KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

A Comparative Study of Convergence of Sequence of Functions in a Banach space

Siba Mohammed Abubakar

2013

KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

A Comparative Study of Convergence of Sequence of Functions in a Banach space

By

Siba Mohammed Abubakar

(Bsc. Mathematics)

A Thesis submitted to the Department of Mathematics, Kwame Nkrumah University of Science and Technology in partial fulfillment of the requirements for the degree of

Master Of Philosophy

College of science

APRIL 2013

Declaration

I hereby declare that this submission is my own work towards the award of the M.Phil degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

Siba Mohammed Abubakar	....	....
Student	Signature	Date

Certified by:

Dr. S. K. Opoku	....	....
Supervisor	Signature	Date

Certified by:

Mr. K. F. Darkwa	....	....
Head of Department	Signature	Date

Acknowledgements

The Prophet Muhammed (PBUH) encouraged us to seek for knowledge even if it is in China. This has always been my driving force and source of inspiration in my career in academia.

An unqualified honour and thanks to my supervisor Dr. S. A. Opoku, whose patience, endurance and suggestions helped me carry out this thesis to a successful end. Greatness shall be your portion. I pray that God Almighty continue to create more of S. A. Opoku.

Another thanks goes to the entire staff of the Mathematics Department, especially, Mr. Kwaku Darkwah for his encouragement and also allowing me do my thesis under Dr. S. A. Opoku, Dr. K. Baah Gyamfi for his crucial assistance to the work.

I reserve this heartfelt gratitude to Dr. Alexis N. Frimpong, of Agogo State College-for his moral support and encouragement.

I owe a dept of gratitude to all my class mates most especially, Gabriel Fosu Obed, Robert Acquah Kofi, Gideon Gogovi, Justice Kwame Appati, Sulemana Rahima Tiyumtaaba, Mohammed Adnan.

Lastly, I would like to thank my family for their diverse support and encouragement which had led to the success of my programme.

Dedication

I dedicate this work to The Prophet Muhammed (PBUH).

Abstract

We discuss four types of convergence of sequence of functions in a Banach space. The convergence considered include point-wise, uniform , strong and weak convergence . It is shown that uniform convergence implies the pointwise convergence and the strong convergence implied the weak convergence.

Contents

Chapter 1

Introduction

1.1 Background

Carothers(2000), modern analysis was formed during the resolution of an important controversy (or, rather, controversies) concerning the representation of 'arbitrary' functions.

Folland(1984) states that the story begins in 1746 with the famous *famous vibrating string problem.* Briefly, an elastic string of length L has each end fastened to one of the end points of the interval$[0,L]$ on the x-axis and its set into motion (as you might pluck a guitar string, for example).

Carothers(2000), the problem is to determine the position $y = f(x) = F(x,0)$ at $t = 0$ where , for simplicity, we assume that the initial velocity $F_t(x,0) = 0$. The function $F(x,t)$ is the solution to D'Alembert's wave equation : $F_{tt} = a^2 F_{xx}$ where a is a positive constant determined by certain physical properties of the string. The initial data for problem is $F(x,0) = f(x), F_t(x,0) = 0$ and $f(0) = 0 = f(L)$

The controversy, initially between D'Alembert and Euler, centre around the nature of the functions f that may be permitted as initial positions. D'Alembert argued that the initial position f must be 'continuous' (in the sense that f must be given by a single analytical expression or 'formula'), while Euler insisted that f could be 'discontinuous' (the initial position might be a series of straight line segements, as when the string is plucked in two or more 'formulas').

$F(x,t) = \sum_{k=1}^{\infty} a_k \sin(\frac{k\pi x}{L}) \cos(\frac{ak\pi t}{L}) \cdots \circledast$

is the solution of the wave equation.

Daniel Bernoulli claimed that equation $\circledast$ is the most general solution to the vibrating string problem. Euler took exception to Bernoulli's solution that accepting $\circledast$ will mean that the initial position f must satisfy $f(x) = \sum_{k=1}^{\infty} a_k \sin(\frac{k\pi x}{L}), \cdots \odot$

Euler pointed out that equation $\odot$ is odd and periodic, whereas no such assumptions can be made of (since a 'function' was understood to be a 'formula,' it was believed that the behaviour of a function on an interval completely determined its behaviour on the whose line). Bernoulli's argument was rejected by most Mathematicians of the time, including Euler and D'Alembert.

Joseph Fourier resurrected Bernoulli's assertion. Fourier presented a paper on heat transfer in which he was able to solve for the steady state temperature $T(x,y)$ of a rectangular metal plate with one edge placed on the interval $[-L,L]$ on the x-axis, and where the initial temperature along the edge $f(x) = T(x,0)$ is known but is again 'arbitrary'. Fourier's solution is based on the premise that an arbitrary function$f$ can be represented as a series of the form

$f(x) = \frac{a_o}{2} + \sum_{k=1}^{\infty}(a_n \cos(\frac{n\pi x}{L}) + b_n \sin(\frac{nk\pi x}{L}))$

If, for simplicity, we take $L = \pi$ then the Fourier series of f over the interval $[-\pi,\pi]$ is given by

$f(x) = \frac{a_o}{2} + \sum_{k=1}^{\infty}(a_n \cos(nx) + b_n \sin(nx)) \cdots \divideontimes$

Fourier argued that if the *fourier coefficients,* $a_o, a_1, \ldots, b_1, b_2, \ldots$ could be determined, that is, if $\divideontimes$ could be solved then it must be valid. Their values determined by

$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(mx) dx$ and $a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) cos(mx) dx$

At the time it was not clear how to define the integral of an 'arbitrary' function. Moreover, term-by-term integration (i.e. the exchange of limits) was not easy to justify. The question of convergence of the series enters the picture. For this reason, Fourier's work was not well received and his ideas on trignometric series went unpublished until the appearance of his classic book, *Theorie Analytique de la chaleur,*.

In particular, Fourier methods allow for a discontinuous function to be written as a sum of continuous functions, which was unthinkable consequence at the time.

Latakus(1976), it was so unthinkable that Cauchy was prompted to set the records straight in his famous *Cours d'Analyse.* Cauchy's refutation of Fourier results, often called *Cauchy's wrong theorem, which states that a convergent sum of continuous functions must again be continuous functions.*

In fact, the general consensus at the time was that both Cauchy and Fourier were right, although a few details would obviously have to be straightened out.

As early as 1826, Abel noted that there were exceptions to Cauchy's theorem and attempted to find the 'safe domain' of Cauchy's results.

But the latent contradiction in Cauchy's theorem was not fully reveallled until Sidel discovered the hidden assumption in Cauchy's proof and, in solving, introduced the concept of *Uniform Convergence.*

Although Fourier was never able to fully justify his less than rigorous arguments, the question raised by his work would inspire Mathematicians for years to come.

Gonzalez-Velasco(1992) states :

It was the success of Fourier's work in applications that made necessary redefinition of the concept of function, the introduction of convergence, a reexamination of the concept of integral, and the ideas of uniform continuity and uniform convergence. It also provided motivation for the theory of sets, was in the background of ideas leading to measure theory and contained the germs of the theory of distribution.

