quad \|T\|^2 \leq \|T^*T\| \qquad$ ...(12)

Thus (11) and (12) $\Rightarrow \|T^*T\| = \|T\|^2$

(g) If T is a non-singular operators on H and T^{-1} is inverse of T, then T^{-1} is also an operator on H. Also

$$TT^{-1} = I = T^{-1}T$$

$$\Rightarrow \quad (TT^{-1})^* = I^* = (T^{-1}T)^*$$

$$\Rightarrow \quad (T^{-1})^* T^* = I = T^*(T^{-1})^*$$

$\Rightarrow \quad T^*$ is invertible and so non-singular and also inverse of T^* is $(T^{-1})^*$.

Hence $(T^*)^{-1} = (T^{-1})^*$

7.4 Self-Adjoint Operator

7.4.1 Definition : A linear operater T on a Hilbert space H is known as self-adjoint or Hermition if $T^* = T$ or in other words, if T is self-adjoint

then $(Tx, y) = (x, T^*y) = (x, Ty)$

Zero operater and Identity operator are examples of self-adjoint operater.

7.4.2 Properties of Self-Adjoint Operator

Theorem 4 : An operator T on H is self-adjoint, then $(Tx, y) = (x, Ty)\ \forall\ x, y \in H$ and conversly.

Proof : If T^* is an adjoint operator of T on H, then by definition we have

$$(Tx, y) = (x, T^*y) \qquad \forall\, x, y \in H$$

If T is self-adjoint, there $T = T^*$. Therefore

$$(Tx, y) = (x, T^*y) = (x, Ty) \qquad \forall\, x, y \in H$$

conversly let us asume that

$$(Tx, y) = (x, Ty) \qquad \forall\, x, y \in H$$

But $\quad (Tx, y) = (x, T^*y)$

So $\quad (x, Ty) = (x, T^*y) \Rightarrow (x, (T - T^*)y) = 0 \qquad \forall\, x, y \in H$

$\therefore \quad x \neq 0 \ \therefore \ (T - T^*)y = 0 \quad \forall\, y \in H$

$\Rightarrow \qquad T = T^* \Rightarrow T$ is self adjoint.

Theorem 5 : Let T be a self-adjoint operator, then $T + T^*$ and T^*T are self-adjoint.

Proof : We have

$$(T + T^*)^* = T^* + T^{**} = T^* + T$$

$$= T + T^*$$

$\Rightarrow \qquad T + T^*$ is self-adjoint

Also $\quad (T^*T)^* = T^*(T^*)^* = T^*T \Rightarrow T^*T$ is self-adjoint

Theorem 6 : If T is an arbitrary operator on Hilbert space H, then $T = 0$ iff $(Tx, y) = 0$ $\forall\, x, y \in H$.

Proof : If $T = 0$, then $(Tx, y) = (0x, y) = 0 \ \forall\, x, y \in H$.

Conversly let $(Tx, y) = 0 \ \forall\, x, y \in H$.

Taking $y = Tx$, we get

$$(Tx, Tx) = 0 \qquad \forall\, x \in H$$

$\Rightarrow \qquad \|Tx\|^2 = 0 \qquad \forall\, x \in H$

$\Rightarrow \qquad Tx = 0 \qquad \forall\, x \in H$

$\Rightarrow \qquad T = 0$

Theorem 7 : If T is an operator on a Hilbert space H, then $(Tx, x) = 0 \ \forall\, x \in H$ iff $T = 0$.

Proof : Let $T = 0$, then $(Tx, x) = (0x, x) = (0, x) = 0 \ \forall\, x \in H$.

Conversely, let $(Tx, x) = 0 \ \forall\, x \in H$.

If $x, y \in H$ and α, β be any scalars, then we have

$$(T(\alpha x + \beta y), \alpha x + \beta y) = (\alpha Tx + \beta Ty, \alpha x + \beta y)$$

$$= (\alpha Tx, \alpha x + \beta y) + (\beta Ty, \alpha x + \beta y)$$

$$= \alpha (Tx, \alpha x + \beta y) + \beta (Ty, \alpha x + \beta y)$$

$$= \alpha (Tx, \alpha x) + \alpha (Tx, \beta y) + \beta (Ty, \alpha x) + \beta (Ty, \beta y)$$

$$= \alpha \bar{\alpha} (Tx, x) + \alpha \bar{\beta} (Tx, y) + \beta \bar{\alpha} (Ty, x) + \beta \bar{\beta} (Ty, y)$$

$$= |\alpha|^2 (Tx, x) + \alpha \bar{\beta} (Tx, y) + \beta \bar{\alpha} (Ty, x) + |\beta|^2 (Ty, y)$$

$$\Rightarrow (T(\alpha x + \beta y), \alpha x + \beta y) - |\alpha|^2 (Tx, x) - |\beta|^2 (Ty, y) = \alpha \bar{\beta} (Tx, y) + \beta \bar{\alpha} (Ty, x) \quad ...(13)$$

Since $(Tx, x) = 0 \ \forall x \in H$, therefore left-hand side of (13) is zero. Hence we get

$$\alpha \bar{\beta} (Tx, y) + \beta \bar{\alpha} (Ty, x) = 0 \quad \forall x, y \in H \text{ and } \alpha, \beta \text{ are any scalars.} \quad ...(14)$$

Taking $\alpha = \beta = 1$ and $\alpha = i$, $\beta = 1$ succesively in (14) we get

$$(Tx, y) + (Ty, x) = 0 \quad ...(15)$$

and $\quad i(Tx, y) - i(Ty, x) = 0$

or $\quad (Tx, y) - (Ty, x) = 0 \quad ...(16)$

Adding (15) and (16) we get

$$2(Tx, y) = 0 \quad \forall x, y \in H$$

or $\quad (Tx, y) = 0 \ \forall x, y \in H \Rightarrow T = 0 \quad$ (by Theorem 6)

Theorem 8 : An operator T on a complex Hilbert space H is self-adjoint iff (Tx, x) is real for all x.

Proof : Let T a self-adjoint operator on H i.e., $T = T^*$. Then for all $x \in H$, we have

$$(Tx, x) = (x, T^* x) = (x, Tx) = \overline{(Tx, x)}$$

Thus (Tx, x) is equal to its own conjugate and is therefore real.

Now suppose that (Tx, x) is real for every $x \in H$.

Since (Tx, x) is real for all $x \in H$, therefore we have

$$(Tx, x) = \overline{(Tx, x)} = \overline{(x, T^* x)} = (T^* x, x)$$

where T^* is adjoint of T which exists for every $x \in H$.

So $\quad (Tx, x) - (T^* x, x) = 0 \ \forall x \in H$

$\Rightarrow \quad (Tx - T^* x, x) = 0 \quad \forall x \in H$

$\Rightarrow \quad ((T - T^*) x, x) = 0 \quad \forall x \in H$

$\Rightarrow \quad T - T^* = 0 \quad$ or $\quad T = T^* \Rightarrow T$ is self-adjoint

Theorem 9 : Let A be the set of all self-adjoint operators in $\beta(H)$. Then A is a closed linear subspace of $\beta(H)$ and therefore A is a real Banach space containing the identity transformation.

Proof : First we note that A is non-empty, since 0 is a self-adjoint operator.

Let $T_1, T_2 \in A$. Then $T_1{*} = T_1$ and $T_2{*} = T_2$

Suppose that α , β be any two real numbes, Then

$$(\alpha T_1 + \beta T_2)* = (\alpha T_1)* + (\beta T_2)*$$

$$= \bar{\alpha} T_1 * + \bar{\beta} T_2 * = \alpha T_1 + \beta T_2$$

$\Rightarrow \quad \alpha T_1 + \beta T_2 \in A$.

Hence A is a real linear subspace of H.

Now we prove that A is closed subset of the Banach space $\beta(H)$,

Let $\{T_n\}$ be a sequence of self-adjoint operators converging to T. Now

$$\|T - T*\| = \|T - T_n + T_n - T_n * + T_n * - T*\|$$

$$\le \|T - T_n\| + \|T_n - T_n *\| + \|T_n * - T*\|$$

$$\le \|T_n - T\| + \|0\| + \|(T_n - T)*\| \qquad (\because T_n \in A \Rightarrow T_n * = T_n)$$

$$\le 2\|T_n - T\| \quad (\because \|T*\| = \|T\|)$$

$$\to 0 \text{ as } T_n \to T$$

$\therefore \quad \|T - T*\| = 0 \Rightarrow T = T*$

$\Rightarrow \quad T$ is self-adjoint $\Rightarrow T \in A$.

$\Rightarrow \quad A$ is a closed subspace of complete Banach space $\beta(H)$.

$\Rightarrow \quad A$ is also complete and hence is a real Banach space.

Also $\quad I* = I \Rightarrow$ the identity operator $I \in A$

7.5 Positive Operators

Since (Tx, x) is real for self-adjoint operators, therefore we can introduce the order relation among them and define positive operator by considering the real values which the self-adjoint operator take.

7.5.1 Definition : A self-adjoint operator T on H is said to be positive if $T \ge 0$ in the order relation.

This means $(Tx, x) \ge 0 \;\; \forall\, x \in H$

From the definition, we have the following properties.

(a) The identity operator I and the zero operator 0 are positive operators, since

$$(I\,x, x) = (x, x) = \|x\|^2 \geq 0$$

and $(0x, x) = (0, x) = 0$

(b) For any arbitrary operators T on H, TT^* and T^*T are positive operators since TT^* and T^*T are self adjoint and

$$(TT^*x, x) = (T^*x, T^*x) = \|T^*x\|^2 \geq 0$$

7.6 Normal Operators

7.6.1 Definition : An operator T on a Hilbert space H is known to be Normal if it commutes with its adjoint i.e. if $TT^* = T^*T$

From the defintiion it is obvious that

(a) Every self-adjoin operators is normal, since

$$T = T^* \Rightarrow TT^* = T^*T$$

(b) Both zero and identity operators are normal operators.

(c) A normal operator is non-necessarily self-adjoint.

7.6.2 Properties of Normal Operators :

Theorem 10 : If T_1 and T_2 are normal operators on H with the property that either commutes with adjoint of the other, then $T_1 + T_2$ and $T_1\,T_2$ are normal.

Proof : Since T_1 and T_2 are normal, therefore

$$T_1\,T_1^* = T_1^*\,T_1 \text{ and } T_2\,T_2^* = T_2^*\,T_2 \qquad ...(17)$$

From hypothesis, we have

$$T_1\,T_2^* = T_2^*\,T_1 \text{ and } T_2\,T_1^* = T_1^*\,T_2 \qquad ...(18)$$

Now $(T_1 + T_2)(T_1 + T_2)^* = (T_1 + T_2)(T_1^* + T_2^*)$

$$= T_1\,T_1^* + T_1\,T_2^* + T_2\,T_1^* + T_2\,T_2^*$$

$$= T_1^*\,T_1 + T_2^*\,T_1 + T_1^*\,T_2 + T_2^*\,T_2 \qquad \text{(from (17) and (18))}$$

$$= T_1^*(T_1 + T_2) + T_2^*(T_1 + T_2)$$

$$= (T_1^* + T_2^*)(T_1 + T_2)$$

$$= (T_1 + T_2)^*(T_1 + T_2)$$

$\Rightarrow$ $(T_1 + T_2)$ is normal.

Also $(T_1 T_2)(T_1 T_2)^* = (T_1 T_2)(T_2^* T_1^*)$

$$= T_1 (T_2 T_2^*) T_1^*$$

$$= (T_1 T_2^*)(T_2 T_1^*)$$

$$= (T_2^* T_1)(T_1^* T_2)$$

$$= T_2^* (T_1 T_1^*) T_2$$

$$= T_2^* (T_1^* T_1) T_2$$

$$= (T_2^* T_1^*)(T_1 T_2)$$

$$= (T_1 T_2)^* (T_1 T_2)$$

$\Rightarrow$ $T_1 T_2$ is normal

Theorem 11 : An operator T on a Hilbert space H is normal iff $\|T^* x\| = \|T x\| \;\; \forall\, x \in H$.

Proof : Let T is normal, Then

$$TT^* = T^*T \Leftrightarrow TT^* - T^*T = 0$$

$$\Leftrightarrow \left((TT^* - T^*T)x, x\right) = 0 \qquad \forall\, x \in H$$

$$\Leftrightarrow (TT^*x, x) = (T^*Tx, x)$$

$$\Leftrightarrow (T^*x, T^*x) = (Tx, T^{**}x)$$

$$\Leftrightarrow (T^*x, T^*x) = (Tx, Tx) \qquad \forall\, x \in H$$

$$\Leftrightarrow \|T^*x\|^2 = \|Tx\|^2$$

$$\Leftrightarrow \|T^*x\| = \|Tx\| \qquad \forall\, x \in H$$

Thoerem 12 : If T is a Normal Operator on H, then $\|T^2\| = \|T\|^2$

Proof : We have

$$\|T^2 x\| = \|TTx\|$$

$$= \|T(Tx)\|$$

$$= \|T^*(Tx)\| \qquad \forall\, x \in H \qquad [\because T \text{ is normal } \therefore \|Tx\| = \|T^*x\| \;\; \forall\, x \in H]$$

Hence $\|T^2 x\| = \|T^*(Tx)\|$ ...(19)

Also $\|T^2\| = \sup\{\|T^2 x\| : \|x\| \le 1\}$

$$= \sup\{\|T^* Tx\| : \|x\| \le 1\}$$

$$= \|T^* T\| = \|T\|^2$$

which completes the proof of the theorem.

Theorem 13 : An arbitrary operator T on a Hilbert space H can be uniquely expressed as $T = T_1 + i\,T_2$ and $T^* = T_1 - i\,T_2$, where T_1 and T_2 are self-adjoint operators.

Proof : Let T^* be the adjoint of T. Define

$$T_1 = \frac{1}{2}(T + T^*) \text{ and } T_2 = \frac{1}{2i}(T - T^*) \qquad ...(20)$$

Then we have

$$T = T_1 + i\,T_2 \text{ and } T^* = T_1 - i\,T_2$$

Again $T_1^* = \left[\frac{1}{2}(T + T^*)\right]^* = \frac{1}{2}(T + T^*)^* = \frac{1}{2}(T^* + T^{**})$

$$= \frac{1}{2}(T^* + T) \qquad (\because T^{**} = T)$$

$$= \frac{1}{2}(T + T^*) = T_1 \Rightarrow T_1 \text{ is self-adjoint.}$$

Also $T_2^* = \left[\frac{1}{2i}(T - T^*)\right]^* = -\frac{1}{2i}(T^* - T)$

$$= \frac{1}{2i}(T - T^*) = T_2 \Rightarrow T_2 \text{ is self-adjoint.}$$

Thus an arbitrary operator T can be expressed in the form (20) where T_1 and T_2 are self-adjoint operators. Next we show that this type of expression is unique. Let the expression is non-unique i.e. let $T = S_1 + i\,S_2$ where S_1 and S_2 and self-adjoint operators on H.

Then $T^* = (S_1 + i\,S_2)^* = S_1^* + \bar{i}\,S_2^* = S_1^* - i\,S_2^*$

Thus $S_1 = \frac{1}{2}(T + T^*) = T_1$ and $S_2 = \frac{1}{2i}(T - T^*) = T_2$

Hence the expression (20) for $T \in \beta(H)$ is unique.

Remark : If T is expressed as $T_1 + i\,T_2$ and $T^* = T_1 - i\,T_2$ where T_1 and T_2 are self-adjoint operators on H then T_1 is called the real part of T and T_2 is called the imaginary part of T.

Theorem 14 : If T is an operator on a Hilbert space H, then T is normal iff its real and imaginary parts commute.

Proof : Let $T = T_1 + i\,T_2$ where T_1 and T_2 are self-adjoint operators on H.

We have $T^* = (T_1 + i\,T_2)^* = T_1{}^* + \bar{i}\,T_2{}^* = T_1{}^* - i\,T_2{}^* = T_1 - i\,T_2$

Now $\quad TT^* = (T_1 + i\,T_2)\,(T_1 - i\,T_2)$

$$= T_1^2 + i(T_2\,T_1 - T_1\,T_2) + T_2^2 \qquad ...(22)$$

and $\quad T^*T = (T_1 - i\,T_2)\,(T_1 + i\,T_2)$

$$= T_1^2 + i(T_1\,T_2 - T_2\,T_1) + T_2^2 \qquad ...(23)$$

If T is normal, then $TT^* = T^*T$

(22) and (23) $\Rightarrow T_1^2 + i(T_2\,T_1 - T_1\,T_2) + T_2^2$

$$= T_1^2 + i(T_1\,T_2 - T_2\,T_1) + T_2^2$$

$\Rightarrow \quad 2i(T_1\,T_2 - T_2\,T_1) = 0$

$\Rightarrow \quad T_1\,T_2 = T_2\,T_1 \Rightarrow$ Real and imaginary parts commute

Conversely if $T_1\,T_2 = T_2\,T_1$, then (22) and (23) gives

$TT^* = T^*\,T \Rightarrow T$ is normal.

Theorem 15 : Show that the set of all normal operators on a Hilbert space H is a closed subset of $\beta(H)$ which contains the set of all self-adjoint operators and is closed under scalar multiplication.

Proof : Let S be the set of all normal operators on a Hilbert space H. We first show that S is closed subset of $\beta(H)$. Let T be any limit point of S. We have to prove that $T \in S$. Since T is a limit point of S, therefore $\exists$ a sequence $\{T_n\}$ of distinct points of S s.t. $T_n \to T$ as $n \to \infty$.

Now $\quad \|T_n{}^* - T^*\| = \|(T_n - T^*)\| = \|T_n - T\| \to 0$ as $n \to \infty$

$\therefore \quad T_n{}^* \to T^*$ as $n \to \infty \qquad ...(24)$

Also $\quad \|TT^* - T^*T\| = \|TT^* - T_n\,T_n{}^* + T_n\,T_n{}^* - T^*T\|$

$$\leq \|TT^* - T_n\,T_n{}^*\| + \|T_n\,T_n{}^* - T^*T\|$$

$$\leq \|TT^* - T_n\,T_n{}^*\| + \|T_n\,T_n{}^* - T_n{}^*\,T_n + T_n{}^*\,T_n - T^*T\|$$

$$\leq \|TT^* - T_n\,T_n{}^*\| + \|T_n\,T_n{}^* - T_n{}^*\,T_n\| + \|T_n{}^*\,T_n - T^*\,T\|$$

$$\leq \|TT^* - T_n\,T_n{}^*\| + \|T_n{}^*\,T_n - T^*\,T\| \qquad ...(25)$$

$$(\because T_n \in S \Rightarrow T_n \text{ is normal} \Rightarrow T_n T_n{}^* = T_n{}^* T_n)$$

Since $T_n \to T$ and $T_n{}^* \to T^*$, the right hand side of (25) tends to zero which implies that

$$\|TT^* - T^*T\| \to 0 \Rightarrow TT^* = T^*T$$

$$\Rightarrow T \text{ is normal} \Rightarrow T \in S$$

This prove that S is a closed subset of $\beta(H)$.

Again every self-adjoint operator is normal. Therefore S is a closed subset of $\beta(H)$ containing the set of all self-adjoint operators.

Finally we prove that S is closed for scalar multiplication i.e. if α is a scalar and $T \in S$, then $\alpha T \in S$ or if T is normal then αT is also normal for any scalar α.

Since T is normal, therefore $TT^* = T^*T$

Now $(\alpha T)(\alpha T)^* = (\alpha T)(\bar{\alpha} T^*)$

$$= \alpha \bar{\alpha} T T^*$$

$$= \bar{\alpha} \alpha \ T^* T = (\bar{\alpha} T^*)(\alpha T)$$

$$= (\alpha T)^* (\alpha T)$$

$\Rightarrow \alpha T$ is normal.

which complete the proof of the theorem.

7.7 Unitary Operators

7.7.1 Definition 1 : An operator U on a Hilbert space H is said to be unitary if $UU^* = U^*U = I$

From the definition it is obvious that

(i) If U is unitary, then it is normal.

(ii) $U^* = U^{-1}$

7.2.2 Definition 2 : An operator T on H is said to be Isometric if $\|Tx - Ty\| = \|x - y\| \ \forall x, y \in H$

Since T is linear, the condition is equivalent to $\|Tx\| = \|x\| \ \forall \ x \in H$

Now we prove a result contained in

Theorem 16 : If T is an operator on a Hilbert space H then the following conditions are equivalent :

(a) $T^*T = I$

(b) $(Tx, Ty) = (x, y) \quad \forall x, y \in H$

(c) $\|T x\| = \|x\| \quad \forall x \in H$

Proof : (a) $\Rightarrow$ (b) : Given that $T^*T = I$

Now $(Tx, Ty) = (x, T^*Ty) = (x, Iy) = (x, y) \quad \forall\, x, y \in H$

(b) $\Rightarrow$ (c) : Given that $(Tx, Ty) = (x, y) \;\; \forall\, x, y \in H$

Taking $x = y$, we get

$$(Tx, Tx) = (x, x)$$

$$\Rightarrow \quad \|Tx\|^2 = \|x\|^2$$

$$\Rightarrow \quad \|Tx\| = \|x\| \qquad \forall\, x \in H$$

(c) $\Rightarrow$ (a) : By (c) we have

$$\|Tx\| = \|x\| \qquad \forall\, x \in H$$

Now $\|Tx\| = \|x\|$

$$\Rightarrow \quad \|Tx\|^2 = \|x^2\|$$

$$\Rightarrow \quad (Tx, Tx) = (x, x) \qquad \forall\, x \in H$$

$$\Rightarrow \quad (Tx, T^{**}x) = (x, x)$$

$$\Rightarrow \quad (T^*Tx, x) = (I\,x, x) \quad \forall\, x \in H \qquad (\because\ T^{**} = T)$$

$$\Rightarrow \quad ((T^*T - I)x, x) = 0 \quad \forall\, x \in H$$

$$\Rightarrow \quad T^*T = I$$

7.7.3 Properties of Unitary Operator

Theorem 17 : An operator T on a Hilbert space H is unitary iff it is an isometric isomorphism of H onto itself.

Proof : Let T be unitary. Then $T^*T = TT^* = I$.

Therefore T is invertible and so T one-one and onto.

Also $\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x)$

$$= (Ix, x) = (x, x)$$

$$= \|x\|^2$$

$$\Rightarrow \quad \|Tx\| = \|x\| \qquad \forall\, x \in H$$

Thus T preserves the norm and so T is an isometric isomorphism of H onto itself.

Conversely suppose that T be an isometric isomorphism of H onto itself. Then T is one-one and onto. Therefore T is invertible i.e. T^{-1} exists.

$$\Rightarrow \quad T\,T^{-1} = T^{-1}T = I \qquad \text{...(26)}$$

Again T preserves the norm, therefore

$$\|Tx\| = \|x\| \qquad \forall\, x \in H$$

$$\Rightarrow \quad T^*T = I$$

$$\Rightarrow \quad (T^*T)T^{-1} = I\,T^{-1}$$

$$\Rightarrow \quad T^*(T\,T^{-1}) = T^{-1} \Rightarrow T^*I = T^{-1}$$

or $\quad T^* = T^{-1} \Rightarrow T\,T^* = TT^{-1} = I$

In a similar manner

$$T^*T = T^{-1} = I$$

Hence $T^*T = TT^* = I \Rightarrow T$ is unitary.

Remark : If T is an unitary operator on H, then $\|T\| = 1$.

Also $\|Tx\| = \|x\|$ so that

$$\|Tx\| = \sup\,\{\|Tx\| : \|x\| \le 1\} = \sup\,\{\|x\| : \|x\| \le 1\}$$

Self-Learning Exercise

In the following questions write T for true and F for false :

1. The conjugate and adjoint operator operate, on functionals in $H^*(T/F)$.

2. If T be an operater on a Hilbert space H, then

   $$(Tx, y) = (x, T^*y) \;\; \forall\; x, y \in H.$$

   where T^* is the adjoint operator of $T.(T/F)$.

3. $\|TT^*\| = \|T\|^2 \qquad (T/F)$

4. O and I are self-adjoint operators (T/F)

5. A normal operator is always a sefl-adjoint (T/F)

6. If N_1 and N_2 are normal operators then $N_1 + N_2$ is also a normal operator (T/F)

7. If N_1 and N_2 are normal operators then $N_1 N_2$ is a normal operator (T/F).

8. If T_1 and T_2 are self-adjoint operators, then their product $T_1 T_2$ is self-adjoint (T/F).

9. If T is an operator on H s.t. $(Tx, x) = 0 \ \forall x \in H$ then $T = 0$ (T/F).

10. If T is a Normal operator and α is a scalar then αT is normal (T/F).

11. If H be an inner product space which is not complete, then H^* necessarity exists (T/F).

Fill in the blanks :

12. Identity is an .... operator

13. $(ST)^* =$ ....

14. If T is a positive operator on Hilbert space H, then $I + T$ is ....

7.8 Summary

In this unit you studied different type of operators on Hilbert spaces and various properties associated with these operators.

7.9 Answers to Self-Learning Exercise

1.	F	2.	T	3.	T	4.	T
5.	F	6.	F	7.	F	8.	F
9.	T	10.	T	11.	F		
12.	Self-adjoint	13.	$T^* S^*$	14.	Non-singular		

7.10 Exercises

1. Define an adjoint operator on a Hilbert space H and give an example.

2. Show that the adjoint operation is one-one onto as a mapping $\beta(H)$ into itself.

3. Prove that every scalar multiple of self-adjoint operator is also normal.

4. Let H be a Hilbert space and T, S be the set of bounded operators on H. Prove that if

   (i) S and T are self-adjoint and $ST = TS$, then ST is self-adjoint.

   (ii) S and T are normal and $ST^* = T^* S$, then ST is normal.

5. If T is an arbitrary operator on a Hilbert space H and α, β are scalars s.t. $|\alpha| = |\beta|$, then show that $\alpha T + \beta T^*$ is normal.

6. If T is a normal operator on a Hilbert space H and λ is a scalar, then show that $T - \lambda I$ is normal.

7. Show that an operator T on a Hilbert space H is unitary iff $T(\{e_i\})$ is a complete orthonormal set whenever $\{e_i\}$ is.

8. Show that the set of unitary operators on a Hilbert space H, forms a multiplicative group.

9. If T is a linear operator on a Hilbert space H, then T is unitary iff adjoint of T exists and $TT^* = T^*T = I$.

10. If T is self-adjoint, any operator S unitarily equivalent to T is also self-adjoint.

11. Let T be normal and A and B be self-adjoint operators s.t. $T = A + iB$. Then prove that $AB = BA$.

□□□

Unit - 8
Projections on a Hilbert Space and Spectral Theory

Structure of the Unit

8.1 Objectives

In this unit first we study projection on a Hilbert space H and properties of the projection operator on H. We also study spectral theory of operators on finite dimensional Hilbert spaces.

8.2 Introduction

The aim of this unit is to study the projection on a Hilbert space. Invariance, reducilility and orthogonal projections will also be studied. Next we shall study to some extent in detail the relation between linear operators on a finite dimensional Hilbert space and matrices as a preliminary step towards the study of spectral theory of operators on finit dimensional Hilbert spaces. After a brief study of the spectrum of an operator and its properties, we shall establish the spectral theorem for normal operators on a finite dimensional Hilbert space and indicate the spectral theorems for self-adjoint, positive and unitary operators.

8.3 Projections

We have already defined a projection on a Banach space B and Hilbert space H i.e. it is an idempotent linear operator P on B s.t. it is a continuous linear transformation from B (or H) into itself with the property $P^2 = P$. It has also been shown that B (or H) $= M \oplus N$ where

$$M = \{Px : x \in B\} \text{ and } N = \{x \in B : Px = \mathbf{0}\}$$

M is called the range and N, the null space of P.

8.3.1 Definition :

Perpendicular Projection : A projection P on a Hilbert space H is known as a perpendicular projection on H if the range M and null space N of P are orthogonal i.e., $M \perp N$. Thus by projection P on H we mean a perpendicular projection on H.

8.3.2 Important Results :

Thoerem 1 : If P is a projection on a Hilbert space H with range M and null space N, then $M \perp N$ iff P is self adjoint, and in this case $N = M^{\perp}$.

Proof : By definition we have $P^2 = P$ and $H = M \oplus N$ Let $M \perp N$. Then we prove that P is self-adjoint. By projection theorem each vector $z \in H$ can be uniquely represented as $z = x + y$, where $x \in M$, $y \in N$ s.t.

$$Pz = P(x+y) = x \text{ and } Py = 0$$

Since $M \perp N$, we have $(x, y) = 0$ ...(1)

Using (1), we get

$$(Pz, z) = (x, z) = (x, x+y) = (x, x) + (x, y) = (x, x) \qquad ...(2)$$

and $$(P^*z, z) = (z, Pz) = (z, x) = (x+y, x) = (x, x) + (y, x) = (x, x) \qquad ...(3)$$

(2) and (3)

$\Rightarrow \quad (Pz, z) = (P^*z, z) \quad \forall z \in H$

$\Rightarrow \quad ((P - P^*)z, z) = 0 \quad \forall z \in H$

$\Rightarrow \quad P - P^* = 0$

$\Rightarrow \quad P^* = P \Rightarrow P$ is self-adjoint.

Conversly suppose that P is self-adjoint i.e., $P^* = P$.

Now, let $x \in M$ and $y \in N$. Then

$$(x, y) = (Px, y) \qquad (P \text{ being projection on } H \text{ and } x \in M \Rightarrow Px = x)$$

$$= (x, P^*y) = (x, Py)$$

$= (x, \mathbf{0}) = 0$ as $y \in N \quad \therefore Py = 0$

$\therefore \qquad (x, y) = 0 \Rightarrow M \perp N$

Lastly we prove that $N = M^{\perp}$ where $M \perp N$

For any $x \in N$ and $N \perp M \Rightarrow x \perp M \Rightarrow x \in M^{\perp} \Rightarrow N \subset M^{\perp}$.

Taking N to be proper closed linear subspace of Hilbert space $M^{\perp}$ i.e. $N \neq M^{\perp}$. So $\exists$ a non-zero vector $z_0 \in M^{\perp}$ s.t. $z_0 \perp N$. Also $z_0 \perp M$.

$\therefore \qquad z_0 \perp M$ and $z_0 \perp N \Rightarrow z_0 \perp M \oplus N = H \Rightarrow z_0 \perp H$

$\Rightarrow \qquad z_0 = \mathbf{0}$ since only zero vector is orthogonal to whole space H.

This contradicts that z_0 is a non zero vector. Hence N cannot be proper subset of $M^{\perp}$ and the only possibility is that $N = M^{\perp}$.

Theorem 2 : If P is the projection on a closed linear subspace M of a Hilbert space H, then

(i) P is the projection on M of $H \Leftrightarrow I - P$ is the projection on $M^{\perp}$

(ii) $x \in M \Leftrightarrow Px = x \Leftrightarrow \|Px\| = \|x\|$

Proof : (i) P is the projection on $H \Leftrightarrow P^2 = P$ and $P^* = P$ ...(4)

Therefore $(I - P)^* = I^* - P^* = I - P$

and $\qquad (I - P)^2 = (I - P)(I - P) = I^2 - IP - PI + P^2 = I - P - P + P = IP$

$\Rightarrow \qquad (I - P)$ is also a projection on H.

Now we prove that if P is defined on M, then $(I - P)$ is defined on $M^{\perp}$. For this let N be the range of $(I - P)$. Then

$x \in N \Rightarrow (I - P)x = x =$ or $x - Px = x \Rightarrow Px = \mathbf{0}$

$\Rightarrow \qquad x \in$ Null space of $P \Rightarrow x \in M^{\perp}$ as $M^{\perp}$ being coincident with null space

$\therefore \qquad N \subset M^{\perp}$

Also $\qquad x \in M^{\perp} \Rightarrow Px = \underline{0} \Rightarrow x - Px = x \Rightarrow (I - P)x = x$

$\Rightarrow x \in$ range of $(I - P) \Rightarrow x \in N$

$\therefore \qquad M^{\perp} \subset N$ ...(6)

(5) and (6)

$\Rightarrow \qquad N = M^{\perp}$

$\Rightarrow$ If P is projection on a closed linear subspace M of H, then $(I-P)$ is projection on $M^{\perp}$.

Conversly : If $(I-P)$ is a projection on $M^{\perp}$ then $I-(I-P)=P$ is the projection on $(M^{\perp})^{\perp}=M^{\perp\perp}=M$, since M is closed.

(ii) If $Px=x$, then Px is the range of P i.e. $x\in$ range of P i.e. $x\in M$.

Conversly, if $x\in M$, then assuming that $Px=y$, we prove that $y=x$.

Now $Px=y\Rightarrow P(Px)=Py\Rightarrow P^2x=Py\Rightarrow Px=Py$ $\quad (\because P^2=P)$

$\Rightarrow P(x-y)=\mathbf{0}\Rightarrow (x-y)\in$ null space of P

$\Rightarrow (x-y)\in M^{\perp}$, as $M^{\perp}$ is the null space of P

$\Rightarrow x-y=z(say)$, where $z\in M^{\perp}$

Now $y=Px\Rightarrow y\in$ range of P i.e. M

Thus $x=y+z$ where $y\in M$ and $z\in M^{\perp}$

Since $x\in M$, we can write $x=x+0$

Since $H=M\oplus M^{\perp}$, we have $z=0$ so that $x=y$

Again If $Px=x\Rightarrow \|Px\|=\|x\|$

Conversly if $\|Px\|=\|x\|$, then we have

$\|x\|^2=\|Px+(I-P)x\|^2$ where $Px\in M$ and $(I-P)x\in M^{\perp}$

as such Px and $(I-P)x$ are orthogonal vecotrs.

Using Phythagorean theorem, we get

$$\|x\|^2=\|Px\|^2+\|(I-P)x\|^2$$

$$=\|x\|^2+\|(I-P)x\|^2$$

$$\Rightarrow \|(I-P)x\|^2=0$$

$$\Rightarrow x-Px=\mathbf{0}\Rightarrow Px=x$$

Thus $x\in M\Leftrightarrow Px=x\Leftrightarrow \|Px\|=\|x\|$

Theorem 3 : If P is a projecton on a Hilbert space H, then prove that

(i) $\|Px\| \le \|x\| \quad \forall x \in H$ (ii) $\|P\| \le I$

(iii) P is a positive operator (iv) $0 \le P \le I$

Proof : (i) We have $\|Px\|^2 + \|(I-P)x\|^2 = \|x\|^2$

$\Rightarrow \|Px\|^2 \le \|x\|^2$ as $\Rightarrow \|(I-P)x\|^2 \ge 0$

$\Rightarrow \|Px\| \le \|x\| \quad \forall x \in H$

(ii) by (i), $\|Px\| \le \|x\|, \quad \forall x \in H$

$\Rightarrow Sup\{\|Px\| : \|x\| \le \perp\} \le 1 \quad \forall x \in H$

$\Rightarrow \|P\| \le \perp, \ x \in H$ being arbitrary.

(iii) For any vector $x \in H$, P being projection on H i.e. $P^* = P, P^2 = P$ and $Px \in$ range of P so that $P(Px) = Px$

We have

$$(Px, P) = (PPx, x) = (Px, P^* x) = (Px, Px) = \|Px\|^2 \ge 0$$

$\Rightarrow P \ge 0$ i.e. P is a positive operator

(iv) Since P and $I - P$ are projections on H, therefore $P \ge 0$ and $I - P \ge 0$ or $P \ge I$. Thus $0 \le P \le I$.

8.4 Invariance and Reducibility

8.4.1 Definitions : Let T be a linear operater on a Hilbert space H, then M is invariant under T if $x \in M \Rightarrow Tx \in M$ i.e. $T(M) \subseteq M$

Obviously M is invariant under Zero operator and every closed subspace is invariant under identity operater I.

Now M being a closed subspace of H, M itself is a Hilhert space so that T may be regarded as operater on M also.

If T on H induces an operator T_M on M and s.t. $T_M(x) = T(x) \forall x \in M$,

then T_M is known as **restriction** of T on M.

we know that H = $H = M \oplus M^{\perp}$

where M is a closed linear subspace of Hilbert space H. Then T is said to be **reduced** by M if both M and $M^{\perp}$ are invasiant under T. We sometimes also say that M reduces T instead of saying that T is reduced by M.

8.4.2 Properties

Theorem 4 : A closed linear subspace M of a Hilbert space H is invariant under an operater $T \Leftrightarrow M^{\perp}$ is invariant under T^*.

Proof : Let M be invariant under T i.e. $x \in M \Rightarrow Tx \in M$.

Suppose that $y \in M^{\perp} \Rightarrow y$ is orthogenal to every vector in M ....(7)

$\Rightarrow y$ is orthogonal to Tx as $Tx \in M$ by **(7)**

$\Rightarrow (Tx, y) = 0$

$\Rightarrow (x, T^* y) = 0, \quad x \in M$

$\Rightarrow T^* y$ is orthogenal to every vector $x \in M$

$\Rightarrow T^* y \in M^{\perp}$

$\Rightarrow M^{\perp}$ is invariant under T^*

Convessly, suppose that $M^{\perp}$ be invariant under T^*.

Since $M^{\perp}$ is a closed linear subspace of H and is invariant under T^*, therefore by the theorem $M^{\perp\perp} = M$ is invariant under $T^{**} = T$.

Theorem 5 : A closed linear subspace M of a Hilbert Sapce H reduces an operator $T \Leftrightarrow M$ is invariant under both T and T*.

Proof : M reduces T $\Rightarrow M$ and $M^{\perp}$ both are invariant under T. But $M^{\perp\perp} = M$ is invariant under T*.

Hence M is invariant both under T and T*.

Conversly, If M is invariant under both T and T*, then M is invariant under T and $M^{\perp}$ is invariant under $T^{**} = T$

$\Rightarrow$ Both M and $M^{\perp}$ are invariant under T.

$\Rightarrow$ M reduces T .

Theorem 6 : If P is the projection on a closed linear subspace M of a Hilbert space H, then

(i) M is invariant under an operator T $\Leftrightarrow$ TP = PTP

(ii) M reduces an operator T $\Leftrightarrow$ TP = PT

Proof (i) Let M be invariant under T and x be an arbtrary vector of H. Then $Px \in M$ (range of P)

$\Rightarrow TPx \in M, M$ is invariant under T.

Since P is a projection and M is the range, therefore

$TPx \in M \Rightarrow P$ maps TPx into itself.

Hence $PTPx = TPx \; \forall x \in H \;\; \Rightarrow PTP = TP$

Conversly : Let PTP = TP

Since P is a projection with range M and an $x \in M$, therefore $Px = x \Rightarrow TPx = Tx$

Using hypothens we have $PTPx = TPx = Tx$

Since P maps elements of M into the same element $P(TPx) = TPx$ means $TPx \in M \Rightarrow Tx = M$. Hence $x \in M \Rightarrow Tx \in M$, therefore M is invariant under T.

(ii) M reduces T $\Rightarrow$ M is invariant under both T and T^*

$$\Rightarrow TP = PTP \text{ and } T^*P = PT^*P \text{ by case (i)}$$

$$\Rightarrow TP = PTP \text{ and } (T^*P)^* = (PT^*P)^*$$

$$\Rightarrow TP = PTP \text{ and } P^*T^{**} = P^*T^{**}P^*$$

$$\Rightarrow TP = PTP \text{ and } PT = PTP \quad (\because T^{**} = T \text{ and } P^* = P)$$

$$\Rightarrow TP = PT$$

Conversly : Suppose that $TP = PT \Rightarrow PTP = PPT = P^2T = PT \quad (\because P^2 = P)$

Also $TPP = PTP \Rightarrow TP^2 = PTP \Rightarrow TP = PTP$

$\because TP = PT \Rightarrow PTP = PT$ and $TP = PTP$

$\Rightarrow M$ reduces T

8.5 Orthogeral Projection

8.5.1 Definition : Two perpendicular projections P and Q on a Hilbert space H are known as orthogonal if PQ = O. In other words P and Q and Q are orthogoral iff their ranges M and N are orthogenal

8.5.2 Important Result :

Theorem : If P and Q are projections on closed linear subspaces M and N of a Hilbert space H, then $M \perp N \Leftrightarrow PQ = O \Leftrightarrow QP = 0$

Proof : P and Q are projections on H $\Rightarrow P^* = P, Q^* = Q$

Also $O^* = O$ and $I^* = I$.

$\therefore PQ = O \Leftrightarrow (PQ)^* = O^* \Leftrightarrow Q^*P^* = O^* \Leftrightarrow QP = 0$

Now we prove that $M \perp N \Leftrightarrow PQ = O$

For any vector $y \in N$ and $M \perp N \Rightarrow y$ is orthogonal to every vector in M

i.e. $y \in N \Rightarrow y \in M^{\perp} \Rightarrow N \subset M^{\perp}$

and for any vector $z \in H$ and Q is projection on H

$\Rightarrow Qz \in N$ (the range of Q) whereas $N = M^{\perp}$

$\Rightarrow Qz \in M^{\perp}$ (the null space of P)

$\Rightarrow P(Qz) = \mathbf{0} \Rightarrow PQz = O \, \forall \, z \in H$

$\Rightarrow PQ = 0$

Conversly : If PQ = O and $x \in M, y \in N$, then

$Px = x$ ($\because$ M being range of P) and $Qy = y$ ($\because$ N is range of Q)

$\therefore (x, y) = (Px, Qy) = (x, P^*Qy) = (x, PQy) (\because P^* = P)$

$= (x, Oy) = (x, \mathbf{0}) = 0$ where $x \in M, y \in N$

$\Rightarrow M \perp N$ i.e. M and N are orthogonal

8.6 Eigenvalue and Eigenvector

8.6.1 Definition : Let T be an operator on a Hilbert space H. Then a scalar λ is called the eignvalue (or characteristic or proper or latent or spectral value) of T if there exists a non-zero vector $x \in H$ s.t. $Tx = \lambda x$

It λ is an eigenvalue of T, then the non zero vector $x \in H$ s.t. $Tx = \lambda x$ is called the eignvector (characteristic or proper or latent or spectral vector) of T.

Each eigenvalue has one or more eigenvector whereas each eigenvector corresponds one eigenvalue. If H has no non-zero vectors, then T cannot have any eigenvector and hence the whole theory reduces to triviality. We therefore develop the spectral theory on the assumption that H $\neq \{0\}$.

The set of all eigenvalues of T is known as the spetrum of T and denoted by $\sigma(T)$.

8.6.2 Properties of eigenvalue and eigenvector

From the definition of eigenvalue and eigenvector, we have the following properties:

Theorem 8 : If x is an eigenvector of T corresponding to eigenvalue λ, and α is a non-zero scalar, then α is also an eigenvector of T corresponding to same eigenvalue.

Proof : Since x is an eigenvector of T corresponding to eigenvalue λ therefore $x \neq 0$ and $Tx = \lambda x$.

If $\alpha \neq 0$ then $\alpha x \neq 0$. Also T$(\alpha x) = \alpha Tx = \alpha \lambda x = \lambda(\alpha x)$

Hence αx is also eigenvector of T coresponding to same eigenvalue λ.

This property tells us that corresponding to a single eigenvalue there may correspond more than one eigenvector.

Theorem 9 : If *x* is an eigenvector of T, then *x* cannot correspond more than one eigenvalue of T.

Proof : If possible, let λ_1 and λ_2 he two distinct eigenvalues of T for eigenvector *x*. Then $Tx = \lambda_1 x$ and $Tx = \lambda_2 x$. Hence $\lambda_1 x = \lambda_2 x \Rightarrow (\lambda_1 - \lambda_2) x = 0 \Rightarrow \lambda_1 - \lambda_2 = 0$ $(\because x \neq 0)$ $\Rightarrow \lambda_1 = \lambda_2$.

Theorem 10 : Let λ be an eigenvalue of an operator T on a Hilbert space. If M_λ is the set consisting of all eigenvectors of T corresponding to the eigenvalue λ and the zero vector 0, then M_λ is a non-zero closed linear subspace of H invariant under T.

Proof : By def. $x \in M_\lambda$ iff $T_x = \lambda_{\bar{x}}$ ...(8)

By hypothens $\mathbf{0} \in M_\lambda$ and $\mathbf{0}$ vector also satisfies (8).

Therefore $M_\lambda = \{x \in H : Tx = \lambda x\}$

$$= \{x \in H : (T - \lambda I)x = 0\}$$

Again if $x, y \in M_\lambda$ and α, β are scalars, then

$Tx = \lambda x$ and $Ty = \lambda y$. We have

$$T(\alpha x + \beta y) = T(\alpha x) + Ty(\beta y)$$

$$= \alpha\, Tx + \beta\, Ty$$

$$= \alpha\, \lambda x + \beta\, \lambda y$$

$$= \lambda(\alpha\, x + \beta\, y)$$

$\Rightarrow \alpha x + \beta y \in M_\lambda$

$\Rightarrow M_\lambda$ is a linear subspace of H.

Also T and I are continuous, M_λ is the null space of the continuus transformation $T - \lambda I$. Hence M_λ is closed.

Further let $x \in M_\lambda$

Since M_λ is a linear subspace of H, therefore $x \in M_\lambda \Rightarrow \lambda x = Tx \in M_\lambda \Rightarrow M_\lambda$ is invariant under T.

The closed subspace M_λ is called the eignspace of T corresponding to the eigenvalue λ .

Theorem 11 : If T is a normal operator on a Hilert space H, then x is an eigenvector of T with eigenvalue λ iff x is an eigenvector of T* with $\bar{\lambda}$ as eigenvalue.