1.2 Statement of the Problem

This work deals with sequence of functions in a Banach space over the domain,S, for which these functions are defined. This leaves a question of whether these functions which converge in the space itself or whether when we take every point, x, in the domain there will be convergence and the relation between them.

Thus we draw the line of distinction between convergence in the space itself and when single points in the domain are taken at a time.

1.3 Objectives of the Study

This work seeks to study the relation between certain types of convergence of sequences in a Banach space. The objectives are as follows;

1. to show whether uniform convergence implies pointwise convergence

2. to investigate whether strong convergence implies weak convergence

1.4 Significance

Fourier series is written for functions which are in L_2 over interval $[0, 2\pi]$ or $[-\pi, \pi]$. For applied mathematics it is done over any closed interval $[a, b]$ except that the function should be periodic then the Fourier series converges in L_2 to the original function. When it comes to application it is used to solve *Boundary Value Problems (BVP).* The boundary conditions are defined so that the functions to be used are continuous and the boundary values are continuous or piecewise continuous. BVP require pointwise convergence. So that leaves the question *under what condition(s) will the Fourier series converge pointwise to the original function?* Then Féjer's theorem answers that question that when there is continuity. Féjer's theorem also guarantees uniform convergence. Hence as shown in this work we have pointwise convergence since uniform convergence implies pointwise convergence.

1.5 Justification

The concepts of weak, strong, uniform and pointwise convergence have been considered and the line of distinction between them drawn.

This will help applied mathematician to easily choose suitable convergence notion(s) when it comes approximation of functions. This work will also help in the study of some fundamental concepts in functional analysis as well as Banach spaces without going through any textbook.

Chapter 2

Literature Review

In this chapter, we try to review some previously done works relating to the thesis.

The combination of the structure of a vector space with the structure of a metric space naturally produces the structure of a *normed space*. If it is a complete metric space, it is called *Banach Space.*

Kreyszig(1978) states in a lemma that the metric derived by a norm satisfies the condition of *translation invariance and homogeneity*. That is $d(x+a,y+a) = d(x,y)$ and $d(\alpha x, \alpha y) = \mid \alpha \mid d(x,y)$. The theory of normed spaces, in particular Banach spaces is one of the most highly developed part of functional analysis.Inner product spaces are special normed spaces.The *Polarization identity* shows that the norm determines the inner product but not every norm is induced by an inner product.
A criterion for determining the relation between the norm and the inner product space is that :" the norm on a normed vector space is obtained by the help of the inner product if and only if the norm satisfies the Paralellogram law."

Gunawan(2002) defines a real-valued function $\langle .,. \mid .,\ldots,. \rangle$ on X^{n+1} satisfying the properties :

$\langle z_1, z_1 \mid z_2,\ldots,z_n \rangle \geq 0$; $\langle z_1, z_1 \mid z_2,\ldots,z_n \rangle = 0$ if and only if $z_1, z_2, \ldots z_n$ are linearly dependent.

$\langle z_1, z_1 \mid z_2,\ldots,z_n \rangle = \langle z_{i_1}, z_{i_1} \mid z_{i_2},\ldots,z_{i_n} \rangle$ for every permutation $(i_1,\ldots,i_n)$ of $(1,\ldots n)$

$\langle x, y \mid z_2,\ldots,z_n \rangle = \langle y, x \mid z_2,\ldots,z_n \rangle$

$\langle \alpha x, y \mid z_2,\ldots,z_n \rangle = \alpha \langle x, y \mid z_2,\ldots,z_n \rangle$,$\alpha \in \mathbb{R}$

$\langle x + x', y \mid z_2,\ldots,z_n \rangle = \langle x, y \mid z_2,\ldots,z_n \rangle + \langle x', y \mid z_2,\ldots,z_n \rangle$

as an *n-inner product on X,* and the pair $(X, \langle .,. \mid .,\ldots,. \rangle)$ an *n-inner product space.* He discusses the notions of strong and weak convergence in n-inner product spaces and studied the relation between them. In particular he showed that strong convergence implies the weak convergence and disprove the converse through a counterexample, by invoking an analogous of Parsevals identity in n-inner product space.

In modern analysis : If a sequence of real valued continuous functions , f_n , converges uniformly to a function f, then$f$ is a continuous function. In this case uniform convergence means that the maximum value of$\mid f_n(x) - f(x) \mid \to 0$ when $n \to \infty$. That is, for each n we choose the ' worst x ', which makes $\mid f_n(x) - f(x) \mid$ as large as possible. If this absolute value still tends to 0 while 'n tends to infinity' thenf_n converges uniformly to f.

Kreyszig(1978) defined a sequence x_n in a normed space $(X, \| \, . \, \|)$ to be weakly convergent if there is an $x \in X$ such that for every $f \in X'$,dual of X, $\lim\limits_{n \in \infty} f(x_n) = f(x)$ and it is denoted by $x_n \rightharpoonup x$, f is a bounded linear functional defined on X. x_n converges to x means the corresponding sequence of scalars. The scalars are obtained by taking the images of x_n under f. So $f(x_n)$ is a sequence of scalars. When such a sequence converges then we say it is weakly convergent.

Wheeden and Zygmund (1977) defined weak convergence in terms of the inner product as $\langle x_n, y\rangle \rightarrow \langle x, y\rangle$ as $n \rightarrow \infty$ for every element y in the inner product space $(X, \langle , \rangle)$. This is because in a Hilbert space H,$x_n \rightharpoonup x$ if and only if $\langle x_n, y\rangle \rightarrow \langle x, y\rangle$ for every y in the space. By Riesz representation theorem, a bounded linear functional f can be represented by $f(x) = \langle x, y\rangle$. Therefore $x_n \rightharpoonup x \Rightarrow f(x_n) \rightarrow f(x)$ for all $f \in H'$.Hence $\langle x_n, y\rangle \rightarrow \langle x, y\rangle$.

Hansen (2008) cited in Kristensen(2008) provided a set of strong results for the case where data is stationary and strong mixing.In this paper, he extends his results in two directions: First, we allow for heterogenous data where the random variables are non-identically distributed but still mixing. Second, data can potentially depend on a parameter and we show uniform convergence also over both the parameter set. The main conclusion of the paper is that as long as the mixing coefficients and suitably moments of data are uniformly bounded as functions of the sample size and the parameter, the results of Hansen (2008) still go through.

Newey(1989) provides uniform convergence results that meet two, related requirements. The first is that the results apply to objects other than sample averages. The second is that minimal pointwise convergences in probability conditions are imposed. Both of these requirements are motivated by the need for results that apply to certain nonparametric and semi parametric models. Estimators for such models often involve objects that are much more complicated than the sample averages, such as preliminary nonparametric regression estimators. To show consistency of such estimators it is useful to have uniform convergence results that apply to these objects. Furthermore, for complicated objects it is helpful to keep the convergence in probability requirement to the minimum of pointwise convergence. In addition, even for sample averages, when the data satisfies complicated dependence restrictions it may be easier to check pointwise convergence in probability, rather than the convergence of various supremums and imfimums . The focus on convergence in probability, rather than the almost sure convergence, is also in keeping with these requirements. For complicated objects it can be more difficult to show almost sure convergence. The paper presents a condition, referred to as uniform stochastic equicontinuity that together with pointwise convergence characterizes uniform convergence to equicontinuous functions on a compact set.