Proof : Let T is a normal operator on H. Then TT*=T*T. Now $T - \lambda I$ is also normal, therefore

$$\|(T - \lambda I)x\| = \|(T - \lambda I)^* x\| \ \forall\, x \in H$$

Also adjoint operation is conjugate linear, therefore

$$(T - \lambda I)^* = T^* - \bar{\lambda} I^* = T^* - \bar{\lambda} I$$

From the above two relations we get

$$\|Tx - \lambda x\| = \|T^* x - \overline{\lambda} x\| \ \forall x \in H$$

Hence $Tx - \lambda x = 0$ iff $T^* x - \overline{\lambda} x = 0$

Thus if x is an eigenvector of T with eigenvalue λ iff λ is an eigenvector of T* with eigenvalue $\overline{\lambda}$.

Theorem 12 : If T is a normal operator on a Hibert space H then each eigenspace of T reduces T.

Proof : Let M_λ be the eigenspace of T corresponding to the eigenvalue λ. To prove that M_λ reduces T. We have to show that M_λ is invariant under both T and T*.

We know that M_λ is invariant under T (see Theorem 10). Let $x \in M_\lambda$. Then $Tx = \lambda x \Rightarrow T^* x = \overline{\lambda} x$. Since M_λ is a subspace, $\overline{\lambda} x \in M_\lambda$ whenever $x \in M_\lambda$. Hence $x \in M_\lambda \Rightarrow T^* x = \overline{\lambda} x \in M_\lambda$. Hence M_λ is invariant under T^*. Thus M_λ reduces T.

Theorem 13 : If T is normal operator on a Hilbert space H, then eigenspaces of T are pairwise orthogonal.

Proof : Let M_i and M_j $(i \neq j)$ be eigenspaces of an operator T on Hilbert space H corresponding to distinct eigenvalues λ_i and λ_j. Let $x_i \in M_i$ and $x_j \in M_j$ so that

$$Tx_i = \lambda_i x_i \text{ and } Tx_j = \lambda_j x_j$$

Now $\lambda_i (x_i, x_j) = (\lambda_i x_i, x_j)$

$$= (Tx_i, x_j)$$

$$= (x_i, T^* x_j)$$

$$= (x_i, \overline{\lambda}_j x_j)$$

$$= \lambda_j (x_i, x_j)$$

$$\Rightarrow (\lambda_i - \lambda_j)(x_i, x_j) = 0$$

$$\Rightarrow (x_i, x_j) = 0 \qquad \because \lambda_i \neq \lambda_j$$

$$\Rightarrow x_i \perp x_j \qquad \forall x_i \in M_i \text{ and } x_j \in \lambda_j$$

$$\Rightarrow M_i \perp M_j \ (i \text{ and } j \text{ are arbitrary})$$

8.7 Existence of Eigenvalues

An immediate question that arises before us is :

Does an arbitrary operator T on a Hilbert space H necessarily have an eigenvalue? We shall give an example to show that it is not necessary for an arbitrary operator T on a Hilbert Space H to possess an eigenvalue.

Consider the Hilbert space l_2 and the operator T or l_2 defined by $T\{x_1, x_2, \ldots\} = \{0, x_1, x_2, \ldots\}$

Let λ be an eigenvalue of T. Then $\exists$ a non-zero vector $y = \{y_1, y_2, y_3, \ldots\}$ in $l_2 s.t.$. $Ty = \lambda y$.

Now $Ty = \lambda y \Rightarrow T\{y_1, y_2, y_3, \ldots\} = \lambda\{y_1, y_2, y_3, \ldots\}$

$$\Rightarrow \{0, y_1, y_2, \ldots\} = \{\lambda y_1, \lambda y_2, \lambda y_3, \ldots\}$$

$$\Rightarrow \lambda y_1 = 0, \lambda y_2 = y_1, \ldots$$

Now y is a non-zero vector $\Rightarrow y_1 \neq 0$. Therefore $\lambda y_1 = 0 \Rightarrow \lambda = 0$. Then $\lambda y_2 = y_1 \Rightarrow y_1 = 0$ and this contradicts the fact that y is a non-zero vector. Thesefore T cannot have an eigenvalue.

But if the Hilbert space H is finite dimensional then T on H will have eigenvalues. It should be recalled that if H is finite dimensional, then every linear transformation on H is continuous and is therefore an operator H. So in this case the set $\beta(H)$ is the collection of all linear transformation on H.

Theorem 14: An operator T on a finite-dimensional Hilbert space H is singular $\Leftrightarrow$ there exists a non-zero vector x in H s.t. $Tx = 0$.

Proof : Let $\exists$ a non-zero vector x on H s.t. $Tx = \mathbf{0}$. We have $Tx = T.0 \Rightarrow x = 0$ but $x \neq 0$ by our assumption i.e. x and 0 are distinct vectors in H so that T is not one-one and hence T is not non-singular i.e.T is singular.

Conversly : Let T be singular. To Show that $\exists$ a non-zero vector $x \in H$ s.t. $Tx = 0$. Now $Tx = \mathbf{0}$ $\Rightarrow x = \mathbf{0} \Rightarrow T$ is one one, since

$$Ty = Tz \Rightarrow T(y - z) = 0 \Rightarrow y - z = 0 \Rightarrow y = z .$$

Since H is finite dimensional, therefore T is one-one implies T is onto and so T is non-singular. This contradicts the hypothesis that T is singular. Hence there must exist a non-zero vector x s.t. $Tx = 0$.

Theorems 15 : If T is an arbitrary operator on a finite dimentional Hilbert space H, then the eigenvalues of T constitute a non empty finite subset of the complex plane. Furthermore, the number of points in this does not exceed the dimension *n* of the space H.

Proof : λ is an eigenvalue of $T \Leftrightarrow \exists$ a non-zero vector x s.t. $Tx = \lambda x$

$\Leftrightarrow$ $\exists$ a non-zero vector x s.t. $(T - \lambda I)x = \mathbf{0}$

$\Leftrightarrow$ operator $T - \lambda I$ is singular

$\Rightarrow$ det $(T - \lambda I) = 0$ i.e. $|T - \lambda I| = 0$

If B be any ordered basis for H, then

$$\det(T - \lambda I) = \det(|T - \lambda I|_B) = \det([T]_B - \lambda[I]_B)$$

$$= \det([T]_B - \lambda[\delta_{ij}]) \text{ where } \delta_{ij} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

$$= \det\left(\left[a_{ij}\right] - \lambda\left[\delta_{ij}\right]\right) \text{ (on setting } T = \left[a_{ij}\right]_{mxn}\text{)}$$

$$\therefore \qquad \det(T - \lambda I) = 0 \Rightarrow \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots\ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots\ldots & a_{nn} - \lambda \end{vmatrix} = 0 \qquad \ldots(9)$$

L.H.S of (9) when expanded, yields a polynomial equation in λ of degree n, with complex coefficients having complex roots. But every operator T on H has a eigenvalue and (9) has exactly n roots in complex plane, some of which may be repeated, therefore has distinct eigen values $\leq$ n i.e. not exceeding *n*, the dimension of H.

8.8 Spectral Theorem

We shall require the following result to prove an important theorem known as **spectral theorem:**

Theorem 16 : If $P_1, P_2, \ldots, P_n$ are the projections on closed linear subspaces $M_1, M_2, \ldots., M_n$ of a Hilbert space H, then $P = P_1 + P_2 + \ldots. + P_n$ is a projection $\Leftrightarrow$ the $P_{\hat{i}}$'s are pairwise orthogenal and then P is the projection on $M = M_1 + M_2 + \ldots.. + M_n$.

Proof : $P_{\hat{i}}$'s are projections on $H \Rightarrow P_i^2 = P$ and $P_i^* = P, i = 1, 2, n$ and $P_i\, P_j = 0$ for $i \neq j$.

Now $P = P_1 + P_2 + \ldots\ldots + P_n \Rightarrow P^* = P_1^* + P_2^* + \ldots\ldots. + P_n^*$

$$= P_1 + P_2 + \ldots. + P_n = P$$

$\Rightarrow P$ is self adjoint.

Also $P^2 = P\,P = (P_1 + P_2 + \ldots.. + P_n)(P_1 + P_2 + \ldots + P_n)$

$$= P_1^2 + P_2^2 + \ldots.. + P_n^2 \qquad \because P_i P_j = 0 \text{ for } i \neq j$$

$$= P_1 + P_2 + \ldots. + P_n = P$$

Hence is a projection on H.

Conversly : If $P_i's$ are projections on H i.e. P is a projection on H or $P^2 = P$ and $P^* = P$. We prove that $P_i P_j = 0$ for $i \neq j$.

For any vector $z \in H$ we have

$$(Pz, z) = (PPz, z) = (Pz, P^* z) = (Pz, Pz) = \|Pz\|^2 \qquad \ldots(10)$$

If for any vector $x \in M_i$ (range of P_i) so that $P_i x = x$,

then $\|x\|^2 = \|P_i x\|^2 \leq \sum_{1=1}^{n} \|P_i x\|^2 = \|P_1 x\|^2 + \ldots. + \|P_n x\|^2$

$$\leq \sum_{1=1}^{n} (P_i x, x)$$

$$\leq (P_1 x, x) + (P_2 x, x) + \ldots. + (P_n x, x)$$

$$\leq ((P_1 + P_2 + \ldots.. + P_n) x, x) = (Px, x) = \|Px\|^2$$

$$\leq \|x\|^2$$

So $\|x\|^2 \leq \|Px\|^2$ and $\|Px\|^2 \leq \|x\|^2$

$\Rightarrow$ the sign of equality holds throughout the above computation, thereby giving that

$\|P_i x\|^2 = \sum_{i=1}^{n} \|P_i x\|^2$ and $\|P_i x\|^2 = 0$ for $i \neq j$

$\Rightarrow \|P_i x\| = 0$ for $i \neq j$

$\Rightarrow$ $P_i x = 0$ for $i \neq j$

$\Rightarrow$ $x \in$ null sapce a $P_i (i \neq j)$ whose range is $M_{\hat{i}}$ and null space is $M_i^{\perp}$.

$\Rightarrow$ $x \in M_i^{\perp}$ for $i \neq j$ with $x \in M_i$

$\Rightarrow$ x is orthogonal to the range M_i for every P_i with $i \neq j$

i.e. $x \in M_j \Rightarrow M_j \perp M_i \ \forall i \neq j$

$\Rightarrow$ every vector in range $P_j (j = 1, \ldots, n)$ is orthogornal to the range M_i for every P_i with $i \neq j$

$\Rightarrow$ range of P_j is orthogonal to the range of every $P_{\hat{i}}$ with $i \neq j$.

Lastly we show that P is the projection on $M = M_1 + M_2 + \ldots.. + M_n$. It will be so if the range of P say $R(P) = M$

Any $x \in R(P) \Rightarrow Px = x$

$\Rightarrow (P_1 + P_2 + \ldots.. + P_n) x = x$

$\Rightarrow$ $P_1 x + P_2 x + \ldots + P_n x = x$

where $P_1 x \in M_1, P_2(x) \in M_2, \ldots, P_n(x) \in M_n$

$\Rightarrow x \in M_1 + M_2 + \ldots. + M_n = M$

$\therefore R(P) \subset M$

Also an $x \in M \Rightarrow x_i \in M_i$ for $1 \leq i \leq n$ with $x = x_1 + x_2 + \ldots. + x_n$ and $M = M_1 + M_2 + \ldots. + M_n$.

$\Rightarrow \|Px_i\|^2 = \|x_i\|^2 \Rightarrow \|Px_i\| = \|x_i\|$

$\Rightarrow Px_i = x_i \Rightarrow x_i \in R(P) \forall i$

$\Rightarrow x_1 + x_2 + ...+ x_n \Rightarrow x \in R(P)$ as $R(P)$ is a linear subspace of H

$\therefore M \subset R(P)$

Hence $M = R(P) \Rightarrow P$ is a projection on M.

Statement of spectral Theorem :

Let T be an operator on a finite dimensional Hilbert space H with $\lambda_1, \lambda_2, ..., \lambda_m$ as the distinct eigenvalues of T and with $M_1, M_2, ..., M_n$ be their corresponding eigenspaces. If $P_1, P_2, ..., P_n$ be the projections on these eigenspace, then following statements are equiralent:

(i) The $M_i's$ are pairwise orthogonal and span H.

(ii) The $P_i's$ are pairwise orthogonal and $P_1 + P_2 + ...+ P_n = I$ and $T = \lambda_1 P_1 + \lambda_2 P_2 + ...+ \lambda_m P_m$

(iii) T is a normal operator on H.

Proof : (i) $\Rightarrow$ (ii) : Since $M_i's$ are pairwise orthogornal and span H, therefore each vector $x \in H$ is uniquely expressible as $x = x_1 + x_2 + ...+ x_m, x_i \in M_i \; \forall i = 1, 2, ..., m$. ...(11)

$M_i's$ are pairwise orthogonal and $P_i's$ are projection on $M_i's$

$\Rightarrow P_i's$ are pairwise orthogonal by Theorem 16

$\Rightarrow \quad P_i P_j = 0, i \neq j$

For any vector $x \in H$, (11) yields

$$P_i x = P_i(x_1 + x_2 + ...+ x_n) = P_i x_1 + P_i x_2 + ...+ P_i x_n \qquad ...(12)$$

M_i being range of P_i and $x_i \in M_i \Rightarrow P_i x_i = x_i$.

If $j \neq i$ and $M_j \perp M_i$ for $j \neq i$

$\Rightarrow x_j \in M_i^{\perp}$ for $j \neq i$

$\Rightarrow P_i x_j = 0, M_i^{\perp}$ being null space of P_i.

So $P_i x_i$ and $P_i x_j = 0 \Rightarrow P_i x = x_i \;\; \forall i = 1, 2, ..., m$ ...(13)

Now $\forall x \in H, Ix = x = x_1 + x_2, ...+ x_m$

$$= P_1 x + P_2 x + ...+ P_m x$$

$$=(P_1+P_2+....+P_m)x$$

$$\Rightarrow I = P_1+P_2+....+P_m=\sum_{i=1}^{n} P_i \qquad ...(14)$$

Also $\forall x \in H, Tx = T(x_1+x_2+....x_m)$

$$=Tx_1+Tx_2+.....+Tx_m$$

$$=\lambda_1 x_1+\lambda_2 x_2+....+\lambda_m x_m$$

as $x_i \in M_i \Rightarrow Tx_i = \lambda_i x_i$

$$=\lambda_1 P_1 x+\lambda_2 P_2 x+.....+\lambda_m P_m x$$

$$=(\lambda_1 P_1+\lambda_2 P_2+.....+\lambda_m P_m)x$$

$$\Rightarrow T=\lambda_1 P_1+\lambda_2 P_2+.....+\lambda_m P_m=\sum_{i=1}^{m}\lambda_i P_i \qquad ....(15)$$

The above expression with (14) is called **spectral resolution** of T.

$(\boldsymbol{ii}) \Rightarrow (\boldsymbol{iii})$: Since each P_i being a projection, we have $P_i^* = P_i$ and $P_i^2 = P_i$, $P_i^{'}s$ are pairwise orthogenal and $i \neq j \Rightarrow P_i P_j = 0$ and given that

$$T=\sum_{i=1}^{m}\lambda_i P_i$$

$$\therefore\ T^* = (\lambda_1 P_1+\lambda_2 P_2+....+\lambda_m P_m)^*$$

$$=\overline{\lambda}_1 P_1^*+\overline{\lambda}_2 P_2^*+.....+\overline{\lambda}_m P_m^*$$

$$=\overline{\lambda}_1 P_1+\overline{\lambda}_2 P_2+.....+\overline{\lambda}_m P_m$$

Therefore $TT^* = |\lambda_1|^2 P_1^2 + |\lambda_2|^2 P_2^2+....+|\lambda_m|^2 P_m^2$ as $P_i P_j = 0$ for $i \neq j$ ...(16)

$$=|\lambda_1|^2 P_1+|\lambda_2|^2 P_2+....+|\lambda_m|^2 P_m$$

Similary $T^* T = |\lambda_1|^2 P_1 + |\lambda_2|^2 P_2+....+\left|\lambda_m{}^2\right| P_m$ ...(17)

(16) and (17) $\Rightarrow TT^* = T^*T \Rightarrow T$ is normal

$(\boldsymbol{iii}) \Rightarrow (\boldsymbol{i})$: Let T be normal. We prove that $M_i{}'s$ are pairwise orthogenal which is true by Theorem 13 as $M_i{}'s$ are eigenspaces of T. Again by Theorem 16, $M_i{}'s$ being pairwise orthogonal and $P_i{}'s$ are projections on $M_i{}'s$, $P_i{}'s$ are pairwise othogenal. Theorem 16 also gives $M = M_1 + M_2+......+M_m$. M being a closed linear subspace of H, then its associate projection $P = P_1 + P_2 +....+ P_m$.

Also T is normal on $H \Rightarrow$ each M_i of T reduces T and P_i being orthogonal projection on closed linear subspace M_i of H, M_i reduces T means $P_i T = TP_i$

$$\therefore \quad TP = T(P_1 + P_2 + ...+ P_m)$$

$$= TP_1 + TP_2 + ...+ TP_m$$

$$= P_1T + P_2T + ...+ P_mT$$

$$= (P_1 + P_2 + ...+ P_m)T = PT$$

Hence $TP = PT$ and P is projection on $M \Rightarrow M$ reduces T and so $M^{\perp}$ is invariant under $T \Rightarrow M^{\perp} \neq \{\underline{0}\}$ and all eigenvectors of T being constrained in M, the restriction T to $M^{\perp}$ say that W is an operator on a non-trivial finite dimensional Hilbert space $M^{\perp}$ and $Wx = Tx \;\; \forall\, x \in M^{\perp}$.

Now x being an eigenvector for W corresponding to the eigenvalue λ, we have $x \in M^{\perp}$ and $Wx = \lambda x$.

Thus $Wx = Tx$ and $Wx = \lambda x \Rightarrow Tx = \lambda x \Rightarrow x$ is also an eigenvector for T. But T has no eigenvector in $M^{\perp}$ since all the eigenvectors for T are in M with $M \cap M^{\perp} = \{\mathbf{0}\}$, therefore W is an operater on a finite dirnrensional Hilbert space $M^{\perp}$, having no eigenvecter and no eigenvalue, therefore $M^{\perp} = \{\mathbf{0}\}$ thereby contradicting the hypothesis $M^{\perp} \neq \{\mathbf{0}\}$ in which case every operater on a non-zero finite dimensional Hilbert space would have an eigenvalue.

Consequently, $M^{\perp} = \{\mathbf{0}\} \Rightarrow M = H$

$$\Rightarrow M_1 + M_2 + ...+ M_m = H$$

$\Rightarrow M_i$'s span H.

Self Learning Exercise

In the following questions write T for true and F for false statement :

1. If P is a profection on a Hilbert space H, then P is a positive operator (T/F/)

2 . $\|Px\| \leq \|x\| \;\forall\, x \in H$ (T/F)

3. If x is an eigenvector of T, then x corresponds more than one eigenvalue of T. (T/F)

4. If T is a normal vector on a Hilbert space H, then each eigenspace of T reduces T. (T/F/)

5. An arbitrary operator T on a Hilbert space H possesses necessarily an eigenvalue (T/F)

6. If P be a projection on a closed linear subspace M of a Hilbert space H then $I - P$ is the projection on $M^{\perp}$ (T/F)

7. Let P be a projection on a closed linear subspace M of a Hilbert space H, then

   $x \in M \Leftrightarrow \|Px\| = ...$

8. Let P be a projection on a Hilbert space H, then

(a) $\|P\| \le ....$ (b) $.... \le \|P\| \le ....$

9. If a closed linear subspace M of the Hilbert space H reduces an operator T $\Leftrightarrow$ $M_i's$ invariant under .... and .....

10. If T is a normal operator on a Hilbert space H then eigenspaces of T are pairwise....

8.9 Summary

In this unit you studied the projection on a Hilbert space, invariance and reducilibilty of an operator on a Hilbert space. Spectral theory in Hilbert space was also discussed.

8.10 Answers to Self-Learning Exercise

1. T 2. T 3. F 4. T 5. F 6. T 7. $\|x\|$ 8. $(a) I$ (b) O *and* I

9. T *and* T^* 10. Orthognal

8.11 Exercises

1. Write a short note on Projection on a Hilbert space
2. Define orthogenal Porjection, reducibility and Invariance of an operator on a Hilbert space.
3. If P and Q are projections on closed linear subspaces M and N of a Hilbert space H, then prove that PQ is a projection iff PQ = QP. Also show that PQ is a projection on $M \cap N$.
4. If P and Q are projections on closed linear subspaces M and N of a Hilbert space H, then prove that following statements are equivalent

   (i) $P \le Q$ (ii) $\|Px\| \le \|Qx\| \; \forall x$

   (iii) $M \subseteq N$ (iv) $QP = P$

   (v) $PQ = P$

5. Show by an example that it is not necessary for an arbitrary operator on a Hilbert space H to possess an eigenvalue
6. Define spectral resolution for an operator on a Hilbert space and prove that spectral resolution of a normal operater on a finite dimensional non-zero Hilbert space is unique.
7. If $M_i's$ are eigenspaces for a normal operator T on a Hilbert space H, then prove that $M_i's$ span H.

□□□

Unit - 9
The Derivative

Structure of the Unit

9.0 Objectives

This unit introduces an important concept of derivative of functions in abstract-spaces, particularly in Banach sapces. We are already know the notion of derivative of a real valued function. Now we need to modify this notion of derivative of functions from Banach spaces to Banach spaces.

9.1 Introduction

A real valued function f on R has a derivative $D\,f(a)$ or $f'(a)$ at a point $a \in R$ if and only if for each $\in > 0$ there exists a $\delta > 0$ such that

$$\left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| < \in \text{ whenever } 0 < |x - a| < \delta$$

Frechet generalized this concept of derivative of a mapping f on a normed linear space N into a normed linear space M. The derivative of f at a point $a \in N$ exists and it is a linear transformation g of N into M if it satisfies the inequality, for $\in > 0$,

$$\| f(x) - f(a) - g(x - a) \| \leq \in \| x - a \|,$$

whenever $\| x - a \| \leq \delta$.

9.2 Derivative

Definition : Let X and Y be any two Banach spaces and V an open subset in X, then two functions $f_1 : V \to Y$ and $f_2 : V \to X$ are said to be **tangential** to each other at a point $v \in V$ if, we have

$$\lim_{\substack{x \to v \\ x \neq v}} \frac{\| f_1(x) - f_2(x) \|}{\| x - v \|} = 0$$

which follows that

$$f_1(v) = f_2(v)$$

If f_1, f_2 are tangential at v and f_2, f_3 are also tangential at v, then f_1, f_3 are tangential at v, since we have the inequality,

$$\|f_1(x) - f_3(x)\| \le \|f_1(x) - f_2(x)\| + \|f_2(x) - f_3(x)\|$$

Hence this relation is an equivalence relation.

Theorem 1 : Let X and Y be any two Banach spaces over the same filed K. In the set of all functions tangential to a function f at $v \in V$, there is at most one function $\phi : X \to Y$, of the form $\phi(x) = f(v) + g(x - v)$, where $g : X \to Y$ is linear, where V is an non-empty open subset of X.

Proof : Suppose there are two functions ϕ and ψ from X into Y given by

$$\phi(x) = f(v) + g(x - v) \text{ and } \psi(x) = f(v) + g_1(x - v)$$

Assume $h(x) = g(x) - g_1(x)$,

then clearly h is linear and

$$\lim_{\substack{x \to 0 \\ x \ne 0}} \frac{\|h(x)\|}{\|x\|} = 0$$

Thus for given $\in > 0$ there exists $a\,\delta > 0$ such that

$$\|h(x)\| \le \in \|x\| \text{ whenever } \|x\| \le \delta$$

But $\in > 0$ is an arbitraril y small so that

$h(x) = 0$ for any x.

$\Rightarrow \quad g = g_1$

Hence $\phi = \psi$

Derivative of a Map :

Definition : Let X and Y be Banach spaces and V be a non-empty open subset of X. A continuous mapping $f : V \to Y$ is said to be differentiable at the point $v \in V$ if there exists a linear mapping $g : X \to Y$ such that the mapping $x \to f(x) - f(v)$ and $x \to g(x - v)$ are tangential at the point v, that is

$$\lim_{\substack{x \to v \\ x \ne v}} \frac{\|f(x) - f(v) - g(x - v)\|}{\|x - v\|} = 0 \qquad ...(1)$$

Let $x = v + h \in V$, we assume

$$n(h) = f(v + h) - f(v) - g(h) \Rightarrow f(v + h) = f(v) + g(h) + n(h) \qquad ...(2)$$

where from equation (1), we have

$$\lim_{h \to 0} \frac{\|n(h)\|}{\|h\|} = 0 \qquad \text{...(3)}$$

n being a function from $N \to Y$, where N is a neighbourhood of $0 \in X$, such that $N + v \subset V$. A function $f : V \to Y$ is said to be differentiable in V if f is differentiable at each point of V.

If f is differentiable in V, then for each point $v \in V$, $D f(v) \in L(x, y)$, which is the space of all linear map from X into Y.

Example 1 : The derivative of the constant function $f : V \to Y$ is the zero linear map, because

$\|f(x) - f(v) - g(x - v)\| = 0$ for any $v, x \in V$, if g is the zero map of $L(x, y)$.

Example2 : The derivative of a continuous linear mapping $f : V \to Y$ is the mapping f itself, because

$$\|f(x) - f(v) - f(x - v)\| = \|f(x) - f(v) - f(x) + f(v)\|$$

$$= 0 \quad , \qquad \forall\, x, v \in V$$

Theorem 2 : Let X and Y be Banach spaces and V be the non-empty open subset of X. Suppose that $f : V \to Y$ and $g : V \to Y$ be differentiable in V and a be any scalar in K. Then the function $(f + g) : V \to Y$ and $\alpha f : V \to Y$ defined by $a f(x) = a f(x)$, $(f + g)(x) = f(x) + g(x)$, are differentiable in V and for all $v \in V$, $D(a f)(v) = a D f(v)$, $D(f + g)(v) = D f(v) + D g(v)$

Let us prove, $D(f + g)(v) = D f(v) + D g(v)$

Proof : Since f and g are differentiable at $v \in V$, so that

$$\lim_{\substack{x \to v \\ x \neq v}} \frac{\|f(x) - f(v) - D f(v)(x - v)\|}{\|x - v\|} = 0$$

and $$\lim_{\substack{x \to v \\ x \neq v}} \frac{\|g(x) - g(v) - D g(v)(x - v)\|}{\|x - v\|} = 0$$

Now,

$$\lim_{\substack{x \to v \\ x \neq v}} \frac{\|(f + g)(x) - (f + g)(v) - (D f(v) + D g(v))(x - v)\|}{\|x - v\|}$$

$$\leq \lim_{\substack{x \to v \\ x \neq v}} \frac{\|f(x) - f(v) - D f(v)(x - v)\|}{\|x - v\|}$$

$$+\lim_{\substack{x\to v\\ x\neq v}}\frac{\|g(x)-g(v)-Dg(v)(x-v)\|}{\|x-v\|}$$

$$=0+0=0$$

The $D(f+g)(v)=Df(v)+Dg(v),\ \forall\, v\in V$

Similarly, we can prove that $D(af)(x)=a\,Df(v)$.

Theorem 3 (Derivative of a composite mapping) : Let X, Y and Z be Banach spaces over the same field K. Suppose that f is a function on an open subset V of X into an open subset W of Y and g is a function on W into Z. If f is differentiable at a point $v\in V$ and g is differentiable at the point $w=f(v)\in W$, then g of is differentiable at v and

$$D(gof)(v)=\left(Dg(f(v))\right)\ o\ Df(v)$$

or $$(gof)'(v)=\left(g'(f(v))\right)\ o\ f'(v)$$

Proof : Let $k\in Y$ be such that $f(v)+k\in W$.

Given that g is differentiable at $f(v)$, so we have

$$g(f(v)+k)=g(f(v))+Dg(f(v)).k+\psi(k) \qquad ...(1)$$

where $\lim_{k\to 0}\frac{\|\psi(k)\|}{\|k\|}=0$

Now let $h\in X$ be such that $v+h\in V$

Given that f is differentiable at $v\in V$, so we have

$$f(v+h)=f(v)+Df(v)\ .\ h+\eta(h)$$

where $\lim_{h\to 0}\frac{\|\eta(h)\|}{\|h\|}=0 \qquad ...(2)$

Now, we have

$$(gof)(v+h)=g\ \{f(v+h)\}$$

$$=g\ \{f(v)+Df(v)\ .\ h+\eta(h)\} \qquad \text{[From eqn. (1)]}$$

Using eqn. (1), we get

$$(gof)(v+h)=g(f(v))+Dg(f(v))\ .\ \{Df(v)\ .\ h+\eta(h)\}+\psi(Df(v)\ .\ h+\eta(h))$$

$$= (g o f)(v) + \left(D g(f(v)) o D f(v)\right) . h + \phi(h) \qquad ...(3)$$

where $\phi(h) = D g(f(v))\, \eta(h) + \psi\left(D f(v) h + \eta(h)\right)$ ...(4)

Now we claim that

$$\lim_{h \to 0} \frac{\|\phi(h)\|}{\|h\|} = 0$$

Let for $\in > 0$ there exists $\mu > 0$ such that $\|k\| \le \mu$,

$$\|\psi(k)\| \le \in \|k\|$$

Also there exists a $\delta > 0$ such that $\|h\| \le \delta$,

$$\|\eta(h)\| \le \in \|h\| \text{ and } \|D f(v).h + \eta(h)\| \le \mu$$

Then for $\|h\| \le \delta$, we have

$$\|\phi(h)\| = \|D g(f(v))\, \eta(h) + \psi\left(D f(v)\right) . h + \eta(h)\|$$

$$\le \|D g(f(v))\, \eta(h)\| + \|\psi\left(D f(v)\right) . h + \eta(h)\|$$

$$\le \|D g(f(v))\|\, \|\eta(h)\| + \in \|D f(v) . h + \eta(h)\|$$

$$\le \|D g(f(v))\| \in \|h\| + \in \|D f(v)\|\, \|h\| + \in . \in \|h\|$$

$$\Rightarrow \quad \|\phi(h)\| \le \left(\|D g(f(v))\| + \|D f(v)\| + \in\right) \in \|h\|$$

$$\Rightarrow \quad \frac{\|\phi(h)\|}{\|h\|} \le \left(\|D g(f(v))\| + \|D f(v)\| + \in\right) \in$$

But $\in > 0$ arbitrary, so that

$$\lim_{h \to 0} \frac{\|\phi(h)\|}{\|h\|} = 0$$

Thus the equation (3) can be written as

$$\lim_{h \to 0} \frac{\|(g o f)(v+h) - (g o f)(v) - \left(D g(v)\right) o\ D f(v) h\|}{\|h\|} = 0$$

Hence $D(g o f)(v) = \left(D g(f(v))\right) o\ D f(v)$

Definition : A bijection f on a Banach space X onto a Banach space Y is said to be a homeomorphism if both f and f^{-1} are continuous on X and Y respectively.

Theorem 4 : Let X and Y be Banach spaces over the same field K of scalars. Let f be a homeororphism of an open subset V of X onto an open subset W of Y and let g be the inverse homoeomorphism of W onto V. If f is differentiable at $a \in V$ and $Df(a)$ is a linear homeomorphism of X onto Y, then g is differentiable at the point $b = f(a) \in W$ and

$$Dg(b) = [Df(a)]^{-1}$$

Proof : It is clear that the linear mapping $Df(a)$ of X onto Y has an inverse linear mapping. Let it be $t = [Df(a)]^{-1}$ of Y onto X. It is also continuous and there is a finite positive real number M such that

$$\|t(y)\| \le M\|y\|, \qquad \forall\, y \in Y \qquad ...(1)$$

Suppose that $h \in X$ be such that $a + h \in V$. Since f is differentiable at a, so that we have

$$f(a+h) = f(a) + Df(a)h + \eta(h) \qquad ...(2)$$

where $\lim\limits_{h \to 0} \dfrac{\|\eta(h)\|}{\|h\|} = 0$

Let for given $0 < \in' \le \dfrac{1}{2M}$, there exists $\delta' > 0$ such that

$$\|\eta(h)\| \le \in'\|h\|, \text{ whenever } \|h\| \le \delta' \qquad ...(3)$$

Since g is continuous at $b = f(a) \in W$, then for given $\delta' > 0$ there exists $\delta > 0$ such that

$$\|g(b+k) - g(b)\| \le \delta', \text{ whenever } \|k\| \le \delta$$

Now, $\|h - t(f(a+h)) - f(a)\| = \|t\,\{Df(a)\,.\,h - (f(a+h) - f(a))\}\|$

$= \|t\,\{Df(a)\,.\,h - f(a+h) + f(a)\}\|$

$= \|t\,(-\eta(h))\|$ from eqn. (2)

$\le M\,\|\eta(h)\|$ from eqn. (1)

$\le M \in'\,\|h\|$ from eqn. (3) ...(4)

Now, $\|h\| = \|h - t(f(a+h)) - f(a) + t(f(a+h)) - f(a)\|$

$\le \|h - t(f(a+h)) - f(a)\| + \|t(f(a+h)) - f(a)\|$

$$\leq M \in' \|h\| + M\| f(a+h) - f(a)\| \qquad \text{from (1) \& (4)}$$

$$\leq \frac{1}{2}\|h\| + M\| f(a+h) - f(a)\|$$

$$\Rightarrow \qquad \|h\| \leq 2M\| f(a+h) - f(a)\| \qquad \text{...(5)}$$

Suppose $g(b+k) = a+h$, then

$$\|g(b+k) - g(b) - t(k)\|$$

$$= \|a + h - a - t(f(a+h)) - f(a)\|$$

$$= \|h - t(f(a+h)) - f(a)\|$$

$$\leq \in' M\|h\| \qquad \text{From (4)}$$

$$\leq \in' M\{2M\|f(a+h) - f(a)\|\} \qquad \text{From (5)}$$

$$= 2 \in' M^2 \|f(a+h) - f(a)\|$$

$$\Rightarrow \qquad \|g(b+k) - g(b) - t(k)\| \leq \in \|k\|, \qquad \in = 2 \in' M^2$$

Hence for given $\in > 0$ there exists $\delta > 0$ such that

$$\|g(b+k) - g(b) - t(k)\| \leq \in \|k\|, \text{ whenever } \|k\| \leq \delta$$

Hence g is differentiable at $b = f(a) \in W$ and $D\,g(b) = t = [D\,f(a)]^{-1}$

9.3 Directional Derivative

Definition : Let X and Y be Banach space over the same field K of scalars and V be an open subset of X. Let f be a function from V into Y and v be a unit vector in V, then the directional derivative of f at $x \in V$ in the direction of unit vector v is denoted by $D_v\, f(x)$ and is defined by

$$D_v\, f(x) = \lim_{h \to 0} \frac{f(x+hv) - f(x)}{h}, \text{ if this limit exists.}$$

Theorem 5 : Suppose that X and Y be Banach spaces over the same filed K of scalars and V be an open subset of X. Let $f : V \to Y$ is differentiable at $x \in V$. Then all the directional derivatives of f exists at x and

$$D_v f(x) = D\,f(x).v \text{, where } v \in V \text{ is a unit vector.}$$

Proof : Suppose $h \in X$ be such that $x + h \in V$. Given that f is differentiable at $x \in V$, so that

$$f(x+h) = f(x) + D\,f(x)\,h + \eta(h),$$

where $\lim_{h \to 0} \frac{\|\eta(h)\|}{\|h\|} = 0$

Since $v \in V$ is a unit vector and let s is arbitrary small, then we have

$$f(x+sv) = f(x) + Df(x).(sv) + \eta(sv) \qquad ...(1)$$

where $\lim_{s \to 0} \frac{\|\eta(sv)\|}{\|sv\|} = 0$

$$\Rightarrow \quad \lim_{s \to 0} \left\| \frac{\eta(sv)}{s} \right\| = 0$$

$$\Rightarrow \quad \lim_{s \to 0} \frac{\eta(sv)}{s} = 0$$

Hence from (1), we have

$$\lim_{s \to 0} \frac{f(x+sv) - f(x)}{s} = \lim_{s \to 0} \left\{ Df(x) \,.\, v + \frac{\eta(sv)}{s} \right\}$$

$$= Df(x) \,.\, v + 0$$

$$\Rightarrow \quad D_v f(x) = Df(x) \,.\, v$$

9.4 Mean Value Theorem and its Applications

In this section we study mean value theorem for a mapping defined on a Banach space.

Theorem 6 : Let X be a Banach space over the field K of scalars and let $f : [a,b] \to X$ and $g : [a,b] \to R$ be continuous and differentiable functions such that $\|Df(t)\| \leq Dg(t)$ at each point $t \in (a,b)$. Then

$$\|f(b) - f(a)\| \leq g(b) - g(a)$$

Proof : Let $\in > 0$ and let T be the set of real numbers $s \in [a,b]$ such that $\forall\, r \in [a,s)$

$$\|f(r) - f(a)\| \leq g(r) - g(a) + \in (r - a + 1) \qquad ...(1)$$

It is given htat

$$\|Df(t)\| \leq Dg(t) \qquad \forall\, t \in (a,b)$$

$$\Rightarrow \quad Dg(t) \geq 0 \qquad \forall\, t \in (a,b)$$

$\Rightarrow$ g is an increasing function on (a,b). Since f is a continuous function in closed interval $[a,b]$ so it is uniformly continuous in $[a,b]$, then there is a real number $p \in (a,b]$ such that $\forall\, q \in [a,p)$

$$\|f(q)-f(a)\| \le \in \qquad ...(2)$$

Since g in increasing function $q \in [a, p]$ so that

$$\in \le g(q)-g(a)+\in(q-a+1) \qquad ...(3)$$

From (2) & (3), we obtain

$$\|f(q)-f(a)\| \le g(q)-g(a)+\in(q-a+1)$$

$\Rightarrow$ $p \in T$ and hence T in non-empty

Now we define a function $h:[a,b]\to R$ as follows

$$h(s)=\|f(s)-f(a)\|-g(s)+g(a)-\in(s-a+1) \qquad ...(4)$$

Then clearly h is continuous in $[a,b]$ and also $h(s)\le 0\ \forall\, s\in T$.

Then T is a closed subset of $[a,b]$ and so it is bounded.

Now T is a non-void bounded subset of R.

Hence supremum of T exists in $[a,b]$.

Let supermum of $S=c$

We shall show that $c=b$.

As contradiction we suppose that $c\ne b$ i.e. $a\le c<b$.

Given that f and g are differentiable in (a,b) and so that there in a real number $q\in(c,b)$ such that $\forall\, s\in(c,q)$

$$\|f(s)-f(c)-Df(c)(s-c)\| \le \in\frac{(s-c)}{2} \qquad ...(5)$$

and $$\|g(s)-g(c)-Dg(c)(s-c)\| \le \in\frac{(s-c)}{2} \qquad ...(6)$$

Now, $\|f(s)-f(c)\|$

$$=\|f(s)-f(c)-Df(c)(s-c)+Df(c)(s-c)\|$$

$$\le \|f(s)-f(c)-Df(c)(s-c)\|+\|Df(c)(s-c)\|$$

$$\le \frac{\in}{2}(s-c)+\|Df(c)\|\,(s-c) \qquad \text{From (5)}$$

$$\le \frac{\in}{2}(s-c)+Dg(c)\,(s-c) \qquad \left|\because \|Df(t)\|\le Dg(t)\ \forall\, t\in(a,b)\right.$$

$$\leq \frac{\in}{2}(s-c)+g(s)-g(c)+\frac{\in}{2}(s-c)$$

$\Rightarrow \quad \|f(s)-f(c)\| \leq g(s)-g(c)+\in(s-c)$

Since $c \in T$, therefore

$$\|f(c)-f(a)\| \leq g(c)-2(a)+\in(c-a+1)$$

Thus $\quad \forall\, s \in [c,q)$

$$\|f(s)-f(a)\| = \|f(s)-f(c)+f(c)-f(a)\|$$

$$\leq \|f(s)-f(c)\| + \|f(c)-f(a)\|$$

$$\leq g(s)-g(c)+\in(s-c)+g(c)-g(a)+\in(c-a+1) \text{ From (7) \& (8)}$$

$\Rightarrow \quad \|f(s)-f(a)\| \leq g(s)-g(a)+\in(s-a+1)$

$\Rightarrow \quad s \in T$ and $s > c$

which is the contradiction to the fact that c is the supermum of T, so our assumption $c < b$ was wrong and hence $c = b$ and

$$\|f(b)-f(a)\| \leq g(b)-g(a)+\in(b-a+1)$$

But $\quad \in > 0$ is an arbitrary small and so

$$\|f(b)-f(a)\| \leq g(b)-g(a)$$

Theorem 7 (Mean value Theorem) : Let X and Y be any two Banach spaces over the same field K of scalars and V be an open subset of X. Let $f : V \to Y$ be continuous function. Let u, v be any two distinct points of V such that $[u,v]\ C\ V$ and f is differentiable in $[u,v]$. Then

$$\|f(v)-f(u)\| \leq \|v-u\| \sup\{\|D f(x)\| : x \in [u,v]\}$$

Proof : We define a mapping $g : [0,1] \to Y$ such that

$$g(t) = f(u+t(v-u)) \qquad \forall\, t \in [0,1]$$

As f is differentiable in $[u,v]$, therefore g is differentiable in $[0,\ 1]$

$\therefore \quad D g(t) = \{D f(u+t(v-u))\}\,.\,(v-u)$

$\Rightarrow \quad \|D g(t)\| = \|(v-u)\ D f(u+t(v-u))\|$

$\Rightarrow \quad \|D g(t)\| \leq \|(v-u)\| \sup\{\|D f(u+t(v-u))\| : t \in [0,\ 1]\}$ ...(1)

Let $c = \|v-u\| \sup \{\|Df(u+t(v-u))\| : t \in [0, 1]\}$

Now we define a mapping $h : [0, 1] \to R$ such that

$$h(t) = ct$$

Then h is obviously continuous and differentiable in $[0, 1]$

$$\therefore \quad Dh(t) = c$$

From (1), we have

$$\|Dg(t)\| \le c = Dh(t)$$

$$\Rightarrow \quad \|Dg(t)\| \le Dh(t) \qquad \forall\, t \in (0,1)$$

Now we know that if $g : [0, 1] \to Y$ and $h : [0, 1] \to R$ are continuous and differentiable such that

$$\|Dg(t)\| \le Dh(t) \quad \text{at each point} \quad t \in (0,1),$$

then $\|g(1) - g(0)\| \le h(1) - h(0)$ [From theorem 6]

$$\Rightarrow \quad \|f(v) - f(u)\| \le c - 0$$

$$\Rightarrow \quad \|f(v) - f(u)\| \le c$$

$$\Rightarrow \quad \|f(v) - f(u)\| \le \|v-u\| \sup \{\|Df(u+t(v-u))\| : t \in [0,1]\}$$

$$\Rightarrow \quad \|f(v) - f(u)\| \le \|v-u\| \sup \{\|Df(x)\| : x \in [u,v]\}$$

Theorem 8 : Let X be a Banach space over the field K of scalars, and V be an open subset of X. Suppose $f : V \to R$ be a function. Let u and v be any two distinct points in V such that $[u,v] \subset V$ and f is differentiable at all points of $[u,v]$. Then

$$f(v) - f(u) = Df(u+t(v-u)) \,.\, (v-u) \quad \text{where } t \in (0,1).$$

Proof : We define a mapping $g : [0, 1] \to R$ such that

$$g(s) = f(u+s(v-u)), \qquad \forall\, s \in [0,1] \qquad ...(1)$$

As f is differentiable in $[u,v]$, therefore g is differentiable in $[0, 1]$, and

$$Dg(s) = Df(u+s(v-u)) \,.\, (v-u), \quad s \in [0,1] \qquad ...(2)$$

Now from Lagrange's mean value theorem, there exists a real number $t \in (0, 1)$ such that

$$\frac{g(1)-g(0)}{1-0}=g'(t), \quad 0<t<1$$

$$\Rightarrow \quad g(1)-g(0)=g'(t)$$

$$\Rightarrow \quad g(1)-g(0)=Df\left(u+t(v-u)\right).(v-u) \qquad ...(3)$$

From (1),

$$g(1)=f(v), \quad g(0)=f(u)$$

Using these in (3), we obtain

$$f(v)-f(u)=Df\left(u+t(v-u)\right).(v-u), \qquad t\in(0,1)$$

Theorem 9 : Let X and Y be any two Banach spaces over the same field K of scalars and V be an open subset of X. Let $f:V\to Y$ be a continuous function and let u and v be any two distinct points in V such that $[u,v]\subset V$ and f is differentable in $[u,v]$. Suppose $g:X\to Y$ be any continuous linear function. Then

$$\|f(v)-f(u)-g(v-u)\|\le c\|v-u\|,$$

where $c\in R$ be such that $\|Df(x)-g\|\le c$, $\forall\, x\in[u,v]$

Proof : We define a mapping $h:V\to Y$ such that

$$h(x)=f(x)-g(x-v), \ \forall\, x\in V \qquad ...(1)$$

Then clearly h is continuous and differentiable in $[u,v]$

and $\quad Dh(x)=Df(x)-g$, $x\in V$, since g is linear ...(2)

Now since $h:V\to Y$ is continuous function and $u,v\in V$ be such that $[u,v]\subset V$ and h is differentiable in $[u,v]$, then from mean value theorem, we have

$$\|h(v)-h(u)\|\le\|v-u\| \sup\left\{\|Dh(x)\| : x\in[u,v]\right\}$$

Using (1) & (2), we have

$$\|f(v)-g(v-v)-f(u)+g(u-v)\|\le\|v-u\| \sup\left\{\|Df(x)-g\| : x\in[u,v]\right\}$$

$$\Rightarrow \quad \|f(v)-f(u)-g(v-u)\|\le c\|v-u\|$$

where $\|Df(x)-g\|\le c \ \forall\, x\in[u,v]$

Self-Learning Exercise

1. Define, when two functions f_1 and f_2 defined on an open subset of a Banach space are tangential at a point.

2. Define derivative on a Banach space.

3. True/Fase Statements :

(a) The derivative of the constant function f on an open subset V of a Banach space X into Banach space Y is the zero map.