Lindgren (2012) in his PhD thesis developed a framework for the analysis of weak convergence and within this framework he analyzed the stochastic heat equation, the stochastic wave equation, and linearized stochastic Cahn-Hilliard, or the linearised Cahn-Hilliard-Cook equation. He was able to show that the rate of weak convergence is twice the rate of strong convergence. Numerical approximations of Stochastic Partial Differential Equations (SPDEs) because, they arise in various applications such as phenomenological studies of phase separation in alloys and modelling of thin fibres in turbulent flow . Secondly, modelling with infinite dimensional stochastic processes in the natural sciences which is underdeveloped in comparison to both deterministic models. The study of numerical SPDEs strengthens the understanding of non-smooth problems in general and the importance of concepts such as weak convergence may come with new insights and ideas to study of also deterministic numerical analysis. Fredrik stated that much less was done for weak convergence of SPDEs.

Yang (2009) stated that the notion of nearness in a number of different variations has found new applications in digital topology, image processing and pattern recognition areas, perhaps due to the fact that those structures are richer than classical topology. The main objectives are to establish a pointwise convergent nearness structure on a function space made of a family of functions from X to Y and to establish two versions of the Ascoli-Arzela theorems for nearness spaces that relate the compactness of the underlying space Y with that of the function space. It is commonly known that the issue came from the fact that a convergent sequence of continuous functions may not converge to a continuous function. So the natural question is: under what conditions the limit of a convergent sequence of continuous functions is still continuous. It turned out that the concept of equicontinuous was used to characterize the condition needed in topological spaces. There are several practical reasons to be interested in this topic. In many cases, a digital image processing algorithm is essentially the application of a sequence of deformation functions to a digital plane. For example, a deletion of a simple point (a point that does not affect the connectness of the digital picture) can be regarded as a sp-continuous function. Hence a thinning algorithm that preserves connectness can be arranged as a sequence of sp-continuous functions. It is also possible to use the tools of function spaces, and the results on convergence of function sequences to study the image processing algorithms, which opens a new set of doors. .

Desmond et al (2002) studies the numerical solution of the stochastic differential equations. The aim was to extend strong mean square convergence theory for numerical SDE simulations beyond the realm of globally Lipchitz problems. Strong convergence theorem for EulerMaruyama (EM) in the case where the vector fields are locally Lipschitz and moment bounds are available. This style of analysis is useful whenever moment bounds can be established, both for EM and for other methods that can be shown to close to EM. It was shown that the optimal rate of convergence can be recovered if the drift coefficient is also assumed to behave like a polynomial.

Takahashi and Yao (2011) prove a weak convergence theorem for Manns iteration for positively homogeneous nonexpansive mappings in a Banach space. In the theorem, the limit of weak convergence is characteraized by using a sunny generalized nonexpansive retraction in convergence theorems for projections in Banach spaces. Further, using the shrinking projection method ,they proved a strong convergence theorem for positively homogeneous nonexpansive mappings in a Banach space. From the two results, they obtained weak and strong convergence theorems for linear contractive mappings in a Banach space. These results are new even if the mappings are linear and contractive.

Hansen (2008) presented a set of rate of uniform consistency results for kernel estimators of density functions and regressions functions. It generally allows for stationary strong mixing multivariate data with infinite support, kernels with unbounded support, and general bandwidth sequences. These results are useful for semiparametric estimation based on a first-stage nonparametric estimator.The main results are the weak and strong uniform convergence of a sample average functional. The conditions imposed on the functional are general. The data are assumed to be a stationary strong mixing time series. The support for the data is allowed to be infinite, and the convergence is uniform over compact sets,expanding sets, or unrestricted euclidean space.

Chapter 3

Methodology

In this chapter we discuss concepts that will help us to compare the types of sequence of functions under consideration.

Preliminary Definitions and Basic concepts.

3.1 Vector Space

There are real vector spaces and complex spaces.The field of real numbers is denoted by $\mathbb{R}$ and the field of complex numbers by $\mathbb{C}$. Elements of $\mathbb{R}$ or $\mathbb{C}$ are called scalars. But the purpose of this work, we will sometimes use F to represent $\mathbb{R}$ or $\mathbb{C}$.

Definition 3.1.1 (Vector Space)

Let F be the field of all real numbers. A non-empty set V is called a vector space over the scalar field F, with a pair of rules for adding elements of V and multiplying elements of V by elements of F such that the following are satisfied:

1. Vector addition is commutative: for all x,y belonging to V, we have

$$x+y=y+x$$

2. Vector addition is associative: for all x,y,z of elements of V, we have

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

3. Vector addition has an identity element: There exists an element 0 the zero vector, such that

$$x + 0 = x$$

$\forall x \in V$

4. Vector addition has an inverse element: For all $x \in V$, there exists an element $-V$, called the additive inverse of x , such that

$$x + (-x) = 0$$

5. Scalar multiplication is distributive over vector addition: For every $\lambda \in F$ and $x, y \in V$

$$\lambda(x + y) = \lambda x + \lambda y$$

6. Vector addition is distributive over scalar multiplication: For all $\lambda, \mu \in F$ and for every $x \in V$

$$(\lambda + \mu)x = \lambda x + \mu x$$

7. For all $\lambda, \mu \in F$ and for every $x \in V$

$$\lambda(\mu x) = (\lambda\mu)x$$

8. Scalar multiplication has an identity element: For all $x \in V$

$$1 \bullet x = x$$

where 1 is the multiplicative identity in F.

Proposition 3.1.1

Let V be a vector space and F be the field of all real numbers. For all $x, y \in V$ and for every $\lambda \in F$, the following assertions and equalities hold:

1. The zero vector, 0 is unique.

2. For every $x \in V$, the additive inverse $-x$ is also unique.

3. If $\lambda \neq 0$ and $\lambda x = 0$ then $x = 0$

4. If $x = 0$ and $\lambda x = 0$ then $\lambda = 0$

5. $0 \bullet x = 0$

6. $(-1) \bullet x = -x$

Example 3.1.1

1. The Euclidean space $\mathbb{R}^n = \{(x_1 \ldots x_n) : x_1 \ldots x_n \in \mathbb{R}\}$. x is an element of $\mathbb{R}$ if and only if $x = (x_1 \ldots x_n)$.
   If $x = (x_1 \ldots x_n)$ and $w = (w_1 \ldots w_n)$ then :

   $$x + w = (x_1 + w_1, \ldots, x_n + w_n) \in \mathbb{R}^n$$

   If $\lambda \in F$ and $x = (x_1 \ldots x_n) \in \mathbb{R}^n$ then:

   $$\lambda x = (\lambda x_1, \ldots, \lambda x_n) \in \mathbb{R}^n.$$

   Thus $\mathbb{R}^n$ is a vector space.