(b) The derivative of a continuous linear mapping f on an open subset V of a Banach space X into a Banach space Y is the mapping f itself.

9.5 Summary

In this unit we studied the notion of derivative of function from one Banach space into another Banach space and concepts of mean value Theorem in Banach spaces.

9.6 Answers to Self-Learning Exercise

1. See text 2. See Text 3. (a) True (b) True

9.7 Exercises

1. Let f be a differentiable function on a non void connected open subset V of a Banach space X over K into a Banach space Y over K such that $Df = 0$. Then f is a constant function.

2. Let $f\ ; [a,b] \to X$ and $g\ ; [a,b] \to R$ are continuous and differentiable function such that $\|Df(t)\| \le Dg(t)$ at each point $t \in (a,b)$, then

$$\|f(b) - f(a)\| \le g(b) - g(a).$$

3. Let X, Y be Banach space over K and let V, W be open subsets in X respectively. Let $f : V \to Y$ be differentiable at a point $a \in U$ and $g : W \to X$ be differentiable at the point $b \in W$, where $b = f(a)$. If $fog = I_y$ and $gof = I_X$. Then

$$Dg(b) = [Df(a)]^{-1}$$

□□□

Unit - 10
Higher Derivatives

Structure of the Unit

10.0 Objectives

In this unit we shall study the concept of higher derivatives of a function on Bahach spaces, which have an important role in the study of these functions.

10.1 Introduction

In this unit we shall introduce higher derivatives of functions defined on Banach spaces and the concept of continuously differentiable maps on Banach spaces (C^n-maps), partial derivatives, Taylor's theorem and existence theorems will be discussed with their applications.

10.2 Continuously differentiable Maps $\left(C^1 - maps\right)$

Definition : Let X and Y be Banach spaces over the same field K and V be an open subset of X. Suppose $f:V \to Y$ is a differentiable function at each point of V. Then f is said to be a continuously differentiable map $\left(C^1 - map\right)$ in V if and only if the function $Df:V \to L(X,Y)$ is continuous.

Definition : Let V be a non-empty open subset of a Banach space $X = X_1 \times X_2$ and let f be a function of V into Y. Suppose $(a_1, a_2) \in V$, we define $V_1 = \{x_1 \in X_1 : (x_1, a_2) \in V\}$. Then V_1 is an open subset of X_1

We also define a mapping $g:V_1 \to Y$ such that $g(x_1) = f(x_1, a_2) \ \forall x_1 \in V_1$

Similarly we define the set

$V_2 = \{x_2 \in X_2 : (a_1, x_2) \in V\}$ and the mapping

$h:V_2 \to Y$ such that

$h(x_2) = f(a_1, x_2) \ \forall \ x_2 \in V_2$

The mapping $f:V\to Y$ is said to be differentiable with respect to the first variable at the point (a_1,a_2) iff g is differentiable at a_1, and we write $Dg(a_1)=D_1f(a_1,a_2)$ or $f_1'(a_1,a_2)$. The derivative $D_1f(a_1,a_2)$ is called the partical derivative of f with respect to the first variable at (a_1,a_2), it is a linear map of X_1 into y.

Similarly, we can define the partial derivative $D_2f(a_1,a_2)$ with respect to the second variable

Thus, we have $D_1f(a_1,a_2)\in L(X_1,Y)$

and $D_2f(a_1,a_2)\in L(X_2,Y)$

Theorem 1 : Let f be a continuous mapping of an open subset V of $X_1\times X_2$ into Y. Then f is a C^1-map in V iff f be differentiable at each point with respect to the first and the second variable. Also the mappings $(a_1,a_2)\to D_1f(a_1,a_2)$ and $(a_1,a_2)\to D_2f(a_1,a_2)$ are continuous on V. Further at each point $(x_1,x_2)\in V$, the derivative of f is given by

$$Df(a_1,a_2)(x_1,x_2)= D_1f(a_1,a_2)x_1+D_2\,f(a_1,a_2)x_2$$

Proof : First suppose that f is C^1-map on V into Y. Let $(a_1,a_2)\in V$ then for $(x_1,x_2)\in V$ and given $\in>0$ there exists $\delta>0$ such that

$$\|f(x_1,x_2)-f(a_1,a_2)-Df(a_1,a_2)((x_1,x_2)-(a_1,a_2))\|$$

$$\le\in\|(x_1,x_2)-(a_1,a_2)\|$$

Put $x_2=a_2$, we get

$$\|f(x_1,a_2)-f(a_1,a_2)-Df(a_1,a_2)((x_1,a_2)-(a_1,a_2))\|$$

$$\le\in\|(x_1,a_2)-(a_1,a_2)\| \qquad ...(1)$$

Since,

$$\|(x_1,a_2)-(a_1,a_2)\|=\|(x_1-a_1,a_2-a_2)\|$$

$$=\|(x_1-a_1,0)\|$$

$$=\|(x_1-a_1)\|$$

Using it in (1), we have

$$\|f(x_1,a_2)-f(a_1,a_2)-Df(a_1,a_2)(x_1-a_1,0)\|\le\in\|(x_1-a_1)\|,$$

$$\text{for } \|(x_1-a_1)\|\le\delta$$

Thus f is differentiable with respect to the first variable at (a_1,a_2)

and $D_1 f(a_1,a_2)x_1 = Df(a_1,a_2).(x_1,0)$

Similarly we have

$$D_2 f(a_1,a_2)x_2 = Df(a_1,a_2).(0,x_2)$$

Now

$$Df(a_1,a_2)(x_1,x_2) = Df(a_1,a_2)\{(x_1,0)+(0,x_2)\}$$

$$= Df(a_1,a_2)(x_1,0) + Df(a_1,a_2)(0,x_2)$$

$$= D_1 f(a_1,a_2)x_1 + D_2 f(a_1,a_2)x_2$$

Which is the required result.

Since Df is continuous, therefore $D_1 f$ and $D_2 f$ are also continuous on V.

Conversely suppose that $D_1 f$ and $D_2 f$ are continuous and differentiable at each point $(a_1,a_2)\in V$.

To prove, f is C^1-map, we have

$$f(a_1+x_1,a_2+x_2)-f(a_1,a_2)-\left(D_1 f(a_1,a_2)x_1 + D_2 f(a_1,a_2)x_2\right)$$

$$= f(a_1+x_1,a_2+x_2)-f(a_1,a_2+x_2)-D_1 f(a_1,a_2)x_1$$

$$+ f(a_1,a_2+x_2)-f(a_1,a_2)-D_2 f(a_1,a_2)x_2 \qquad ....(2)$$

Let

$$g(z)\equiv f(a_1+z,a_2+x_2)-D_1 f(a_1,a_2)z,$$

Where $z=tx$ and $t\in(0,1)$

So that

$$Dg(z)=D_1 f(a_1+z,a_2+x_2)-D_1 f(a_1,a_2)$$

Since $D_1 f$ is continuous so for any $\in>0$ there is open ball of radius η_1, and centered at (a_1,a_2) such that for all $(x_1,x_2)\in B((a_1,a_2);\eta_1)$ we have

$$\left\|D_1 f(a_1+z,a_2+x_2)-Df_1(a_1,a_2)\right\|\le\in,$$

for $z=tx_1, t\in(0,1)$

So by the mean value theorem,

$$\|g(x_1)-g(0)\| \le \in \|x_1\|, \quad z \in (0, x_1)$$

$$\Rightarrow \|f(a_1+x_1, a_2+x_2) - f(a_1, a_2)x_2 - D_1 f(a_1, a_2)x_1\| \le \in \|x_1\| \qquad ...(3)$$

Since $D_2 f$ is also continuous, similarly,

We have, for

$$(x_1, x_2) \in B((a_1, a_2); \eta_2),$$

$$\|f(a_1, a_2+x_2) - f(a_1, a_2) - D_2 f(a_1, a_2)x_2\| \le \in \|x_2\| \qquad ...(4)$$

Let us take $\eta = \min(\eta_1, \eta_2)$

Now, from equation (2), we have

$$\|f(a_1+x_1, a_2+x_2) - f(a_1, a_2) - (D_1 f(a_1, a_2)x_1 + D_2 f(a_1, a_2)x_2)\|$$

$$= \|f(a_1+x_1, a_2+x_2) - f(a_1, a_2+x_2) - D_1 f(a_1, a_2)x_1$$

$$+ f(a_1, a_2+x_2) - f(a_1, a_2) - D_2 f(a_1, a_2)x_2\|$$

$$\le \|f(a_1+x_1, a_2+x_2) - f(a_1, a_2+x_2) - D_1 f(a_1, a_2)x_1\|$$

$$+ \|f(a_1, a_2+x_2) - f(a_1, a_2) - D_2 f(a_1, a_2)x_2\|$$

Using (3) and (4), we obtain

$$\|f(a_1+x_1, a_2+x_2) - f(a_1, a_2) - (D_1 f(a_1, a_2)x_1 + D_2 f(a_1, a_2)x_2)\|$$

$$\le \in \|x_1\| + \in \|x_2\|$$

$$= \in (\|x_1\| + \|x_2\|)$$

Since $\in$ is arbitrary small positive quantity, therefore f is differentiable at (a_1, a_2)

Since $D_1 f$ and $D_2 f$ are continuous then Df is also continuous in V.

Hence f is C^1 map.

10.3 Higher Derivatives

Suppose X and Y be Banach spaces over the same field K of scalars and V be an open non-void subset of X. $f : V \to Y$ is a $C^1 - map$ then the map $Df : V \to L(X, Y)$ is continuous. If the map Df is differentiable at a given point $v \in V$, then $D(Df(v))$ will be a linear map $X \to L(X, Y)$. This map is called the second derivative of f at v and is denoted by $D^2 f(v)$. The map Df is continuous implies that $D^2 f(v)$ is a continuous linear map i.e. $D^2 f(v) \in L(X, L(X, Y))$. If Df is differentiable on V, then we have a map $D^2 f : V \to L(X, L(X, Y))$. It this map is continuous, we say that f is a $C^2 - map$.

Since $L(X, L(X, Y)) \cong L(X^2, Y)$

We write $D^2 f(v) \in L(X^2, Y)$

Continuing in this manner, f is a $C^{n-1}-map$ then the map $D^{n-1} f : V \to L(X^{n-1}, Y)$ in continuous. Its derivative, if it exists at $v \in V$ is called the n^{th} derivative of f at v and is denoted by $D^n f(v)$ and it is an element of $L(X^n, Y)$.

If $D^{n-1} f$ in differentiable on V, then we have the map $D^n f : V \to L(X^n, Y)$.

It this map is continuous, then we say that f is $C^n - map$.

For each $v \in V$ and each $(x_1, x_2, ..., x_n) \in X^n$, we have

$$D^n f(v).(x_1, x_2 ....., x_n) = D(D^{n-1} f(v).x_1) \ (x_2, x_3, ..., x_n).$$

From the definition of higher derivatives, we obtain the following properties :

1. Let $f : V \to Y$ in m-times differentiable in V and $D^m f$ is n-times differentiable in V. Then by induction f is $(m+n)$ times differentiable in V and

$$D^n(D^m f) = D^{m+n} f$$

2. Let $f : V \to Y$ and $g : V \to Y$ are n-times differentiable in V. Then $f + g$ is also n-times differetiable in V and

$$D^n(f + g) = D^n f + D^n g$$

Moreover for all $k \in K$, kf is n-times differentiable in V and

$$D^n(kf) = k\, D^n f$$

Theorem 2 : Let X and Y be Banach spaces over the same field K of scalars and V be an open subset of X. Let $f : V \to Y$ is twice differentiable at a point $v \in V$. Then $D^2 f(v) \in L(X^2, Y)$ is a bilinear symmetric mapping i.e. for all $(x, y) \in X \times X$,

$$D^2 f(v).(x, y) = D^2 f(v)(y, x)$$

Proof : We define a mapping g as follows :

$$g(x, y) = f(v + x + y) - f(v + x) - f(v + y) + f(v)$$

Then clearly g is a symmetri function in (x, y).

Also

$$\left\| g(x,y) - \left(D^2 f(v).y\right).x \right\|$$

$$= \left\| g(x,y) - Df(v+y).x + Df(v).x + Df(v+y).x - Df(v).x - \left(D^2 f(v).y\right).x \right\|$$

$$\leq \left\| g(x,y) - Df(v+y).x + Df(v).x \right\| + \left\| Df(v+y).x - Df(v).x - \left(D^2 f(v).y\right).x \right\| \quad ...(1)$$

As Df is differentiable at $v \in V$ then for given $v \in V$ then for given $\in > 0$ there exists a $\delta > 0$ such that

$$\left\| Df(v+y) - Df(v) - D^2 f(v).y \right\| \leq \in \|y\| \text{ for } \|y\| \leq \delta$$

therefore

$$\left\| Df(v+y).x - Df(v).x - \left(D^2 f(v)y\right).x \right\| \leq \in \|y\| \, \|x\|$$

$$\leq \in \|x\| \left(\|y\| + \|x\| \right)$$

for $\|x\| \leq \frac{\delta}{2}$ and $\|y\| \leq \frac{\delta}{2}$

Now suppose

$$s(x) = f(v+x+y) - f(v+x) - Df(v+y).x + Df(v).x \quad ...(3)$$

From mean volue theorem, we have

$$\left\| s(x) - s(0) \right\| \leq \|x\| \sup\left\{ \left\| s'(tx) \right\| : t \in [0,1] \right\}$$

From (3), we get

$$s'(x) = Df(v+x+y) - Df(v+x) - Df(v+y) + Df(v)$$

Using it in above, we have

$$\left\| s(x) - s(0) \right\| \leq \|x\| \sup\left\{ \left\| Df(v+tx+y) - Df(v+tx) - Df(v+y) + Df(v) \right\| : t \in [0,1] \right\} \quad ...(4)$$

Now

$$\left\| Df(v+tx+y) - Df(v+tx) - Df(v+y) + Df(v) \right\|$$

$$= \left\| \left\{ Df(v+y+tx) - Df(v) - D^2 f(v)(y+tx) \right\} - \left\{ Df(v+tx) - Df(v) - D^2 f(v).(tx) \right\} - \left\{ Df(v+y) - Df(v) - D^2 f(v).y \right\} \right\|$$

$$\leq \in (\|y+tx\|) + \in \|y\| + \in \|tx\|$$

Since

$$\|y+tx\| \leq \|y\| + \|tx\|$$

$$\leq \|y\| + \|x\| \qquad \because t \in (0,1) \Rightarrow \|tx\| \leq \|x\|$$

Using it in above, we have

$$\|Df(v+y+tx) - Df(v+tx) - Df(v+y) + Df(v)\|$$

$$\leq \in (\|y\| + \|x\|) + \in \|y\| + \in \|x\|$$

$$= 2 \in (\|x\| + \|y\|)$$

Using it in (4), we get

$$\|s(x) - s(0)\| \leq \|x\|.2 \in (\|x\| + \|y\|)$$

Substituting the values of $s(x)$ and $s(0)$ from (3), we have

$$\|f(v+x+y) - f(v+x) - Df(v+y).x + Df(v)x - f(v+y) - f(v)\|$$

$$\leq 2 \in \|x\| (\|x\| + \|y\|) \qquad ...(5)$$

Now from (1) and (2), we get

$$\|g(x,y) - (D^2 f(v).y).x\| \leq \|f(v+x+y) - f(v+x) - f(v+y)$$

$$+ f(v) - Df(v+y).x + Df(v).x\|$$

$$+ \in \|x\| (\|x\| + \|y\|)$$

Using (5), we obtain

$$\|g(x,y) - (D^2 f(v).y).x\| \leq 2 \in \|x\| (\|x\| + \|y\|)$$

$$+ \in \|x\| (\|x\| + \|y\|)$$

$$\Rightarrow \|g(x,y) - (D^2 f(v).y).x\| \leq 3 \in \|x\| (\|x\| + \|y\|) \qquad ...(6)$$

Interchanging x and y, we obtain

$$\|g(y,x) - (D^2 f(v).x).y\| \leq 3 \in \|y\| (\|x\| + \|y\|) \qquad ...(7)$$

Now,

$$\left\| \left(D^2 f(v).x\right).y - \left(D^2 f(v).y\right).x \right\|$$

$$= \left\| g(x,y) - \left(D^2 f(v).y\right).x - g(y,x) + \left(D^2 f(v).x\right).y \right\|$$

$$\leq \left\| g(x,y) - \left(D^2 f(v).y\right).x \right\| + \left\| g(y,x) - \left(D^2 f(v).x\right).y \right\|$$

$$\leq 3 \in \|x\| \left(\|x\| + \|y\|\right) + 3 \in \|y\| \left(\|x\| + \|y\|\right)$$

$$= 3 \in \left(\|x\| + \|y\|\right)^2 \qquad \text{From (6) and (7)}$$

But $\in$ is an arbitrary small, therefore

$$\left(D^2 f(v).x\right).y = \left(D^2 f(v).y\right)x$$

$$\Rightarrow D^2 f(v)(x,y) = D^2 f(v)(y,x)$$

Proved

Theorem 3 : Let X and Y be Banach spaces over the same field K of scalars and V be an open subset of X. Suppose $f : V \to Y$ be an n-times differentiable function on V. Then for each permutation p of $\underline{n}$ and each point $(x_1, x_2, ...x_n) \in X^n$ and each $v \in V$,

$$D^n f(v)\left(x_{p(1)}, x_{p(2)}, ..., x_{p(n)}\right) = D^n f(v)(x_1, x_2, ..., x_n)$$

Proof : We shall prove this result by induction on n. For n = 2, this reduces to theorem (2) i.e.

$D^2 f(v)(x_1, x_2) = D^2 f(v)(x_2, x_1)$, which we have already proved.

Let us assume that the result is true for $(n-1)$ i.e. $D^{n-1} f(v)$ is a symetric member of $L\left(X^{n-1}, Y\right)$

Now suppose $x_1 \in X$, then

for $(x_2, ..., x_m) \in X^{n-1}$, we have

$$D^n f(v).(x_1, x_2, ..., x_n) = D\left(D^{n-1} f(v).x_1\right)(x_2 ....., x_n)$$

Now we know that each permutation of $\underline{n}$ is a composition of consecutive transpoitions $(r, r+1)$ of $\underline{n}$. Since by hypothesis $D^{n-1} f(v)$ is a symmetric function of X^{n-1} into Y, for $r = 2, 3, ...n.$

$$\therefore \quad D^n f(v).(x_1, x_2, ...x_r, x_{r+1} ...., x_n) = D^n f(v).(x_1, x_2 ..., x_{r+1}, x_r ....x_n)$$

So now, it is sufficient to show that

$$D^n f(v).(x_1, x_2 ...., x_n) = D^n f(v).(x_2, x_1 ....., x_n)$$

But we know that $D^n f(v) = D^2(D^{n-2} f)(v)$, and so that

$$(D^n f(v).x_1).x_2 = (D^n f(v).x_2).x_1$$

Consqeuently

$$D^n f(v).(x_1, x_2, ..., x_n) = D^n f(v).(x_2, x_1, ..., x_n)$$

Proved

10.4 Taylor's Theorem

Theorem 4 : Let f be a function defined on the interval $[a,b]$ of R into R such that f is m times differentiable in $[a,b]$ and $(m+1)$ times differentiable in interval (a,b). Then

$$f(b) = f(a) + (b-a)Df(a) + ... + \frac{(b-a)^m}{m!} D^m f(a) + \frac{(b-a)^{m+1}}{(m+1)!} D^{m+1} f(c)$$

Where $c \in (a,b)$

Proof : Given that $f:[a,b] \to R$ be a function. We define a function g on $[a,b]$ as follows :

$$g(x) = f(b) - f(x) - (b-x)Df(x) ...... \frac{(b-x)^m}{m!} D^m f(x)$$

$$- A\frac{(b-x)^{m+1}}{(m+1)!} \qquad \forall x \in [a,b] \qquad ...(1)$$

Where A is a constant can be determined by putting

$$g(a) = g(b)$$

Put $x = b$ in eq^n (1), we get

$g(b) = 0 \Rightarrow g(a) = 0$

Put $x = a$ in eq^n (1), we obtain

$$g(a) = f(b) - f(a) - (b-a)Df(a) ..... \frac{(b-a)^m}{m!} D^m f(a)$$

$$- A.\frac{(b-a)^{m+1}}{(m+1)!}$$

$$\Rightarrow f(b) = f(a) + (b-a)Df(a) + ... + \frac{(b-a)^m}{m!} D^m f(a) + A.\frac{(b-a)^{m+1}}{(m+1)!} \qquad ...(2)$$

$\because g(a) = 0$

Now from eq^h (1), it is clear that

(i) $g(x)$ is continuous in $[a,b]$

(ii) $g(x)$ is differentiable in (a,b) and

(iii) $g(a)=g(b)$

Hence by Rolle's theorem there exists $c\in(a,b)$ such that

$g'(c)=0$

Differentiate eq^n (1) w.r. $t\ \ x$, we get

$$g'(x)=-f'(x)+f'(x).....-\frac{(b-x)^m}{m!}D^{m+1}f(x)+A.\frac{(m+1)(b-x)^m}{(m+1)!}$$

putting $x=c$, we have

$$g'(c)=-\frac{(b-c)^m}{m!}D^{m+1}f(c)+A.\frac{(b-c)^m}{m!}$$

$\Rightarrow\ \ A=D^{m+1}f(c)$ $\qquad\qquad\because g'(c)=0$ and $b-c\neq 0$

Substituting the value of A in (2), we get

$$f(b)=f(a)+(b-a)Df(a)+....+\frac{(b-a)^m}{m!}D^m f(a)$$

$$+\frac{(b-a)^{m+1}}{(m+1)!}D^{m+1}f(c)$$

Where $c\in(a,b)$

Theorem 5 : Let X he a Banach space over the field K of scalars, and let I be an open interval in R containing $[0,1]$. If $\psi:I\to X$ is $(n+1)$ times continuously differentiable function of a single variable $t\in I$. Then

$$\psi(1)=\psi(0)+\psi^{|}(0)+\frac{\psi^{n}(0)}{2!}...+\frac{\psi^{n}(0)}{n!}+\int_0^1\frac{(1-t)^n}{n!}\psi_{(t)}^{n+1}dt$$

Proof : We know that if the function f on $[0,1]$ has a continuous derivative f', then

$$f(1)-f(0)=\int_0^1 f'(t)dt \qquad ...(1)$$

We define a function f on I as follows :

$$f(t) = \psi(t) + (1-t)\psi'(t) + ...+ \frac{(1-t)^n}{n!}\psi^n(t) \qquad ...(2)$$

$$\therefore \quad f'(t) = \psi'(t) + (1-t)\psi''(t) - \psi'(t) + ...+ \frac{(1-t)^n}{n!}\psi^{n+1}(t)$$

$$\Rightarrow \qquad \frac{(1-t)^n}{n!}\psi^{n+1}(t) = f'(t)$$

$$\Rightarrow \qquad \int_0^1 \frac{(1-t)^n}{n!}\psi^{n+1}(t)dt = \int_0^1 f'(t)dt$$

Using eq^n (1), we get

$$f(1) - f(0) = \int_0^1 \frac{(1-t)^n}{n!}\psi^{n+1}(t)dt$$

Using (2), we get

$$\psi(1) - \psi(0) - \psi'(0) - \frac{\psi''(0)}{2!} ..... - \frac{\psi^n(0)}{n!} = \int_0^1 \frac{(1-t)^n}{n!}\psi^{n+1}(t)dt$$

$$\Rightarrow \qquad \psi(1) = \psi(0) + \psi'(0) - \frac{\psi''(0)}{2!} + ...+ \frac{\psi^n(0)}{n!} + \int_0^1 \frac{(1-t)^n}{n!}\psi^{n+1}(t)dt$$

Proved

Theorem 6 : Let X be a Banach space over a field K of scalars and let I be an open interval in R containing $[0,1]$.

If $\psi : I \to X$ is an $(n+1)$ times differentiable function of a single variable $t \in I$ and if $\|\psi^{n+1}(t)\| \le M$ for $t \in [0,1]$.

Then

$$\left\| \psi(1) - \psi(0) - \psi'(0) - \frac{\psi''(0)}{2!} ..... - \frac{\psi^n(0)}{n!} \right\| \le \frac{M}{(n+1)!}$$

Proof : We define two functions $f : [0,1] \to X$ and $g : [0,1] \to R$ as follows :

$$f(t) = \psi(t) + (1-t)\psi'(t) + ...+ \frac{(1-t)^n}{n!}\psi^n(t) \qquad ...(1)$$

and,

$$g(t) = \frac{-M(1-t)^{n+1}}{(n+1)!} \qquad \forall\, t \in [0,1] \qquad ....(2)$$

From equation (1), we have

$$\|Df(t)\| = \left\| \frac{(1-t)^n}{n!} \psi^{n+1}(t) \right\|$$

$$= \frac{(1-t)^n}{n!} \left\| \psi^{n+1}(t) \right\|$$

$$\|Df(t)\| \le \frac{(1-t)^n}{n!} . M \qquad ....(3)$$

From equation (2), we have

$$Dg(t) = \frac{M(1-t)^n}{(n+1)!} .(n+1)$$

$$\Rightarrow \qquad Dg(t) = \frac{M(1-t)^n}{n!} \qquad ...(4)$$

From (3) and (4), we have

$\|Df(t)\| \le Dg(t)$, for all $t \in \{0,1\}$

Now we know that if $f:[0,1] \to Y$ and $g:[0,1] \to R$ are continuous and differentiable functions such that $\|Df(t)\| \le Dg(t), \ \forall\, t \in (0,1)$, then we have

$\|f(1) - f(0)\| \le g(1) - g(0)$, by theorem (6) unit (9)

Using (1) and (2), we get

$$\left\| \psi(1) - \psi(0) - \psi'(0) - \frac{\psi''(0)}{2!} .... \frac{\psi^n(0)}{n!} \right\| \le 0 - \left(- \frac{M}{(n+1)!} \right)$$

$$\Rightarrow \qquad \left\| \psi(1) - \psi(0) - \psi'(0) - \frac{\psi''(0)}{2!} .... - \frac{\psi^n(0)}{n!} \right\| \le \frac{M}{(n+1)!}$$

Theorem 7 (Taylor's formula with Lagrange's Reminder) : Let X and Y be Banach space over the same field K of scalars and V be an open subset of X. Let $f: V \to Y$ be an $(n+1)$ times differentiable function. If the interval $[a, a+h]$ is contained in V and if $\|f^{n+1}(x)\| \le M$, $x \in V$. Then

$$\left\| f(a+h) - f(a) - f'(a)h - ... - \frac{f^n(a)}{n!} h^n \right\| \le \frac{M\|h\|^{n+1}}{(n+1)!}$$

Proof : We define a mapping $\phi : [0,1] \to Y$ as follows :

$$\phi(t) = f(a + t\,h),. \qquad \forall\, t \in [0,\, 1] \qquad ...(1)$$

$$\phi^{n+1}(t) = h^{n+1} f^{\,n+1}(a + t\,h)$$

$$\Rightarrow \qquad \left\|\phi^{n+1}(t)\right\| = \left\|h^{n+1} f^{\,n+1}(a + t\,h)\right\|$$

$$\Rightarrow \qquad \left\|\phi^{n+1}(t)\right\| \le M \left\|h\right\|^{n+1} \qquad ...(2)$$

Now suppose that

$$\psi(t) = \phi(t) + (1-t)\phi'(t) + ... + \frac{(1-t)^n}{!n}\phi^n(t), \;\; t \in [0,\, 1] \qquad ...(3)$$

Therefore,

$$\psi'(t) = \frac{(1-t)^n}{n!}\; \phi^{n+1}(t)$$

$$\Rightarrow \qquad \left\|\psi'(t)\right\| = \frac{(1-t)^n}{n!}\; \left\|\phi^{n+1}(t)\right\|$$

$$\Rightarrow \qquad \left\|\psi'(t)\right\| \le \frac{(1-t)^n}{n!}\; M \left\|h\right\|^{n+1} \qquad \text{[From (2)]} \qquad ...(4)$$

We define again $g : [0,1] \to R$ as follows :

$$g(t) = -\, M \frac{(1-t)^n}{(n+1)!}\; \left\|h\right\|^{n+1} \qquad ...(5)$$

so that

$$D\,g(t) = M \frac{(1-t)^n}{n!}\; \left\|h\right\|^{n+1} \qquad ...(6)$$

Using (6) in (4), we get

$$\left\|D\,\psi(t)\right\| \le D\,g(t), \qquad t \in [0,1]$$

Then by mean value theorem, we have

$$\left\|\psi(1) - \psi(0)\right\| \le g(1) - g(0)$$

using (3) and (5), we obtain

$$\left\|\phi(1)-\phi(0)-\phi'(0)-...-\frac{\phi^n(0)}{n!}\right\| \le 0-\left\{-\frac{M.\|h\|^{n+1}}{(n+1)!}\right\}$$

Using (1), we get

$$\left\|f(a+h)-f(0)-h\,f'(0)-...-\frac{h^n}{n!}f^n(0)\right\| \le \frac{M\|h\|^{n+1}}{(n+1)!}$$

Theorem 8 (Taylor's Formula with Integral Remainder) : Let X and Y be Banach space over the same field K of scalars and V be an open subset of X. Suppose $f:V\to Y$ be a function of class C^{n+1}. If the closed interval $[a,a+h]$ is contained in V. Then

$$f(a+h)=f(a)+h\,f'(a)+\frac{h^2}{2!}f''(a)+...+\frac{h^n}{n!}f^n(a)$$

$$+\int_0^1\frac{(1-t)^n}{n!}f^{n+1}(a+t\,h)\,.\,h^{n+1}\,dt$$

Proof : We define a function $\psi:[0,1]\to Y$ as follows :

$$\psi(t)=f(a+t\,h),\ \forall\, t\in[0,1] \qquad ...(1)$$

$$\left.\begin{array}{l}\psi'(t)=h\,f'(a+t\,h)\\ \psi''(t)=h^2 f''(a+t\,h)\\ \quad\vdots \qquad \vdots \qquad \vdots\\ \psi^n(t)=h^n f^n(a+t\,h)\end{array}\right\} \qquad ...(2)$$

Suppose,

$$f(t)=\psi(t)+(1-t)\psi'(t)+...+\frac{(1-t)^n}{n!}\psi^n(t)$$

Since,

$f(1)-f(0)=\int_0^1 f'(t)\,dt$, then by theorem (5), we have

$$\psi(1)-\psi(0)-\psi'(0)-...-\frac{\psi^n(0)}{n!}=\int_0^1\frac{(1-t)^n}{n!}\psi^n(t)\,dt$$

Using (1) and (2), we obtain

$$f(a+h)=f(0)+h\,f'(0)+\frac{h^2}{2!}f''(0)+...+\frac{h^n}{n!}f^n(0)$$

$$+\int_0^1\frac{h^{n+1}}{n!}(1-t)^n f^{n+1}(a+t\,h)\,dt$$

10.5 Existence Theorems on Differentiable Functions

In this section we shall prove the implicit function theorem and the inverse function theorem.

Theorem 9 (Implicit function Theorem) : Let X, Y and Z be Banach space over field K, let f be a continuous function on an open subset W of $X \times Y$ into Z such that at each point $(x, y) \in W$ the partial derivative $D_2 f(x, y)$ exists and $D_2 f$ is a continuous function on W into $L(Y, Z)$ and let $(u, v) \in W$ be such that $f(u,v) = 0$ and $D_2 f(u,v)$ is a linear homoeomorphism of Y onto Z. Then there exists an open neighbourhood U of u and an open neighbourhood V of v such that $U \times V \subset W$ and a unique continuous function g on U into V such that

$$g(u) = v \text{ and for each } x \in U,\ f(x, g(x)) = 0.$$

If f is differentiable at (u,v) then g is differentiable at u and

$$D g(u) = -(D_2 f(u,v))^{-1} o D_1 f(u,v).$$

Proof : We define a function $h : W \to Y$ as follows :

$$h(x, y) = y - (D_2 f(u,v))^{-1}(f(x, y)), \quad \forall (x, y) \in W \qquad ...(1)$$

$$\Rightarrow \qquad h(x, y) = y \text{ iff } f(x, y) = 0. \qquad ...(2)$$

h is continuous in W, therefore $D_2 h(x, y)$ exists at each $(x, y) \in W$ and $D_2 h$ is continous on W.

Let U' be an open ball with centre u and radius $\in$ and V' be a closed ball with centre v and radius $\in$ such that

$$U' \times V' \subset W \text{ and for all } (x, y) \in U' \times V'$$

$$\|D_2 h(x, y)\| \le \frac{1}{2} \qquad ...(3)$$

Now by mean value theorem,

$$\|h(x, y) - h(x, y')\| \le \frac{1}{2}\|y - y'\| \quad \forall x \in U',\ y, y' \in V'$$

Let U'' is an open bell with centre u and it contained in U', then

$$\|h(x, v) - v\| \le \frac{\in}{2} \qquad ...(4)$$

Now,

$$\|h(x, y) - v\| = \|h(x, y) - h(x, v) + h(x, v) - v\|$$

$$\le \|h(x, y) - h(x, v)\| + \|h(x, v) - v\|$$

$$\leq \frac{1}{2}\|y-v\|+\frac{\in}{2} \qquad \text{[From (4) and (5)]}$$

$\Rightarrow \quad \|h(x,y)-v\| \leq \in \qquad$ [As $y, v \in V'$, $\Rightarrow \|y-v\| \leq \in$, $\because \in$ is the radius of V']

Since V' is closed and Y is complete and $V' \subset Y$, therefore V' is also complete.

Then by Banach fixed point theorem there exists a unique linear transformation $g' : U'' \rightarrow V'$ such that

$$h(x, g'(x)) = g'(x)$$

Using (2), we obtain

$$f(x, g'(x)) = 0 \qquad \text{...(6)}$$

Let V be the interior of V' and let $U = g^{-1}(V)$ and let g be the restriction of g' to the set U, then

$$g'(x) = g(x), \ x \in V$$

Then, we have

$$f(x, g(x)) = 0 \text{ and } g(u) = v.$$

Now let f is differentiable at (u,v) and x be any element in X such that $u+x \in U$

Let $y = g(u+x) - g(u)$, then

$$f(u+x, g(u+x)) = 0$$

$$\Rightarrow \quad f(u+x, g(u)+y) = 0 \qquad \text{...(7)}$$

As f is differentiable at (u,v), therefore given $\in > 0$, there exists $\delta > 0$ such that $\|x\| \leq \delta$.

Therefore,

$$\|f(u+x, g(u)+y) - f(u, g(u)) - px - qy\| \leq \in (\|x\| + \|y\|)$$

where $\quad p = D_1 f(u, g(u))$

$$q = D_2 f(u, g(u))$$

then we have

$$\|px + qy\| \leq \in (\|x\| + \|y\|) \qquad \text{...(8)}$$

Since $f(u+x, g(u+x)) = 0 = f(u, g(u))$

Given that $q = D_2 f$ is a linear homeomorphism, therefore

$$\left\|\left(q^{-1}op\right)x+y\right\| \le \left\|q^{-1}\right\| \, \|px+qy\|$$

$$\le \in \left\|q^{-1}\right\| \left(\|x\|+\|y\|\right) \qquad ...(9)$$

Let $\in \left\|q^{-1}\right\| = \dfrac{1}{2}$

Now,

$$\|y\| = \left\|y+\left(q^{-1}op\right)x-\left(q^{-1}op\right)x\right\|$$

$$\le \left\|y+\left(q^{-1}op\right)x\right\| + \left\|q^{-1}op\right\| \, \|x\|$$

$\Rightarrow \quad \|y\| \le \dfrac{1}{2}\left(\|x\|+\|y\|\right)+\left\|q^{-1}op\right\| \, \|x\| \qquad$ from (9)

Thus

$$\|x\|+\|y\| \le 2\|x\| \left(1+\left\|q^{-1}op\right\|\right) \qquad ...(10)$$

Now,

$$\left\|g(u+x)-g(u)-\left\{-\left(q^{-1}op\right)x\right\}\right\|$$

$$= \left\|y+\left(q^{-1}op\right)x\right\|$$

$$\le \in \left\|q^{-1}\right\| \left(\|x\|+\|y\|\right) \qquad \text{[From (9)]}$$

$$\le \in \left\|q^{-1}\right\| \, 2\|x\| \left(1+\left\|q^{-1}op\right\|\right) \qquad \text{[From (10)]}$$

$$\Rightarrow \quad \frac{\left\|g(u+x)-g(u)-\left\{-\left(q^{-1}op\right)\right\}x\right\|}{\|x\|} \le 2 \in \left\|q^{-1}\right\| \left(1+\left\|q^{-1}op\right\|\right)$$

But $\in$ is arbitrary number therefore g is differentiable and

$$D\,g(u) = -q^{-1}op$$

$$= -\left\{D_2\, f(u,v)\right\}^{-1} o\, D_1\, f(u,v)$$

Theorem 10 (Inverse Function Theorem) : Let X and Y be Banach spaces over the same field K of scalars and W be an open subset of X. Let $w \in W$ be such that $D\,f(w)$ is a linear homeomorphism of X into Y. Then there exists an open neighbourhood U of w contained in W and an open neighbourhood V of $f(w)$ contained in Y such that f' the restriction of f to the set U is C^1 homeomorphism of U onto V, its inverse in a C^1 homeomorphism of V onto U, and

$$D f'^{-1}(f(w)) = (D f'(w))^{-1}$$

If $D f(x)$ is a linear homeomorphism of X into Y, for all $x \in W$, then f is an open mapping of W into Y. If $D f(x)$ is a linear homeomorphism of X onto Y for all $x \in W$ and f is injective then f is a C^1 homeomorphism of $f(W)$ onto W.

Proof : We define a function $h : W \times Y \to Y$ as follows :

$$h(x, y) = f(x) - y$$

Then $D_1 h(x, y) = D f(x)$ and $D_2 h(x, y) = -I_y$

$$\forall (x, y) \in W \times Y$$

$\Rightarrow$ h is a C^1 map on $W \times Y$

Then by implicit function theorem, there exists an open neighbourhood U' of w contained in W, an open neighbourhood V of $f(w)$ contained in Y and a C^1 map $g : V \to U'$ such that $f(g(y)) = y$, $\forall y \in V$ and $g(f(w)) = w$.

We take $U = g(V)$

Then $U \subset U'$, g is a bijection of V onto U and $U = U' \cap f^{-1}(\mathrm{v})$, which is an open subset of X.

Let f' is an inverse of g, and f' is a C^1 homeomorphism of U and V, g is a C^1 homeomorphism of V onto U and $D f'^{-1}(f(w)) = (D f'(w))^{-1}$

Now suppose $f(x)$ is a linear homeomorphism of X onto Y, $\forall x \in W$. Then by the first part, for each $x \in W$, there is an open neighbourhood U of x contained in W, such that restriction of f to U is a homeomorphism of U onto its image. Hence f is an open mapping of W into Y.

Moreover, let f is also injective. Then f is a bijection of W onto $f(W)$ and so that a homeomorphism of W onto $f(W)$.

Self-Learning Exercise

1. Define C^1 map
2. Define higher derivatives

10.6 Summary

In this unit we studied higher derivatives of functions defined on Banach spaces. We also studied the Taylor's theorem and existence theorems on differentiable function.

10.7 Answers to Self-Learning Exercise

1. See text 2. See text

10.8 Exercises

1. Let $f : W \to Y$, where W is an open subset of the product $X_1 \times X_2 \times ... \times X_n$ of Banach spaces $X_1, X_2, ..., X_n$ over field K such that f is twice differentiable at $w \in W$. Then for

$$i, j = 1,2,........,n$$

$$D_i\left(D_j f\right)(w) = D_j\left(D_i f\right)(w)$$

2. Let X and Y be Banach spaces over field K and let f be a C^n-map of an open subset W of X into $L(X,Y)$. Then the map $(w,x) \to (f(w),x)$ is also C^n-map.

□□□

Unit - 11
The Integral in a Banach Space

Structure of the Unit

11.0 Objectives

In this unit, we introduce integral of a regulated function through step function and discuss some of its basic properties.

11.1 Introduction

At elementary stage, the subject of integration is generally introduced as the inverse of differentiation, so that a function F is called an integral of a given function f if $F'(x) = f(x)$, for all values of x belonging to the domain of the function f. The reference to integration from summation point of view was always associated with the geometric concepts. To formulate an independent theory of integration, the German mathematician, Riemann, gave a purely arithmatic treatment to the subject and developed the subject entirely free from the intuitive dependence on geometrical concepts.

The Riemann integral depends very explicitly on the order structure of the real line. Accordingly, we have studied the integration of real valued function of real variable in under-gradark courses. In this unit, we courider the integral of a function of one variable into a Banach space. To study integration of such functions, we take slightly different approach than for real valued functions of a real variable. First we detive the integral of a regulated function through step functions and then prove some basic properties of the integrals.

11.2 Subdivision

Let $[a,b]$ be a compact interval of the real line. A set of points $\{a_1, a_2, ..., a_n\}$ of $[a,b]$ is called a subdivision of $[a,b]$

if, $\quad a = a_1 \le .... \le a_n = b$.

The subdivision consists of n poitns. A subdivision S_2 of $[a,b]$ is said to be refinement of a subdivision S_1 of $[a,b]$ iff each point of S_1 is a point of S_2 i.e, $S_1 \subset S_2$. Let S_1 and S_2 be two subdivisions of $[a,b]$, then there exists a unique subdivision $S(= S_1 \cup S_2)$ whose points are the point of S_1 or S_2 and which is a refinement of both S_1 and S_2.

11.3 Step Function

Let $[a,b]$ be compact interval of R and let X be a Banach space over K. Then a function $f:[a,b] \to X$ is called a **step function** with respect to a subdivision $(a_i : i \in \underline{n})$ of $[a,b]$ iff for each i in $\underline{n-i}$, $f(a_i, a_{i+1})$ is a singleton. We say that f is a **Step function** on $[a,b]$ into X iff it is a step function with respect to some subdivision of $[a,b]$.

Thus a function f on $[a,b]$ into X is a step function on $[a,b]$ into X iff there exists a subdivision $(a_i : i \in \underline{n})$ of $[a,b]$, $n \geq 2$ and there exists a list $(x_i : i \in \underline{n-1})$ of points of X such that for each i in $\underline{n-1}$ and each t in (a_i, a_{i+1}), $f(t) = x_i$.

11.4 Integral of a Step Function

Let f be a step function on compact interval $[a,b]$ of R into a Banach space X. Let $S = (a_i : i \in \underline{n})$ be subdivision of $[a,b]$ such that f is a step function with respect to S. For each i in $\underline{n-1}$, let x_i be a point of X such that foll all t in (a_i, a_{i+1}), $f(t) = x_i$. Then we put

$$I_S(f) = \sum_{i=1}^{n-1} (a_{i+1} - a_i)\, x_i$$

Now let j be a fixed element of $\underline{n-1}$ and let c_j be any point of (a_j, a_{j+1}). Then

$$S_1 = (a_1, \ldots, a_j, c_j, a_{j+1}, \ldots, a_n)$$

is a subdivision of $[a,b]$ such that S_1 is refinement of S. Moreover

$$I_{S_1}(f) = (a_2 - a_1)x_1 + \ldots + (c_j - a_j)x_j + (a_{j+1} - c_j)x_j + \ldots$$

$$+ (a_n - a_{n-1})x_{n-1} = I_S(f)$$

Now let T be a subdivision of $[a,b]$ such that T is a refinement of S. Then by induction,

$$I_S(f) = I_T(f)$$

Finally let U be any other subdivision of $[a,b]$ with respect to which f is a step function. Then by definition 11.2, there exists a subdivision V of $[a,b]$ such that V is a refinement of both S and U. Hence

$$I_S(f) = I_U(f) = I_V(f).$$

Consequently, we define the **integral of a step function** f on $[a,b]$ into X as the vector $\sum_{i=1}^{n-1}(a_{i+1}-a_i)\,x_i$,

where $(a_i : i \in \underline{n})$ is a subdivision of $[a,b]$ such that f is a step function with respect to this subdivision, and where for each i in $\underline{n-1}$, there is a vector $x_i \in X$ such that

$$f(t) = x_i \text{ for all } t \in (a_i, a_{i+1})$$

and denote it by $\int_a^b f$ or $\int_a^b f(t)\,dt$

Evidently

$$\left\| \int_a^b f \right\| \le \sum_{i=1}^{n-1} |a_{i+1}-a_i|\,\|x_i\|$$

$$\le \sum_{i=1}^{n-1} |a_{i+1}-a_i| \sup\{\|f(t)\| : t \in [a,b]\}$$

$$= (b-a)\sup\{\|f(t)\| : t \in [a,b]\}$$

$$= (b-a)\,\|f\|$$

Now if we consider the set $S([a,b],X)$ of all step functions on a compact internal $[a,b] \subset R$ into a Banach space X then this set tuns out to be a Banach space with norm as stated above. This is evident form the following theorem.

Theorem 1 : Let $[a,b]$ be a compact interval of R and let X be a Banach space over K. Then the set $S([a,b],X)$ of all step functions on $[a,b]$ into X is a vector subspace of the Banach space $B([a,b],\, X)$ of all bounded functions on $[a,b]$ into X with Sup. norm

$$f \to \|f\| = \sup\ \{\|f(t)\| : t \in [a,b]\}$$

and the map $f \to \int_a^b f$ is a continuous linear map of $S([a,b],X)$ into X.