2. Let a, b be real numbers such that $a < b$. Denote by $C[a, b]$ the set of all continuous real-valued functions on the closed interval $[a, b]$.
   For every pair $f, g \in C[a, b]$ the element $f + g \in C[a, b]$ is defined by:

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

$\forall x \in [a,b]$.

For every $\lambda \in \mathbb{R}$ and for every $f \in C[a,b]$ λf is defined by :

$$(\lambda f)(x) = \lambda f(x).$$

Thus $C[a,b]$ is a vector space.

3.2 Linear Independence and Basis

Let V be a vector space over a field F.

1. A non empty subset A of V is said to be *linearly independent* over F if for every finitely many distinct elements $\{a_1 \dots a_n\}$ of A and scalars $\lambda_1 \dots \lambda_n \in F$ the condition

$$\sum_{j=1}^{n} \lambda_j a_j = 0$$

implies that $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$

2. A non empty set D is said to be *linearly dependent* if it is not linearly independent. Thus D is linearly dependent over F if and only if there exist finitely many distinct elements $d_1 \dots d_q \in D$ and scalars $\gamma_1, \dots, \gamma_q \in F$ such that :

$$\sum_{j=1}^{q} \lambda_j d_j = 0$$

and at least one of $\{\lambda_1, \dots, \lambda_q\}$ is not 0.

3. A non empty subset S of V is said to *span* V if for every $x \in V$ there exist finitely many elements $u_1, \dots, u_m \in S$ and scalars $\gamma_1, \dots, \gamma_m \in F$ such that

$$x = \sum_{j=1}^{m} \gamma_j u_j = 0.$$

Thus a basis B of a vector space V is linearly independent subset of V that spans V.

3.3 Normed Vector Space

The concept of a norm is an abstract generalization of the length of a vector.

Definition 3.3.1 (Norm)

Let V be a vector space over F a mapping $\|,\|: V \to \mathbb{R}$ is called a norm on V if these four conditions are satisfied:

1. $\| x \| \geq 0$ for every $x \in V$.

2. For $w \in V$, $\| w \| = 0$ if and only if $w = 0$, the zero in V.

3. $\| \lambda x \| = | \lambda | \| x \|$ for all $\lambda \in F$ and for all $x \in V$.

4. $\| x + y \| \leq \| x \| + \| y \|$ (triangle inequality) for all $x, y \in V$.

Definition 3.3.2 (Normed Vector Space)

A normed vector space is a vector space with a norm. A normed vector space is an ordered pair $(V, \|,\|)$ where V is a vector space and $\|,\|$ is a norm defined on V.

Example 3.3.1

1. A norm $\| \bullet \|$ is defined on $\mathbb{R}^n$ by the formula $\| x \| = max\{| x_1 |, \ldots, | x_n |\}$ if $x = (x_1, \ldots, x_n)$.

   Proof

   i. $\| x \| \geq | x_1 | \geq 0$ for every $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

   ii. When $w = (w_1, \ldots, w_n) \in \mathbb{R}^n$ then $\| w \| = 0$ if and only if

   $$0 \leq | w_j | \leq \| w \| \leq 0$$

   for every $j \in 1, \ldots, n$. It follows that $\| w \| = 0$ if and only if $w = (0, \ldots, 0)$.

iii. If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. and $\lambda \in \mathbb{R}$ choose $k \in \{1, \ldots, n\}$ such that $| x_k | = \| x \|$ then:

$$\| \lambda x \| = max\{| \lambda x_1 |, \ldots, | \lambda x_n |\} = | \lambda | \| x \|$$

iv. If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $x = (y_1, \ldots, y_n) \in \mathbb{R}^n$ choose $k \in \{1, \ldots, n\}$ such that $\| x + y \| = | x_k + y_k |$. Then:

$$\| x + y \| = | x_k + y_k | \leq | x_k | + | y_k | \leq \| x \| + \| y \|$$

Thus all four axioms which define a norm are satisfied by $\|, \|$

$\mathbb{R}^n$ together with this norm is called *Euclidean n-space.*

Another norm $\| x \| = \sqrt{\sum_{j=1}^{n} | x_j |^2}$ can also be defined on $\mathbb{R}^n$. This is often called the *Euclidean norm.*

2. Define a mapping $\|, \| : \mathbb{R}^n \to \mathbb{R}$ such that for arbitrary $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ the norm $\| x \| = \sqrt[p]{\sum_{j=1}^{n} | x_j |^p}$ where $1 \leq p < \infty$

Proof

i. $\| x \| = \sqrt[p]{\sum_{j=1}^{n} | x_j |^p} \geq 0$ by definition.

ii. If $\| x \| = 0$ then $\sqrt[p]{\sum_{j=1}^{n} | x_j |^p} = 0$

$\Rightarrow x_j = 0, j = 1, \ldots, n$

If $x = 0$ then $\| x \| = \sqrt[p]{\sum_{j=1}^{n} | x_j |^p} = 0$

iii. If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

$$\begin{aligned}
\| \lambda x \| &= \sqrt[p]{\sum_{j=1}^{n} | \lambda x_j |^p} \\
&= \sqrt[p]{| \lambda |^p | x_j |^p} \\
&= \sqrt[p]{| \lambda |^p} \sqrt[p]{\sum_{j=1}^{n} | x_j |^p} \\
&= | \lambda | \| x \|
\end{aligned}$$

iv. If $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ then

$$\begin{aligned}
\| x+y \| &= \sqrt[p]{\sum_{j=1}^{n} | x_j + y_j |^p} \\
\| x+y \|^p &= \sum_{j=1}^{n} | x_j + y_j |^p \\
&= \sum_{j=1}^{n} | x_j + y_j || x_j + y_j |^{p-1} \\
&= \sum_{j=1}^{n} [| x_j || x_j + y_j |^{p-1} + | y_j || x_j + y_j |^{p-1}] \\
&= \sum_{j=1}^{n} | x_j || x_j + y_j |^{p-1} + \sum_{j=1}^{n} | y_j || x_j + y_j |^{p-1}
\end{aligned}$$

Applying the Holder's inequality we will get

$$\begin{aligned}
&\leq \left(\sum_{j=1}^{n} | x_j |^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^{n} | x_j + y_j |^{p(p-1)} \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} | y_j |^{p(p-1)} \right)^{\frac{1}{p}} \left(\sum_{j=1}^{n} | x_j + y_j |^p \right)^{\frac{1}{p}} \\
&\leq \left(\sum_{j=1}^{n} | x_j + y_j |^{p(p-1)} \right)^{\frac{1}{p}} \left[\left(\sum_{j=1}^{n} | x_j |^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} | y_j |^p \right)^{\frac{1}{p}} \right] \\
&\leq \| x+y \|^{\frac{p}{q}} \left[\left(\sum_{j=1}^{n} | x_j |^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} | y_j |^p \right)^{\frac{1}{p}} \right]
\end{aligned}$$

Dividing through by

$$\| x+y \|^{\frac{p}{q}}$$

we obtain

$$\| x+y \| \leq \left(\sum_{j=1}^{n} | x_j |^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{n} | y_j |^p \right)^{\frac{1}{p}}$$

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

Thus the four axioms about norm satisfied.