Proof : Let $f, g \in S\,([a,b],\, X)$ i.e., f and g be step functions on $[a,b]$ into X with respect to subdivisions S and T of $[a,b]$ respectively. Let $U = (a_i : i \in \underline{n})$ be a refinement of both S and T, then f and g are step functions with respect to U also and so for each i in $\underline{n-1}$ there are vectors x_i and y_i in X such that for each $t \in (a_i, a_{i+1})$, $f(t) = x_i$ and $g(t) = y_i$

$$f(t) + g(t) = x_i + y_i$$

i.e., $\quad (f+g)t = x_i + y_i$ for each $t \in (a_i, a_{i+1})$

Hence $f+g$ is a step function on $[a,b]$ into X with respect to U and so $f+g \in S([a,b],X)$.

Also for each $\alpha \in K$, αf is a step function on $[a,b]$ into X with respect to S i.e. $\alpha f \in S([a,b],\, X)$.

For each f in $S([a,b],\, X)$, $I_m(f)$ is a finite subset of X and so it is bounded. Thus $S([a,b],\, X)$ is a vector subspace of $B([a,b],\, X)$. The function $f \to \int_a^b f$ on $S([a,b],\, X)$ into X is clearly linear and continuous with a norm less than or equal to $b-a$.

11.5 Regulated Function

Let $[a,b]$ be a compact interval of R and X be a Banach space over K. Then a member of the closure of the vector subspace $S([a,b],\, X)$ of all step functions an $[a,b]$ into X in the Banach space $B([a,b],X)$ is called a **regulated function** on $[a,b]$ into X.

The unique continuous linear extension of the map $f \to \int_a^b f \in X$, $f \in S([a,b],X)$ to closure of $S([a,b],X)$ will be denoted by the same symbol $f \to \int_a^b f$ and for each regulated function f on $[a,b]$ into X, $\int_a^b f$ will be called the **integral** of f.

The class of regulated functions is larger than the class of continuous functions. In support of this we have the following theorem.

Theorem 2 : Let f be a function on a compact interval $[a,b]$ of R into a Banach space X over K. Then f is regulated iff the following conditions are satisfied.

(i) for each point $c \in [a,b)$

$$\lim_{\substack{t\to c\\ t>c}} f(t) \text{ exists}$$

(ii) for each point $c \in (a,b]$

$$\lim_{\substack{t\to c\\ t<c}} f(t) \text{ exists.}$$

In particular, if f is continuous, then f is regulated.

Proof : First by, let f be regulated. Let r be any positive real number, then there is a g in $S([a,b],X)$ such that

$$\|f-g\| \le \frac{r}{3}$$

Let c be any point of $[a,b]$. Since g is a step function on $[a,b]$, there is a d in $[a,b]$ such that $c<d$ and for each t and t^1 in $[c,d]$ $\|g(t)-g(t^1)\| \leq \frac{r}{3}$.

and so,

$$\|f(t)-f(t')\| \leq \|f(t)-g(t)\| + \|g(t)-g(t')\| + \|g(t')-f(t')\| \leq \frac{r}{3}+\frac{r}{3}+\frac{r}{3}=r$$

Hence

$\lim\limits_{\substack{t\to c\\ t>c}} f(t)$ exists, as X is complete.

Similarly, we can prove that $\lim\limits_{\substack{t\to c\\ t<c}} f(t)$ exists.

Next suppose that f satisfies condition (i) and (ii). Let r be any positive real number. Then for each c in $[a,b]$, there exists real number $p(c)$ and $q(c)$ such that the open interval $(p(c),q(c))$ contains c and for all pairs of point t and t' in $[a,b]\cap(p(c),c)$ or $[a,b]\cap(c,q(c))$

$$\|f(t)-f(t')\| \leq r.$$

Since $[a,b]$ is compact, there exists a finite subset C of $[a,b]$ such that $[a,b] \subset \underset{c\in C}{U} (p(c),q(c))$.

Let $\{a_i : i \in \underline{n}\}$ be a set of points in the finite set

$$(a,b) \cup \left\{[a,b] \cap \left(\bigcup_{c\in C} (p(c),q(c)) \right)\right\}$$

arranged in increasing order. For each i in $\underline{n-1}$, (a_i,a_{i+1}) is contained in $(p(c),c)$ or $(c,q(c))$ for some c in C and

$$\|f(t)-f(t')\| \leq r$$

for all pairs of points t and t' in (a_i,a_{i+1}).

Define a function g on $[a,b]$ into X such that for each $i \in \underline{n}$, $g(a_i)=f(a_i)$ and for each $i \in \underline{n-1}$ and for each $t \in (a_i,a_{i+1})$, $g(t)=f(s_i)$, where s_i is the middle point of (a_i,a_{i+1}). Then g is a step function on $[a,b]$ into X and for each $t \in [a,b]$,

$$\|f(t)-g(t)\| \leq r$$

Hence $\|f-g\| \leq r$ and so f is regulated.

Remark : (1) Let $[a,b]$ be a compact interval of R and let X be a Banach space for any two regulated functions f and g on $[a,b]$ into X and any scalar $\alpha \in K$

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

$$\int_a^b (\alpha f) = \alpha \int_a^b f, \text{ since the map } f \to \int_a^b f \text{ is linear.}$$

(2) Given any sequence $\{f_n : n \in N\}$ in closure of $(S[a,b], X)$ converging to f in $B([a,b], X)$, then f is in closure of $S([a,b], X)$ and the sequence $\left\{\int_a^b f_n : n \in N\right\}$ in X converses to $\int_a^b f$ in X.

11.6 Basic Properties of Integrals

Theorem 3 : Let f be a regulated function on a compact interval $[a,b]$ of R into a Banach space X over K, and c be any point of $[a,b]$. Then the restriction of f to $[a,c]$ (respectively $[c,b]$) is a regulated function on $[a,c]$ (respectively $[c,b]$) into X and $\int_a^b f = \int_a^c f + \int_c^b f$.

Proof : Let f be a step function on $[a,b]$, then clearly restriction of f to $[a,c]$ (respectively $[c,b]$) is a step function on $[a,c]$ (respectively $[c,b]$) and

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Now let f be a regulated function on $[a,b]$

Then there is a sequence $\{f_n : n \in N\}$ of step function on $[a,b]$ converging to f in $B([a,b], X)$. Then by the aobve remark $\int_a^b f = \int_a^c f + \int_c^b f$.

Theorem 4 : Let f be a regulated function on a compact interval $[a,b]$ of R into a Banach space X over K and g be a continuous linear map of X into a Banach space Y over K. Then gof is regulated and

$$\int_a^b (gof) = g\left(\int_a^b f\right).$$

Proof : Let f be a step function, then clearly gof is also a step function and

$$\int_a^b (gof) = g\left(\int_a^b f\right)$$

Now let f be a regulated function in $[a,b]$. Then there is a sequence $\{f_n : n \in N\}$ of step functions on $[a,b]$ into X converging to f in $B = ([a,b], X)$. For each $n \in N$

$\|gof - gof_n\| \leq \|g\| \, \|f - f_n\|$.

$\Rightarrow$ $\|gof - gof_n\| \to 0$ as $n \to \infty$, since $\|f - f_n\| \to 0$ as $n \to \infty$.

$\Rightarrow$ the sequence $\{g_0 f_n : x \in N\}$ of step functions on $[a,b]$ in to Y converges to gof in $B([a,b], X)$ ...(1)

Here gof is regulated.

By the definition of the integral, the sequence $\left\{\int_a^b f_n : n \in N\right\}$ converges to $\int_a^b f$ in X.

Since g is continuous and linear, the sequence

$$\left\{\left(g\int_a^b f\right) : n \in N\right\} \text{ converges to } g\int_a^b f \text{ in } Y \qquad ...(2)$$

but for each $n \in N$

$$\int_a^b (gof_n) = g\left(\int_a^b f_n\right)$$

Hence $\int_a^b (gof) = g\left(\int_a^b f\right)$

Definition : Let f be a regulated function on a compact interval $[a,b]$ of R into a Banach space X. Let c and d be any points of $[a,b]$ such that $c < d$. Then we define

$$\int_c^d f = -\int_d^c f$$

Theorem 5 : Let f be a regulated function on a compact interval $[a,b]$ of R into a Banach space X. Then at each $t \in [a,b]$, the function $F : [a,b] \to X$, $F(t) = \int_a^t f$, $t \in [a,b]$ is continuous.

Proof : f be a regulated function on $[a,b]$, so there is a sequence $\{f_n : n \in N\}$ of step functions on $[a,b]$ convergin to f in $B([a,b], X)$. Therefore

$$\left\|\int_a^t f_n\right\| \leq (t-a) \, \|f_n\| \; \forall n \in N$$

$$\Rightarrow \quad \left\|\int_a^t f\right\| \leq |(t-a)| \, \|f\| \qquad ...(1)$$

Consequently the function

$F : [a,b] \to X$, $F(t) = \int_a^t f$, $t \in [a,b]$ is continuous.

Theorem 6 : Let f be a continuous function on a comapact interval $[a,b]$ of R into a Banach space X over K. Let F be the function $t \to \int_a^t f$ on $[a,b]$ into X. Let g be any differentiable function on $[a,b]$ into X such that $Dg = f$. Then F is differentiable, $DF = f$ and $\int_a^b f = F(b) - F(a) = g(b) - g(a)$.

Proof : Let c be any point of $[a,b]$ and let t be any real number such that $c + t \in [a,b]$ and $t \neq 0$

Then

$$\int_c^{c+t} f(c)\,dt = t\,f(c),$$

where $f(c)$ is the constant function on $[a,b]$, assigning $f(c)$ to all points of $[a,b]$. Hence

$$\frac{F(c+t) - F(c)}{t} - f(c) = \frac{1}{t}\int_c^{c+t} \left(f(t) - f(c)\right) dt$$

and so by (1) of Theorem 5, we get

$$\left\| \frac{F(c+t) - F(c)}{t} - f(c) \right\| \le \frac{1}{|t|} \cdot \|t\| \, \|f(t) - f(c)\| \to 0 \text{ as } t \to c \; [\because f \text{ is continuous}]$$

Hence $DF(c) = f(c)$, so that

$$DF = f = Dg \qquad \{\because Dg = f \text{ given}\}$$

$\Rightarrow \quad D(F - g) = 0$

$\Rightarrow \quad F - g$ is constant function on $[a,b]$. But $f(a) = 0$, and

Hence $\int_a^b f = F(b) - F(a) = g(b) - g(a)$.

Theorem 7 : Let f be a C^1 map on a compact interval $[a,b]$ into a compact interval $[c,d]$ of R and let g be a continuous function on $[c,d]$ into a Banach space X over K. Then

$$\int_a^b \left(D f(s)\right) g\left(f(s)\right) ds = \int_{f(a)}^{f(b)} g(t)\,dt.$$

Proof : Let $h : [c,d] \to X$ be defined by

$$h(t) = \int_c^t g(u)\,du,\; t \in [c,d]$$

Then by Theorem 6, $Dh = g$ and

$$\int_{f(a)}^{f(b)} g(t)\,dt = h\left(f(b)\right) - h\left(f(a)\right) \qquad \text{...(1)}$$

By chain rule for each $s \in [a,b]$

$$D(hof)(s) = Dh(f(s)) o D f(s)$$

$$= (D f(s)) \; Dh(f(s))$$

$$= (D f(s)) \; g(f(s)) \qquad [\because Dh = g]$$

Hence again by Theorem 6, we have

$$\int_a^b (D f(s)) \; g(f(s)) ds = h(f(b)) - h(f(a))$$

$$= \int_{f(a)}^{f(b)} g(t) dt \qquad \text{from (1)}$$

Theorem 8 : Let U be an open subset of a Banach space X over K, let $[a,b]$ be a compact interval of R, let f be a continuous function on $U \times [a,b]$ into a Banach space Y over K and let $g : U \to Y$ be defined as

$$g(x) = \int_a^b f(x,t) dt, \quad x \in U,$$

then g is continuous. If $D f$ exists as a continuous function on $U \times [a,b]$ into $L(X,Y)$, then g is C^1 map and for each $x \in U$, $D g(x) = \int_a^b D_1 f(x,t) dt$.

Proof : Since f is continuous in $U \times [a,b]$ and $[a,b]$ is compact, for each positive real number r and each point $x \in U$, there is a positive real number r' such that for all $t \in [a,b]$ and for all $x' \in U$ such that $\|x' - x\| \le r'$,

$$\|f(x',t) - f(x,t)\| \le r \qquad \text{...(1)}$$

and so $\|g(x') - g(x)\| \le \int_a^b \|f(x',t) - f(x,t)\| dt$ [by definition of g]

$$\le r(b-a) \qquad \text{[by (1)]}$$

Hence g is conitnuous in U.

Next suppose that $D_1 f$ exists as a continuous function on $U \times [a,b]$. Let r be any positive real number and let x be any point in U. Since $D_1 f$ is continuous in $U \times [a,b]$ and $[a,b]$ is compact, there is a positive real number t' such that for all $x \in [a,b]$ and for all $x' \in U$ such that $\|x' - x\| \le r'$.

$$\|D_1 f(x',t) - D_1 f(x,t)\| \le r$$

Then for all $t \in [a,b]$ and for all $x' \in X$ such that $\|x' - x\| \le r'$.

$$\|f(x+x',t)-f(x,t)-D_1 f(x,t).x'\| \le r\|x'\| \qquad ...(2)$$

and so

$$\left\|g(x+x')-g(x)-\int_a^b D_1 f(x,t).x'dt\right\|$$

$$=\left\|\int_a^b f(x+x',t)\,dt-\int_a^b f(x,t)\,dt-\int_a^b D_1 f(x,t)x'\,dt\right\| \qquad \text{[by def. of } g\text{]}$$

$$=\left\|\int_a^b \left(f(x+x',t)-f(x,t)-D_1 f(x,t).x'\right)\,dt\right\|$$

$$\le r\|x'\|(b-a) \qquad \text{[by (2)]}$$

But by Theorem 4 for all $t\in[a,b]$ and for all $x'\in U$

$$\int_a^b D_1 f(x,t)x'dt=\left(\int_a^b D_1 f(x,t)dt\right)x'$$

as $u\to u(x')$ is continuous and linear function on $L(X,Y)$.

Hence $Dg(x)=\int_a^b D_1 f(x,t)\,dt$.

Theorem 9 : Let f be a regulated function on a compact interval $[a,b]$ of R into R such that $a<b$ and for all t in $[a,b]$, $f(t)\ge 0$. Then $\int_a^b f(t)\,dt\ge 0$.

If f is continuous at a point c of $[a,b]$ and $f(c)>0$, then

$$\int_a^b f(t)\,dt>0$$

Proof : Given f is a regulated function, so there exists a sequence $\{f_n : n\in N\}$ of step functions on $[a,b]$ converging to f in $B([a,b],X)$ such that for each $n\in N$ and for each $t\in[a,b]$, $f_n(t)\ge 0$ and so $\int_a^b f_n(t)\,dt\ge 0$

Hence $\int_a^b f(t)\,dt\ge 0$

Next suppose that f is continuous at a point c of $[a,b]$ and $f(c)>0$. Then there is a positive real number r such that for all $t\in[a,b]$ with $|t-c|<r$ implies that

$$\frac{1}{2}f(c)<f(t).$$

If $c=a$, choose a positive real number $s<r$ such that $[a,a+s]\subset[a,b]$, then

$$\int_a^b f(t)\,dt = \int_a^{a+s} f(t)\,dt + \int_{a+s}^b f(t)\,dt$$

$$\geq \frac{1}{2} s\, f(c)$$

$$> 0 \qquad [\because f(c) > 0]$$

If $c \neq a$, choose a positive real number $s < r$ such that $(c-s, c) \subset [a,b]$ then

$$\int_a^b f(t)\,dt = \int_a^{c-s} f(t)\,dt + \int_{c-s}^c f(t)\,dt + \int_c^b f(t)\,dt$$

$$\geq \frac{1}{2} s\, f(c) > 0$$

Theorem 10 : Let f be a continuous function on a compact interval $[a,b]$ of R into the topological dual $X*$ of a Banach space X over R such that $a < b$ and for each C^1-map g on $[a,b]$ into X with $g(a) = g(b) = 0$ and

$$\int_a^b < g(t), f(t) > dt = 0.$$

Then $f(t) = 0$ for each $t \in [a,b]$.

Proof : As a contradiction, let $f(r) \neq 0$ for some $r \in [a,b]$. Since f is continuous, we may suppose that r is different from a and b. Since $f(r) \neq 0$, there is an $x \in X$ such that $< x, f(r) > \neq 0$.

Let $< x, f(r) > > 0$. Since f is continuous, there is a positive real number s such that $[r-s, r+s] \subset [a,b]$ and for each $t \in [r-s, r+s]$, $< x, f(t) >> 0$.

Let h be a C^1-map on $[r-s, r+s]$ into R such that $h \geq 0$ (for example $h(t) = 0$ if $t \notin (r-s, r+s)$ adn $h(t) = \left((t-r)^2 - s^2\right)^2$ if $t \in (r-s, r+s)$.

Let g be the function on $[a,b]$ such that $g(t) = h(t)\,.\,x$ for all $t \in [a,b]$. Then g is a C^1-map on $[a,b]$ into X such that $g(a) = g(b) = 0$ and $< g(t), f(t) >> 0$ for each $t \in (r-s, r+s)$,

Hence by Theorem 9, we have

$$\int_a^b < g(t),\, f(t) > dt > 0,$$

which is not possible, so our assumption was wrong.

Hence $f(t) = 0$ for each $t \in [a,b]$.

Theroem 11 : Let $[a,b]$ be a compact interval, let g be a regulated function on $[a,b]$ into $\{r \in R : r \geq 0\}$ and let h be a continuous function on $[a,b]$ into R such that for all $t \in [a,b]$

$$h(t) \le g(t) + c\int_a^t h(s)\,ds \qquad ...(1)$$

where c is a positive real number. Then for all $t \in [a,b]$

$$h(t) \le g(t) + c\int_a^t g(s)e^{C(t-s)}ds$$

Proof : Let $j : [a,b] \to R$ be defined as

$$j(t) = \int_a^t h(s)\,ds,\ t \in [a,b] \qquad ...(2)$$

Then for each $t \in [a,b]$

$$D\,j(t) = h(t)$$

$$\le g(t) + c\,j(t) \qquad \text{[by (1) and (2)]} \qquad ...(3)$$

Let $k : [a,b] \to R$ be defined as

$$k(t) = j(t)\ e^{-c}(t-a),\ t \in [a,b]$$

Then for each $t \in [a,b]$

$$D\,k(t) = (D\,j(t) - c\,j(t))e^{-c(t-a)}$$

$$\le g(t)e^{-c(t-a)} \qquad \text{[by (3)]}$$

By the definition of j and k, it is clear that

$$j(a) = 0 = k(a)$$

By mean value theorem, for all $t \in [a,b]$, we have

$$k(t) \le \int_a^b g(s)e^{-c(s-a)}ds \qquad \text{Then}$$

$$j(t) = k(t)e^{c(t-a)}$$

$$\le e^{c(t-a)}\int_a^t g(s)\ e^{-c(s-a)}ds \qquad ...(4)$$

Hence by (1) and (2), we have

$$h(t) \le g(t) + c\,j(t)$$

$$\le g(t) + c\,e^{(t-a)}\int_a^t g(s)\ e^{-c(s-a)}ds \qquad \text{[by (4)]}$$

$$= g(t) + c\int_a^t g(s)\ e^{c(t-s)}ds$$

Self-Learning Exercise

1. Write whether the following statements are true or false.

   (a) A subdivision S of a compact interval $[a,b]$ is said to be refinement of a subdivision T of $[a,b]$ iff each point of S is a point of T i.e., $S \subset T$.

(b) The function $f \to \int_a^b f$ is not a continuous linear map of the set $S([a,b], X)$ (the set of all step functions on $[a,b]$ into X) into X.

(c) If f is continuous, then f is regulated.

(d) If any sequence $\{f_n : n \in N\}$ in $Cl\,(S[a,b], X)$ converging to f in $B([a,b], X)$, then $f \in S([a,b], X)$.

(e) If f and g be regulated functions on a compact interval $[a,b]$ of R into a Banach space X over K, then gof is also regulated.

11.7 Summary

In this unit, we have seen that by taking slightly different approach than for real valued function of a real variable, we can find the integral of a function of one variable into a Banach space. We also discuss various properties of such integrals.

11.8 Answers to Self-Learning Exercise

1. (a) False (b) False (c) True
   (d) False (e) False

11.9 Exercises

1. Define integral of a step function. If $[a,b]$ be compact interval of R and X be Banach space. Prove that the set $S([a,b], X)$ of all step functions on $[a,b]$ into X is a vector subspace of the Banach space $B([a,b], X)$ of all bounded functions on $[a,b]$ into X with sup norm $f \to \|f\| = \sup\{\|f(t)\| : t \in [a,b]\}$.

2. Define regulated function. If f be a regulated function on a compact interval $[a,b]$ of R into a Banach space X. Prove that at each $t \in [a,b]$ the function $F : [a,b] \to X$, $F(t) = \int_a^t f$, $t \in [a,b]$ is continuous.

3. Let f be a regulated function defined on a compact interval $[a,b]$ of R into a Banach space X. Show that for each positive real number $\in$, there is a positive real number δ such that for any increasing sequence $a = a_1 \le t_1 \le a_2 \le, \ldots, \le a_i \le t_i \le a_{i+1} \le \ldots \le a_n = b$ of points of $[a,b]$. such that $|a_{i+1} - a_i| \le \delta$, $i \in \underline{n-1}$, $\left\| \int_a^b f(t)\,dt - \sum_{i=1}^{n-1} f(t_i)(a_{i+1} - a_i) \right\| \le \in$.

□□□

Unit - 12
Differential Equations

Structure of the Unit

12.0 Objectives

The present unit is devoted to differential equations. Existence and uniqueness theorems for ordinary differential equations are proved.

12.1 Introduction

In many practical problems we come across with a differential equation which cannot be solved by one of the standard methods known so far. Vaious methods have been formulated for getting to any desired degree of accuracy the numerical solution of the above mentioned type of differential equation with numerical confficients and given conditions. We have studied the Picard's integration method for finding an approximate solution of the initial value of the form

$$\frac{dy}{dx}=f(x,y),\quad y(x_0)=y_0 \qquad ...(1)$$

Theorems which state the conditions under which an initial value problem of the form (1) has at least one solution, only one solution are called existence theorem and uniqueness theorem respectively. The purpose of this unit is to introduce differential equations. Starting with the definitions of a differantial equation and its solution, existence and uniqueness theorems for ordinary differential equations are obtained.

12.2 First order Differential Equations

Throughout this unit X denotes a Banach space over the real field R, f denotes a function of a single real variable t with values in X. Further if f is differentiable, its derivative $\frac{df}{dt}$ will again be considered as a function with values in X.

Definition : Let I be an interval of R, W be a subset of a Bananch space X over K and let g be a continuous map of $I \times U$ into X. Then an equation of the type

$$\frac{dx}{dt} = g(t,x), \quad (t,x) \in I \times W$$

is called a differential equation.