Definition 3.3.3

Let $(V, \| \, . \, \|)$ be a normed vector space and x_n a sequence in V. We say that a sequence x_n is convergent in V if there exists $x \in V$ satisfying the condition:

To every positive real number ε, there corresponds a positive integer p such that

$$\| x_n - x \| \leq \varepsilon \quad \forall x \geq p.$$

Under such circumstances we write

$$x_n \to x \qquad as \qquad n \to \infty$$

or $\lim_{n\to\infty} x_n = x.$

Proposition 3.3.1

For all sequences defined on the normed vector space $(V, \| \, . \, \|)$ the following assertions are valid:

1. Uniqueness of a limit Let x_n be a sequence in V.

   If $x, y \in V$, $x_n \to x$ and $y_n \to y$ as $n \to \infty$ then $x = y$

   Proof

   Assume that $x \neq y$

   Then $x - y \neq 0$ and so $\| x - y \|$ is a positive number.

   Choose positive integers q_1, q_2 such that

$$\begin{aligned} \| x_n - x \| &< \frac{1}{2} \| x - y \| \quad \forall n \geq q_1 \\ \| x_n - x \| &< \frac{1}{2} \| x - y \| \quad \forall n \geq q_2 \end{aligned}$$

Let $q = q_1 + q_2$

$$\begin{aligned} \| x - y \| &= \| x - x_q + x_q - y \| \\ &\leq \| x - x_q \| + \| x_q - y \| \\ &< \| x - y \| \end{aligned}$$

This is a contradiction. The assumption is false. Hence $x = y$

2. If $x_n \to x$ and $y_n \to y$ then $x_n + y_n \to x + y$

   Proof

   Given $\varepsilon > 0$ choose positive integers p_1 and p_2 such that

$$\begin{aligned} \| x_n - x \| &< \frac{\varepsilon}{2} \quad \forall n \geq p_1 \\ \| y_n - y \| &< \frac{\varepsilon}{2} \quad \forall n \geq p_1 \end{aligned}$$

   Let $p = p_1 + p_2$

$$\begin{aligned} \| (x_n + y_n) - (x + y) \| &= \| x_n - x + y_n - y \| \\ &\leq \| x_n - x \| + \| y_n - y \| \\ &< \varepsilon \end{aligned}$$

   Thus $x_n + y_n \to x + y$ as $n \to \infty$

Definition 3.3.4 (Cauchy sequence)

Let $(V, \| \, . \, \|)$ be a normed vector space. A sequence x_n in V is called a Cauchy sequence in $(V, \| \, . \, \|)$ if to every positive real number ε, there corresponds a positive integer p such that for all $m \geq p$ and $n \geq p$ implies

$$\| x_m - x_n \| < \varepsilon$$

More briefly, x_n is a Cauchy sequence if $\lim_{m,n \to \infty} \| x_m - x_n \| = 0$

Definition 3.3.5

If a_n is a convergent sequence in $(V, \parallel . \parallel)$ then a_n is a Cauchy sequence in $(V, \parallel . \parallel)$

Proof

Let $\lim_{n\to\infty} a_n = a$ Given a positive real number ε, choose positive integer p such that

$$\parallel a_n - a \parallel < \frac{\varepsilon}{2} \quad \forall n \geq p$$

Then

$$\begin{aligned} \parallel a_m - a_n \parallel &= \parallel a_m - a + a - a_n \parallel \\ &\leq \parallel a_m - a \parallel + \parallel a - a_n \parallel \\ &< \varepsilon \quad \forall m, n \geq p \end{aligned}$$

Theorem 3.3.2

If x_n is a Cauchy sequence in a normed vector space, then the sequence of norms $\parallel x_n \parallel$ converges.

Proof

Since

$$\parallel x \parallel - \parallel y \parallel \quad \leq \quad \parallel x - y \parallel$$

we have

$\parallel x_m \parallel - \parallel x_n \parallel \leq \parallel x_m - x_n \parallel \to 0$ as $m, n \to \infty$

This shows that the sequence of norms is a Cauchy sequence of real numbers, hence convergent.

Definition 3.3.6

Let a_n be a sequence in a set S. If n_j is a non decreasing sequence of positive integers then a_{n_j} is called a subsequence of a_n.

$a_{n_j} \to \infty$ *as* $j \to \infty$

Theorem 3.3.3

Suppose a_n is a convergent sequence in a normed vector space $(V, \| \,.\, \|)$ and $a = \lim\limits_{n\to\infty} a_n$. If a_{n_j} is a subsequence of a_n then a_{n_j} is also convergent and $a = \lim\limits_{j\to\infty} a_{n_j}$

Proof

Given $\varepsilon > 0$, choose a positive integer p such that $\| a_n - a \| < \varepsilon \quad \forall n \geq p$

Then $\| a_{n_j} - a \| < \varepsilon \quad \forall j \geq p \quad$ since $n_j \geq j \quad \forall j \geq 1$

Thus $a_{n_j} \to a$ as $j \to \infty$

Theorem 3.3.4

Let $(V, \| \,.\, \|)$ be a normed vector space over F. Suppose a_n is a Cauchy sequence in V. If a_{n_j} is a convergent subsequence of a_n and $a = \lim\limits_{j\to\infty} a_{n_j}$ then a_n is convergent and $a = \lim\limits_{n\to\infty} a_n$.

Proof

Given a positive real number ε, choose positive integers t_1, t_2 such that

$\| a_m - a_n \| < \frac{\varepsilon}{2}$ for all $m \geq t_1$ $n \geq t_2$

while $\| a_{n_j} - a \| < \frac{\varepsilon}{2}$ for all $j \geq t_2$

Let $t = t_1 + t_2$

Then $\| a_n - a \| \leq \| a_n - a_{n_t} \| + \| a_{n_t} - a \| < \varepsilon$ for all $n \geq t$

Thus $a_n \to a$ as $n \to \infty$

Definition 3.3.7 (Inner Product) *Let V be a vector space over a field F. A mapping $\langle , \rangle : V \times V \to F$ is called an inner product space if these condition are satisfied:*

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in V$.

2. $\langle x,y\rangle \geq 0$ for $x,y \in V$.

3. For $x \in V$, $\langle x,x\rangle = 0$ if and only if $x = 0$ the zero in V.

4. $\langle \lambda x,y\rangle = \lambda\langle x,y\rangle$ whenever $\lambda \in F$ and $x,y \in V$.

5. $\langle x+z,y\rangle = \langle x,y\rangle + \langle z,y\rangle$ for all $x,y,z \in V$.

   Under such circumstances we call the ordered pair $(V,\langle . \rangle)$ is called an inner product space.

Theorem 3.3.5

Let $(V,\langle . \rangle)$ be an inner product space over F. If $x,y \in V$ and $\lambda \in F$ then $\langle x,\lambda y\rangle = \bar{\lambda}\langle x,y\rangle$

Proof

$$\begin{aligned} \langle x,\lambda y\rangle &= \overline{\lambda\langle y,x\rangle} \\ &= \bar{\lambda}\overline{\langle y,x\rangle} \\ &= \bar{\lambda}\langle x,y\rangle \end{aligned}$$

Hence the proof.

Theorem 3.3.6 (Continuity of inner product space) *If $x_n \to x$ and $y_n \to y$ then $\langle x_n,y_n\rangle \to \langle x,y\rangle$*

Proof

$$\begin{aligned} |\langle x_n,y_n\rangle - \langle x,y\rangle| &= |\langle x_n,y_n\rangle - \langle x_n,y\rangle + \langle x_n,y\rangle - \langle x,y\rangle| \\ &\leq |\langle x_n,y_n-y\rangle| + |\langle x_n-x,y\rangle| \\ &\leq \| x_n \|\| y_n-y \| + \| x_n-x \|\| y \| \to 0 \end{aligned}$$

Since $y_n - y \to 0$ and $x_n - x \to 0$ as $n \to \infty$

Theorem 3.3.7 *Given an inner product space $(V, \langle . \rangle)$ define a norm $\| \, . \, \|$ on V by $\| x \| = \sqrt{\langle x, x \rangle}$ for all $x \in V$. Then $\| \, . \, \|$ is a norm on V.*

Proof

1. By definition $\| x \| \geq 0$
2. For $x \in V$ $\| x \| = 0$ if and only if $\sqrt{\langle x, x \rangle} = 0$
   if and only if $\langle x, x \rangle = 0$ if and only if $x = 0$
3. Let $\lambda \in F$ and $x \in V$ then:
   $\| \lambda x \|^2 = \langle \lambda x, \lambda x \rangle = | \lambda |^2 \langle x, x \rangle$
   Therefore $\| \lambda x \| = | \lambda | \| x \|$
4. Let $x, y \in V$ then

$$\begin{aligned}
\| x + y \|^2 &= \langle x + y, x + y \rangle \\
&= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&= \| x \|^2 + 2\Re\langle x, y \rangle + \| y \|^2 \\
&\leq \| x \|^2 + 2 \, | \langle x, y \rangle | + \| y \|^2 \\
&\leq \| x \|^2 + 2 \| x \| \| y \| + \| y \|^2 \\
\| x + y \|^2 &\leq (\| x + y \|)^2 \\
\Rightarrow \| x + y \| &\leq \| x \| + \| y \|
\end{aligned}$$

Proposition 3.3.8

If a_n is a Cauchy sequence in a normed vector space V, a_n need not to be a convergent sequence in V.

The most commonly used are sequences in the Euclidean line $\mathbb{R}$ and Euclidean spaces $\mathbb{R}^n$.

Definition 3.3.8

A sequence a_n of real numbers is said to be convergent if there exists a real number a satisfying the condition:

to every positive real number ε there corresponds a positive integer p such that

$\mid a_n - a \mid < \varepsilon$ for all $n \geq p$.

Under such circumstances we write $a_n \to a$ as $n \to \infty$.

Definition 3.3.9

Let u_n be a sequence of real numbers then

i. u_n is said to be strictly increasing if $u_1 < u_2 < u_3 < \ldots$

ii. u_n is said to be monotonic non-decreasing if $u_1 \leq u_2 \leq u_3 \leq \ldots$

iii. u_n is said to be strictly decreasing if $u_1 > u_2 > u_3 > \ldots$

iv. u_n is said to be monotonic non-decreasing if $u_1 \geq u_2 \geq u_3 \geq \ldots$

Definition 3.3.10

If a_n is a sequence and n_k is a strictly increasing sequence of positive integers then a_{n_k} is called a subsequence of a_n.

Definition 3.3.11 *If a_n is a convergent sequence of real numbers and $\lim_{n\to\infty} a_n = a$ then $a_{n_{j\to a}}$ as $j \to \infty$ for every subsequence a_{n_j} of a_n.*

Proof

Given a positive real number ε, choose a positive integer p such that $\mid a_n - a \mid < \varepsilon$ for all $n \geq p$.

Then $\mid a_{n_j} - a \mid < \varepsilon$ for all $j \geq p$.

Example 3.3.2

1. Given a real number a, let $a_n = a$ for all $n \geq 1$. Then a_n is a convergent sequence and $\lim_{n\to\infty} a_n = a$

Proof

If ε is a positive real number then $\mid a_n - a \mid = 0 < \varepsilon$ for all $n \geq 1$

2. If $b_n = (-1)^n$ for every positive integer n, then b_n is not convergent.

   Proof

   Assume that b_n is convergent

   Let $b = \lim\limits_{n \to \infty} b_n$

   Then by the uniqueness, $1 = \lim\limits_{k \to \infty} b_{2k} = b = \lim\limits_{k \to \infty} b_{2k+1} = -1$.

   This is a contradiction. The assumption is false. Hence b_n is not convergent.

Theorem 3.3.9 *If a_n is a Cauchy sequence of real numbers, then a_n is a bounded sequence.*

Proof

Choose a positive integer t such that $\mid a_m - a_n \mid < 1$ for all $m \geq q$ and for all $n \geq q$

Then $\mid a_m \mid = \mid a_m - a_t + a_t \mid \leq \mid a_m - a_t \mid + \mid a_t \mid < 1 + \mid a_t \mid$ for all $m \geq q$

Let $\delta = 1 + \sum_{j=1}^{t} \mid a_j \mid$ then $\mid a_n \mid < \delta$ for every positive integer n.

3.4 Pointwise Convergence

Let S be a non-empty set. Suppose f_n is a sequence of real-valued functions on X. Given $a \in S$ we say that $f_n(a)$ is convergent if there exists a real-valued function f such that these conditions are satisfied:

To every positive real number ε there corresponds a positive integer t such that

$\mid f_n(a) - f(a) \mid < \varepsilon$ for all $n \geq t$

This is equivalent to assertion that $f_n(a) \to f(a)$ as $n \to \infty$.

3.5 Uniform Convergence

Let X be a non-empty set and f_n a sequence of functions on X. We say that f_n is uniformly convergent on X if there exists a function f on X satisfying the condition:
To every positive real number ε there corresponds a positive integer p such that
$| f_n(x) - f(x) | < \varepsilon$ for all $x \in X$ and for all $n \geq q$

3.5.1 Test For Uniform Convergence of Sequence of Functions

In order to test whether a given sequence f_n is uniformly convergent or not in a given interval, so far we have the definition of uniform convergence. Accordingly, we have to try to get $m \in N$, independent of x, which is not easy in practice. This method can be replaced by an easy method given in the following theorem:

Theorem 3.5.1 $[M_n - Test]$

Let f_n be a sequence of functions defined on an interval I such that
$\lim_{n\to\infty} f_n = f(x) \quad \forall x \in [a,b]$ and let $M_n = sup\{| f_n(x) - f(x) | : x \in [a,b]\}$.
Then f_n converges uniformly on $[a,b]$ if and only if $M_n \to 0 \quad as \quad n \to \infty$.