Definition : A differentiable map $f ; I \to W$ is called an integral solution of the differential equation $\frac{dx}{dt} = g(t,x)$ if and only if $Df(t) = g(t, f(t))$ for each $t \in I$. An integral solution of the differential equation is also called an integral solution for g.

Now let (t_0, x_0) be an interior point of $I \times W$. Let I' be an open subset of W containing x_0. Then a differentiable map h : $I' \to W'$ is called an integral solution for g at (t_0, x_0) if $h(t_0) = x_0$ and h is an intgral solution for restriction of g to $I' \times W'$.

A map $f ; I' \times W' \to W$ is called a local flow for g at (t_0, x_0) iff for each $x \in W'$, $f(t_0, x) = x$ and the map $\phi : I' \to W$, $\phi(t) = f(t,x)$ for $t \in I'$ is an integral solution for g at (t_0, x). Thus f is a local flow for g at (t_0, x_0) iff $f(t_0, x) = x$ and $D_1 f(t,x) = g(t, f(t))$ for each $x \in W'$.

Theorem 1 : Let I be an open integral of $R,$ let W be an open subset of a Banach space X over K. Let (t_0, x_0) be point of $I \times W$ and let g be a continuous map of $I \times W$ into X. Then a continuous map h : $I \to W$ is an integral solution for g at (t_0, x_0) iff for each $t \in I$

$$h(t) = x_0 + \int_{t_0}^{t} g(s, h(s)) ds.$$

Proof : Given that g and h are continuous so the map $s \to g(s, h(s))$ of I into X is continuous.

Firstly let h be an integral solution for g at (t_0, x_0), then clearly

$$h(t) = x_0 + \int_{t_0}^{t} g(s, h(s)) ds$$

for each $t \in I$

Next let for each $t \in I$

$$h(t) = x_0 + \int_{t_0}^{t} g(s, h(s)) ds,$$

then h is differentiable in I and its derivative is the map $s \to g(s, h(s))$ and so h is an integral solution for g at (t_0, x_0).

12.3 Approximate Solution

Let $\in > 0$ be a real number. A differentiable map $f : I \to W$ is an approximate solution within $\in$ or an $\in$-approxmate solution for the differential equation $\frac{dx}{dt} = g(t,x)$ if $\|f'(t) - g(t, f(t))\| \le \in$ for all $t \in I$.

12.4 Lipschitz's Property

Let f be a function on a subset W of a Banach space X over K into a Banach space X over K. Let V be any subset of W and let c be any positive real number, then f is said to be $c-$ lipschitz on V iff for all x and x' in V

$$\|f(x) - f(x')\| \le c\|x - x'\|$$

Let I be any subset of R and let g be a function on $I \times W$ into Y. Then g is said to be $c-$ Lipschitz on V uniformly with respect to I iff for all $t \in I$ and all x and x' in V

$$\|g(t,x) - g(t,x')\| \le c\|x - x'\|$$

Now we shall prove a lemma which compares two approximate solutions of a differential equation. We first prove an anxillany lemma.

Lemma 2 : let u be a non-negative continuous function on an interval $\{0,c\},(c>0)$ satisfying the inequality

$$u(t) \le \text{ at} + k\int_0^t u(s)ds \qquad \text{...(1)}$$

for all $t \in [0,c]$, then

$$u(t) \le \frac{a}{k}\left(e^{kt} - 1\right) \text{ for } t \in [0,c]$$

Proof : Let $v(t) = \int_0^t u(s)ds$,

then $v'(t) = u(t), v(0) = 0$ and so inequality (1) reduces to $v'(t) \le \text{at} + k\, v(t)$, ...(2)

which is a differential inquality.

Taking $w(t) = \bar{e}^{kt} v(t)$, then

$$w'(t) = \bar{e}^{kt}\left(v'(t) - kv(t)\right)$$

$$\le \bar{e}^{kt} \text{ at} \qquad \{by(2)\}$$

Since $w(0) = 0$, the mean value inquality gives $w(t) \le \int_0^t \text{ as } \bar{e}^{ks} ds$

Integrating the right hand side by parts, we obtain

$$w(t) \le \frac{a}{k^2}\left(1 - \bar{e}^{kt} - kt\,\bar{e}^{kt}\right) \qquad ...(3)$$

Therefore $v(t) = e^{kt} w(t)$ $\{by(3)\}$

$$\le \frac{a}{k^2}\left(e^{kt} - 1 - kt\right)$$

since $u(t) = v'(t) \le at + kv(t)$, therefore

$$u(t) \le at + \frac{a}{k}\left(e^{kt} - 1 - kt\right)$$

$$= \frac{a}{k}\left(e^{kt} - 1\right)$$

Lemma 3 : (Fundamental Lemma) :

Let I be an open interval of R. Let W be an open subset of a real Banach space X and let g be a continuous map of $I \times W$ into X such that g is c-lipschitz on W uniformly with respect to I, where c is a poritive real number. Let r_1 and r_2 be two positive real numbers such that for all $t \in I$

$$\left\|Df_1(t) - g\left(t, f_1(t)\right)\right\| \le r_1 \text{ and } \left\|Df_2(t) - g\left(t, f_2(t)\right)\right\| \le r_2 \qquad ...(4)$$

i.e., f_1, f_2 are r_1 – approximate solution and r_2 – approximate solution of the equation $\frac{dx}{dt} = g(t, x)$ respectively. Then for all s and t in I

$$\left\|f_1(t) - f_2(t)\right\| \le \left\|f_1(s) - f_2(s)\right\| e^{c|t-s|} + (r_1 + r_2)\left(\frac{e^{c|t-s|} - 1}{c}\right)$$

Proof : We can assume that $s = 0$ and $t > 0$. Then

$$\left\|Df_1(t) - Df_2(t)\right\| = \left\|Df_1(t) - g\left(t, f_1(t)\right) + g\left(t, f_1(t) - g\left(t, f_2(t)\right) + g\left(t, f_2(t)\right) - Df_2(t)\right)\right\|$$

$$\le \left\|Df_1(t) - g\left(t, \delta_{1(t)}\right)\right\| + \left\|Df_2(t) - g\left(t, f_2(t)\right)\right\|$$

$$+ \left\|g\left(t, f_1(t)\right) - g\left(t, f_2(t)\right)\right\|$$

$$\le r_1 + r_2 + \left\|g\left(t, f_1(t)\right) - g\left(t, f_2(t)\right)\right\| \qquad \{by(1)\}$$

$$\le r_1 + r_2 + c\left\|f_1(t) - f_2(t)\right\| \qquad (\because g \text{ is c-Lipschilz on W})$$

Taking $r = r_1 + r_2$ and $f(t) = f_1(t) - f_2(t)$, we have

$$\left\|Df(t)\right\| \le r + c\left\|f(t)\right\|$$

So by mean value inquality for $t > 0$, we have

$$\|f(t) - f(0)\| \le \int_0^t \left(r + c\|f(u)\|\right) du$$

But $\|f(u)\| \le \|f(u) - f(0)\| + \|f(0)\|$

Hence for each u in $[0, t]$, we have

$$\|f(t) - f(0)\| \le \left(r + c\|f(0)\|\right)t + c\int_0^t \|f(u) - f(0)\| dx \qquad ...(2)$$

Putting $\|f(t) - f(0)\| = h(t)$ and $r + c\|\|f(0)\| = b$,

Then (2) reduces to

$$h(t) \le bt + c\int_0^t h(u) du$$

and therefore by Lemma 2, we have

$$h(t) \le \frac{b}{c}\left(e^{ct} - 1\right) \text{ for } t \in I$$

Rewriting the values of $h(t)$ and b, we have

$$\|f(t) - f(0)\| \le \frac{r + c\|f(0)\|}{c}\left(e^{ct} - 1\right)$$

Hence

$$\|f(t)\| \le \|f(t) - f(0)\| + \|f(0)\|$$

$$\le \frac{r + c\|f(0)\|}{c}\left(e^{ct} - 1\right) + \|f(0)\|$$

$$= \frac{r}{c}\left(e^{ct} - 1\right) + \|f(0)\| e^{ct}$$

Hence again rewriting the values of $f(t)$ and r, we have

$$\|f_1(t) - f_2(t)\| \le \frac{r_1 + r_2}{c}\left(e^{c|t-s|} - 1\right) + \|f_1(s) - f_2(s)\| e^{c|t-s|}$$

$$[\because s = 0,\ t > 0]$$

If $x_1 = f_1(s)$ and $x_2 = f_2(s)$ be their initial values at $s \in I$. Then for all $t \in I$, we have

$$\|f_1(t) - f_2(t)\| \le \frac{(r_1 + r_2)}{c}\left(e^{c|t-s|} - 1\right) + \|x_1 - x_2\| e^{c|t-s|} \qquad ....(3)$$

Now we shall make use of the fundamental lemma in proving the following theorems

Theorem 4 : (Uniqueness Theorem)

Let I be an interval in R, W a subset of a Banach space X over K and let $g: I \times W \rightarrow X$ be a continuous function c-Lipsehitz in $x \in X$, If there are two exact solutions, f_1 and $f_2 : I \rightarrow X$ of the differential equation $\frac{dx}{dt} = g(t,x)$ and if $f_1(s) = f_2(s)$ for $s \in I$, then the functions f_1 and f_2 are identical in the interval I.

Proof : Putting $f_1(s) = f_2(s)$, $r_1 = 0, r_2 = 0$ in the inquelity of the fundamental lemma, we have

$$\|f_1(t) - f_2(t)\| = 0 \text{ for } t \in I$$

$$\Rightarrow f_1 = f_2$$

Hence the theorem.

Theorem 5 (Existence Theorem)

Let I be a closed interval in R, W be a closed set in a Banach space X and $g : I \times W \rightarrow X$ be a continuous function which is c-Lipschitz in $x \in X$. Let $(s, x_0) \in I \times W$, for given $r > 0$, let $f: I \rightarrow X$ be an r-approximate solution of the differential equation $\frac{dx}{dt} = g(t,x)$ such that $f(s) = x_0$, then there exists in I an exact solution $\phi: I \rightarrow X$ of the differential equation such that $\phi(s) = x_0$.

Proof : Let $\{r_n\}$ be a decrearing squence of positive real numbers such that $\lim_{n\rightarrow\infty} r_n = 0$. For each $n \in N$, let $f_n: I \rightarrow X$ be an r_n – approximate solution such that $f_n(s) = x_0$. By the fundamental lemma, we have

$$\|f_n(t) - f_p(t)\| \leq \frac{r_n + r_p}{c}\left(e^{c(r-s)} - 1\right) \text{ for all } t \in I.$$

Let $\frac{e^{c(t-s)} - 1}{c} \leq m$ for all $t \in I$. Then

$$\|f_n(t) - f_p(t)\| \leq (r_n + r_p)m$$

and thus $\{f_n\}$ is a cauchy squance. Therefore the squence $\{f_n\}$ has a limit.

Let $\lim_{n\rightarrow\infty} f_n = \phi$, then ϕ is a continuous function $I \rightarrow X$.

Since $(t, f_n(t)) \in I \times U$ for all $n \in N$ all $t \in I$ and since U is a closed set in X, therefore $(t, \phi(t)) \in I \times U$ for all $t \in I$.

Also each f_n is an r_n-approximate solution, therefore

$$\left\|Df_n(t)-g(t,f_n(t))\right\|\le r_n,$$

and so by the mean value inquality, we have

$$\left\|f_n(t)-x_0-\int_s^t g(u,f_n(u))du\right\|\le r_n|t-s|$$

since $\lim_{n\to\infty} f_n(t)=\phi(t)$

and $\lim_{n\to\infty}\int_{t_0}^t g(u,f_n(u))du=\int_s^t g(u,\phi(u))du$

Therefore in the limit $n\to\infty$, the above inquality reduces to

$$\phi(t)=x_0+\int_s^t g(u,\phi(u))du$$

Which gives $\phi'(t)=g(t,\phi(t))$

Hence ϕ is an exact solution of the differential equation with $\phi(s)=f_n(s)=x_0$.

12.5 Locally Lipschitz

Let I be an interval in R, W be a subset of a real Banach space X. A function $g:I\times W\to X$ is locally Lipschitz if for each point $(t_0,x_0)\in I\times W$, there exists a neighbourhood $J\times V$ of $(t_0,x_0)\in I\times W$ and $c>0$ such that

$$\left\|g(t,x_1)-g(t,x_2)\right\|\le c\left\|x_1-x_2\right\|$$

for each $t\in J$ and $x_1,x_2\in V$.

In other words g is locally Lipschitz, if the restriction of g to $J\times V$ is c-Lipschitz in $x\in X$.

Theorem 6 (Global Uniqueness Theorem) :

Let I be an interval in R, W be a subset of a Banach space X and $g:I\times W\to X$ be a locally Lipschitz function. If there are two exact solutions f_1 and f_2 : $I\to X$ of the differential equation $\frac{dx}{dt}=g(t,x)$ and if they are equal for one value $t_0\in I$, then they are identical in the entire I.

Proof : Let J be a subset of I given by

$$J\ \{t\in I: f_1(t)=f_2(t)\}$$

Now we shall establish that the set J is simultaneously open and closed in I.

Since the function f_1-f_2 is continuous, therefore J is a closed set.

Now let $f_1(t_0)=f_2(t_0)=x_0$

Since g is locally Lipschitz, there exists a neigbhbourhood N of (t_0, x_0) in $I \times W$ as well as a real number $k > 0$ such that g is k-Lipschitz in N. Let $\alpha > 0$ be such that $t \in I$ and $|t - t_0| \le \alpha$ imply that $(t, f_1(t))$ and $(t, f_2(t))$ are in N. Then the uniqueness theorem 4 yields $f_1(t) = f_2(t)$ for all $t \in I \cap [t_0 - \alpha, t_o + \alpha]$. This shows that J is open in I.

Since I is connected set and J is both open and a closed set in I, the theorem is proved.

Theorem 7 : Let I be an open interval of R, let W be an open subset of a real Banach space X, let g be a continuous map of $I \times W$ into X such that there esixts two real numbers b and c both greater than 0,

$$\sup(\|g(t,x)\| : (t,x) \in I \times W) \le b$$

and g is $c-$Lipschitz on W uniformly with respect to all compact subsets of I. Let (t_o, x_0) be any point of $I \times W$ and let r be a real number such that $0 < r < 1$ and the closed ball $\overline{B}(x_0, 2r) \subset W$. Let a be a real number such that $0 < a < r / bc$ and the interval $J = \{t_0 - a, t_0 + a\} \subset I$.

Then for each $x \in \overline{B}(x_0, r)$, there exists a unique map $h_x : J \to \overline{B}(x_0, 2r)$ such that h_x is an integral solution for g at (t_0, x). Moreover the map

$f : J \times \overline{B}(x_0, r) \to \overline{B}(x_0, 2r), f(t,x) = h_x(t)$ is a continuous local flow at (t_0, x_0) for g.

Furthermore, there is a positive real number e such that the function $x \to h_x$ is a e-Lipschitz on $\overline{B}(x_0, r)$.

Proof : Let $x \in \overline{B}(x_0, 2r)$ and H_x be the set of all continuous functions $h : J \to \overline{B}(x_0, 2r)$ such that $h(t_0) = x$. Then H_x is a closed subset of the complete metric space Z of all continuous functions of J into $\overline{B}(x_0, 2r)$ with the topology of uniform convergence on J and so H_x is itself is a complex subspace of Z.

Let e_x be the function on H_x such that for all $h \in H_x$ and all $t \in J$

$$(e_x(h))(t) = x + \int_{t_0}^{t} g(s, h(s)) ds,$$

then for all $t \in J$, $(e_x(h))(t) \in \overline{B}(x_0, 2r)$ and so $e_x(h)$ is a continuous map of J into $\overline{B}(x_0, 2r)$.

Since $(e_x(h))(t_0) = x$, e_x is a map of H_x into itself for each h_1 and h_2 in H_x, we have

$$\|e_x(h_1) - e_x(h_2)\| \le r \|h_1 - h_2\| \qquad ...(1)$$

therefore by Banach fixed point theorem, there is a unique $h_x \in H_x$ such that

$$h_x(t) - (e_x(h_x))(t) = x + \int_{t_0}^{t} g(s, h_x(s)) ds$$

for each $t \in J$ and $h_x(t_0) = x$ and so h_x is an integral solution for g at (t_0, x).

It also follows that f is a local flow for g at (t_0, x_0).

To prove the last part of the theorem, let x and y be any two points of $\overline{B}(x_0, r)$. Then

$$\|h_x - e_y(h_x)\| = \|e_x(h_x) - e_y(h_x)\|$$

$$= \sup\{\|(e_x(h_x))(t) - (e_y(h_x))(t)\| : t \in J\}$$

$$= \|x - y\| \qquad ...(2)$$

Hence

$$\|e_y(h_x) - e_y(e_y(h_x))\| \le r\|h_x - e_y(h_x)t\| \quad \{by(1)\}$$

$$= r\|x - y\|, \qquad \{by(2)\}$$

therefore for any natural number n

$$\|e_y^n(h_x) - e_y^{n+1}(h_x)\| \le r^n \|x - y\| \qquad ...(3)$$

and so

$$\|h_x - e_y^{n+1}(h_x)\| \le \|h_x - e_y(h_x)\| + \|e_y(h_x) - e_y^2(h_x)\| + \ldots\ldots$$

$$+ \|e_y^n(h_x) - e_y^{n+1}(h_x)\|$$

$$\le (1 + r + \ldots. + r^n)\|x - y\| \qquad \{by(2)\, and\, (3)\}$$

$$\le \frac{1}{1-r}\|x - y\| \qquad ...(4)$$

Again for any natural number n

$$\|e_y^n(h_x) - h_y\| = \|e_y^n(h_x) - e_y^n(h_y)\|$$

$$\le r\|e_y^{n-1}(h_x) - e_y^{n-1}(h_y)\| \qquad \{by(1)\}$$

$$\le r^n \|h_x - h_y\|$$

therefore

$$\lim_{n\to\infty}\|e_y^n(h_x) - h_y\| = 0 \qquad [\because 0 < r < 1]$$

$$\Rightarrow \lim_{n\to\infty} e_y^n(h_x) = h_y \qquad ...(5)$$

Consequently

$$\|h_x - h_y\| = \lim_{n\to\infty} \|h_x - e_y^{n+1}(h_x)\| \qquad \{by(5)\}$$

$$\le \frac{1}{1-r}\|x-y\| \qquad \{by(4)\}$$

and so the map $x \to h_x$ is e-Lipschitz on $\overline{B}(x_0, r)$, where $e = \dfrac{1}{1-r}$. Now the map $x \to h_x$ is e-Lipschitz implies that it is continuous and that f is continuous.

12.6 Maximal Intergal Solution

Let I be an open interval of R, let W be an open subset of a real Banach space X, and let g be a continuous map of $I \times W$ into X such that g is c-Lipschitz on W uniformly with respect to all compact subsets of I, where c is a positive real number. Let (t_0, x_0) be any point of $I \times W$, and F be the set of all integral solutions for g at (t_0, x_0) with open domains. If f_1 and f_2 are any two members of F then by Remark 1, f_1 and f_2 coincide on $D_m(f_1) \cap D_m(f_2)$. Let f be the union of all members of F and J be the union of domains of the members of F. Then J is an open interval of I and f is an integral solution for g at (t_0, x_0), called the maximal integral solution for g at (t_0, x_0).

Theorem 8 : Let (a,b) be an open interval of R, let W be an open subset of a real Banach space X and let g be a continuous map of $(a,b)\times W$ into X such that g is c-Lipschitz on W uniformly with respect to (a,b). Let f be the maximal integral solution for g at a point (t_0, x_0) of $(a,b)\times W$ with domain (a',b') such that there exists a positive real number r with the property that $(a', a'+r)$ and $(b'-r, b')$ are contained in (a,b), the closures of $f[(a', a'+r)]$ and $f[(b'-r, b')]$ are contained in W, and there is a positive real number m such that

$$\|g(t, f(t))\| \le m$$

for all $t \in (a', b')$. Then $a' = a$ and $b' = b$

Proof: Since f is the maximal integral solution for g at (t_0, x_0), therefore by theorem 1 for each $t \in (a', b')$, we have

$$f(t) = x_0 + \int_{t_0}^{t} g(s, f(s))ds$$

and $\|f(t_1) - f(t_2)\| \le m|(t_1 - t_2)|$ for all $t_1, t_2 \in (a', b')$

Hence $\lim_{t\to a'} f(t)$ and $\lim_{t\to b'} f(t)$ exist and belong to W.

If possible, suppose that $a' \ne a$. Then by theorem 7, there is an integral solution f' for g at $\left(a', \lim_{t\to a'} f(t)\right)$. So $Df' = Df$ on $(a', a'+t)$, where t' is a positive real number and so $f' - f$ is constant on $(a', a'+t)$.

Now $\lim_{t \to a'} f'(t) = \lim_{t \to a'} f(t)$

therefore $f' - f = 0$ on $(a', a' + t)$.

This shows that f defined on (a', b') is not a maximal integral solution for g at (t_0, x_0), which is a contradiction to the given condition.

Hence $a' = a$. Similarly, we can show that $b' = b$.

Self-Learning Exercise

1. Write whether the following statements are true or false.

   (a) Theorem under which an initial value problems of the form

   $$\frac{dy}{dx} = f(x, y),\ y(x_0) = y_0$$ has at least one solution is called uniqueness theorem.

   (b) If f_1 and f_2 be two integral solutions for a continuous map $g : I \times W \to X$ at a point $(t_0, x_0) \in I \times W$ with open domains I_1 and I_2 respectively. Then f_1 and f_2 coincide on $I_1 \cup I_2$.

   (c) A function $g : I \times W \to X$ is locally lipschitz if for each point $(t_0, x_0) \in I \times W$, there exists a nbd $J \times V$ of (t_0, x_0) and $c > 0$ such that

   $$\|g(t, x_1) - g(t, x_2)\| \le c\,\|x_1 - x_2\|$$

   for each $t \in J$ and $x_1, x_2 \in V$.

12.7 Summary

In this unit, we have seen that a fundamental lemma compares two approximate solutions of a differential equation. We also see that a local flow always exists if a continuous map $g : I \times W \to X$ satisties Lipschitz condition.

12.8 Answers to Self-Learning Exercise

1. (a) False (b) False (c) True

12.9 Exercises

1. Let I be an interval in R, W a subset of Banach space X and let $g : I \times W \to X$ be a continuous function c-Lipschitz in $x \in X$. If there are two exact solutions f_1 and $f_2 : I \to X$ of the differential equation $\frac{dx}{dt} = g(t, x)$ and if $f_1(t_0) = f_2(t_0)$, $t_0 \in I$ then the function f_1 and f_2 are identical in the interval I.

2. Let $g(t, x)$ be a real valued continuous function defined in the set $|t| \le a$, $|x| \le b$ in R^2, such that $g(t, x) < 0$ for $t, x > 0$ and $g(t, x) > 0$ for $t, x < 0$. Show that $x = 0$ is the unique solution of the differential equation $\frac{dx}{dt} = g(t, x)$ defined in a nbd of 0 and such that $x(0) = 0$.

□□□