Example 3.5.1

Prove that the sequence f_n , where $f_n(x) = \frac{x}{(1+nx^2)}$ converges uniformly on any closed intervals *I*.

SOLUTION

Here pointwise limit is given by $\lim_{n\to\infty} f_n(x) = 0, \quad \forall x \in I$

$| f_n(x) - f(x) | = | \frac{x}{1+nx^2} - 0 | = | \frac{x}{1+nx^2} | = | y |$,

where $y = \frac{x}{1+nx^2}$

$\frac{dy}{dx} = \frac{(1+nx^2)\bullet 1 - x \bullet 2nx}{(1+nx^2)^2} = \frac{1-nx^2}{(1+nx^2)^2}$

For maximum and minimum value of y, we have $\frac{dy}{dx} = 0$

Which implies $1 - nx^2 = 0$, and hence we get $x = \frac{1}{\sqrt{n}} \in I$.
$\frac{d^2y}{dx^2} = \frac{(1+nx^2)^2 \bullet (-2nx) - (1-nx^2) \bullet 2(1+nx^2) \bullet 2nx}{(1+nx^2)^4}$
$\frac{d^2y}{dx^2} = \frac{-2nx(1+nx^2) - 4nx(1-nx^2)}{(1+nx^2)^3}$
Substituting $x = \frac{1}{\sqrt{n}}$,
$\frac{d^2y}{dx^2} = \frac{-\sqrt{n}}{2} < 0$ showing that y is maximum when $x = \frac{1}{\sqrt{n}}$, and
$y = \frac{1}{2\sqrt{n}}$ after substituting $x = \frac{1}{\sqrt{n}}$.
$\therefore M_n = sup\{| f_n(x) - f(x) |: x \in I\} = sup\{| y |: x \in I\} = \frac{1}{2\sqrt{n}}$
Since $M_n \to 0 \quad as \quad n \to \infty$, f_n is uniformly convergent on any closed interval I.

Theorem 3.5.2

Let X be a non-empty set and a_n a sequence of real-valued functions on X. Then these two statements are equivalent:

1. a_n is uniformly convergent

2. (Cauchy Criterion) to every positive real number ε there corresponds a positive integer p such that $| a_m(x) - a_n(x) | < \varepsilon$ for all $x \in X$ for all $m, n \geq p$.

Proof

If (1) is true we let $a_n \to a$ uniformly on X.

Given a positive real number ε choose positive integer p such that

$| a_m(x) - a(x) | < \frac{\varepsilon}{2}$ for all $x \in X$ for all $m \geq p$. Then

$$\begin{aligned} | a_m(x) - a_n(x) | &= | a_m(x) - a(x) + a(x) - a_n(x) | \\ &\leq | a_n(x) - a(x) | + | a(x) - a_m(x) | \\ &< \varepsilon \end{aligned}$$

for all $x \in X$ and for all $m, n \geq p$

Hence $1 \Rightarrow 2$

If (2) is true then $a_n(x)$ is a Cauchy sequence of real numbers at each point $x \in X$.

Then $a_n(x)$ is a convergent sequence of real numbers at each point $x \in X$.

Let $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x \in X$.

Given a positive real number ε choose a positive integer p such that

$\mid a_m(x) - a_n(x) \mid < \frac{\varepsilon}{2}$ for all $x \in X$ and for all $m, n \geq p$ let $m \to \infty$.

Then $\mid a_m(x) - a_n(x) \mid \leq \frac{\varepsilon}{2} < \varepsilon$ for all $n \geq p$ and for all $x \in X$.

$2 \Rightarrow 1$

Corollary 3.5.3

Suppose that f_n is a sequence of continuous function on a set S; if the sequence f_n converges uniformly on S to a function f, then f is continuous on S.

Notation 3.5.1

If a, b are real numbers such that $a < b$ then $]a, b[= \{x \in \mathbb{R} | a < x < b\}$ is called an open interval in $\mathbb{R}$ and $[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$ is called a closed interval in $\mathbb{R}$.

If $I = [a, b]$ or $I =]a, b[$ then $b - a$ is called the *length of the interval I.*

Definition 3.5.1

If $[a_n, b_n]$ is a sequence of closed intervals in $\mathbb{R}$ such that $[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset \ldots$ and $b_n - a_n \to 0$ as $n \to \infty$. Then the intervals $[a_n, b_n]$ are called nested intervals.

Lemma 3.5.4 *If $[a_n, b_n]$ is a nested interval then $\cap_{n=1}^{\infty} [a_n, b_n] \neq \varnothing$*

Theorem 3.5.5

If I_n are nested intervals then $\cap_{n=1}^{\infty} I_n$ contains exactly one element.

Proof

Choose $v \in \cap_{n=1}^{\infty} [a_n, b_n]$ using the lemma above.

Assume that $\cap_{n=1}^{\infty} [a_n, b_n] \neq \{v\}$.

Choose $w \in \cap_{n=1}^{\infty} [a_n, b_n]$ such that $v \neq w$.

Then $v - w \neq 0$ and so $\mid v - w \mid$ is a positive real number.

Choose a positive integer p such that $b_n - a_n <| v - w |$ for all $n \geq p$.

Then both v and w are elements of $[a_p, b_p]$

There is a contradiction $| v - w | \leq b_p - a_p <| v - w |$

The assumption is false. Hence $\cap_{n=1}^{\infty}[a_n, b_n] = \{v\}$.

3.5.2 Bolzano-Weierstrass Theorem

Every bounded sequence of real numbers has a convergent subsequence.

Proof

Given a bounded sequence of real numbers a_n we choose a strictly increasing sequence $x_n \to x$ as $n \to \infty$ and $\lim_{n\to\infty} x_n = x$ of positive integers and a real number v such that $a_{n_k} \to v$ as $k \to \infty$.

Choose real numbers a, b such that $a < b$, a is a lower bound of $\{a_n \mid n = 1, 2, \dots\}$ and b is an upper bound of $\{a_n \mid n = 1, 2, \dots\}$

If $a_n \in [a, \frac{a+b}{2}]$ for infinitely many n,

Let $I_1 = [a, \frac{a+b}{2}]$ or $I_1 = [\frac{a+b}{2}, b]$, then

the length of $I_1 = \frac{b-a}{2}$ and $a_n \in I_1$ for infinitely many n.

Next choose I_2 such that $I_1 \supset I_2$, then the length of $I_2 = \frac{b-a}{2^2}$ and $a_n \in I_2$ for infinitely many n.

Continuing the process we obtain by induction a sequence I_j of closed intervals such that $I_1 \supset I_2 \supset I_3 \supset \dots$ then the length of $I_j = \frac{b-a}{2^j}$ for all j and $a_n \in I_j$ for infinitely many n.

$\frac{b-a}{2^j} \to 0$ as $j \to \infty$ and so let v be the unique element of $\cap_{j=1}^{\infty} I_j$

Finally choose a positive integer n_1 such that $a_{n_1} \in I_1$. Choose $n_2 > n_1$ such that $a_{n_2} \in I_2$.

For a positive integer k choose n_k such that $a_{n_k} \in I_k$ then $a_n \in I_{k+1}$ for infinitely many n.

Choose $n_{k+1} > n_k$ such that $a_{n_k+1} \in I_{k+1}$.

Then by induction a subsequence a_{n_k} of a_n has been chosen and $a_{n_k} \to v$ as $k \to \infty$

Theorem 3.5.6

These two statements about a sequence a_n of real numbers are equivalent:

(a) a_n is convergent

(b) a_n is a Cauchy sequence.

Proof

Suppose (a) is true and let $a = \lim_{n\to\infty} a_n$.

Given a positive real number ε choose a positive integer q such that

$\mid a_n - a \mid < \frac{\varepsilon}{2}$ for all $n \geq q$. Then

$$\begin{aligned} \mid a_m - a_n \mid &= \mid a_m - a + a - a_n \mid \\ &\leq \mid a_m - a \mid + \mid a - a_n \mid \\ &< \varepsilon \end{aligned}$$

for all $m \geq q$ ans for all $n \geq q$.Thus $a \Rightarrow b$

Suppose (b) is true

Since a_n is a Cauchy sequence then its bounded.

Using Bolzano-Weierstrass Theorem choose a subsequence a_{n_j} and let $a = \lim_{j\to\infty} a_{n_j}$

Given a positive real number ε choose a positive integers t_1, t_2 such that $\mid a_m - a_n \mid < \frac{\varepsilon}{2}$

for all $m \geq t_1$ and for all $n \geq t_1$

while $\mid a_{n_j} - a \mid < \frac{\varepsilon}{2}$ for all $j \geq t_2$

Let $q = t_1 + t_2$ Then

$$\begin{aligned} \mid a_n - a \mid &= \mid a_n - a_{n_q} + a_{n_q} - a \mid \\ &\leq \mid a_n - a_{n_q} \mid + \mid a_{n_q} - a \mid \\ &< \varepsilon \end{aligned}$$

for all $n \geq q$.Thus $b \Rightarrow a$

Definition 3.5.2 (Metric)

Let X be a non-empty set. A mapping $d : X \times X \to \mathbb{R}$ *is called* a metric on X *if these four conditions are satisfied:*

1. $d(x,y) = d(y,x)$ for every pair $x,y \in X$.

2. $d(x,y) \geq 0$ for every pair $x,y \in X$.

3. $d(a,b) = 0$ for $a,b \in X$ if and only if, $a = b$

4. Triangle inequality: $d(x,y) \leq d(x,z) + d(z,y)$ for all $x,y,z \in X$.

Note that $d(x_1,x_n) \leq d(x_1,x_2) + d(x_2,x_3) + \cdots + d(x_{n-1},x_n)$ is called the *Generalized triangle inequality*

Under such circumstances the ordered pair (X,d) is called a *metric space.* $d(x,y)$ is regarded as the *distance* between *x* and *y*.

A subspace $(Y,\bar{d})$ of (X,d) is obtained if we take a subset $Y \subset X$ and restrict d to $Y \times Y$: thus the metric on Y is the restriction $\bar{d} = d\mid_{Y \times Y}$. Thus $\bar{d}$ is called *the metric induced on Y by d.*

Next we look at the metric defined on various spaces:

1. The Euclidean line $\mathbb{R}$ is the set of all real numbers with metric by $d(x,y) =\mid x-y \mid$

2. The metric space $\mathbb{R}^2$, called the Euclidean plane, is obtained if we take the set of ordered pair of real numbers, written $x = (x_1,x_2)$, and $y = (y_1,y_2)$ for all $x,y \in \mathbb{R}^2$. The Euclidean metric is defined by $d(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2}$ OR $d_1(x,y) =\mid x_1-y_1 \mid + \mid x_2-y_2 \mid$. The metric space d_1 does not have a standard name but sometimes called *taxicab metric* because $\mathbb{R}^2$ is sometimes denoted by E^2.

3. The Euclidean space $\mathbb{R}^n$ has a metric defined by $d(x,y) = \sqrt{(x_1-y_1)^2 + \cdots + (x_n-y_n)^2}$ where $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$. for all $x,y \in \mathbb{R}^n$.

4. The space ℓ_p. Let $p \geq 1$ be a fixed real number. By definition each element in the space ℓ_p is a sequence $x = (x_1, x_2, \dots)$ of numbers such that $\mid x_1 \mid^p + \mid x_2 \mid^p + \dots$ converges. That is $\sum_{j=1}^{\infty} \mid x_j \mid^p < \infty$ and the metric is defined by
$d(x,y) = (\sum_{j=1}^{\infty} \mid x_j - y_j \mid^p)^{\frac{1}{p}}$ where $y = (y_1, y_2, \dots)$ and $\sum_{j=1}^{\infty} \mid y_j \mid^p < \infty$
If $p = 2$ we have the famous Hilbert space ℓ_2 with metric $d(x,y) = \sqrt{\sum_{j=1}^{\infty} \mid x_j - y_j \mid^2}$.

Definition 3.5.3

Given a normed vector space $(V, \| \, . \, \|)$ define $d : V \times V \to \mathbb{R}$ by $d(w,z) = \| w - z \|$. Then d is a metric on V.

Proof

i. $d(w,z) = \| w - z \| = \| z - w \| = d(z,w)$ for every pair $w, z \in V$.

ii. $d(w,z) = \| w - z \| \geq 0$ for every pair $w, z \in V$.

iii. For $a, b \in V$ the condition $d(a,b) = 0$ is equivalent to $\| a - b \| = 0$ if and only if $a - b = 0$ if and only if $a = b$

iv. If $u, w, z \in V$ then

$$
\begin{aligned}
d(w,z) &= \| w - z \| \\
&= \| w - u + u - z \| \\
&\leq \| w - u \| + \| u - z \| \\
&\leq d(w,u) + d(u,z)
\end{aligned}
$$

Thus d is a metric on V.

d is called *the metric induced on V by the norm* $\| \, . \, \|$

Definition 3.5.4

Given a point $x_o \in X$ and a real number $r > 0$ we define three sets as follows:

1. $B(x_o, r) = \{x \in X \mid d(x, x_o) < r\}$ is called an open ball.

2. $\tilde{B}(x_o, r) = \{x \in X \mid d(x, x_o) \leq r\}$ is called a closed ball.

3. $S(x_o, r) = \{x \in X \mid d(x, x_o) = r\}$ is called a sphere.

We see that an open ball of radius r is the set of all points in X whose distance from the centre of the ball is less than r.

An open ball $B(x_o, \varepsilon)$ is often called an $\varepsilon-$neighbourhood of x_0 $(\varepsilon > 0)$. By neighbourhood of x_o we mean any subset of X which contains an $\varepsilon-$neighbourhood of x_o. We see directly from that definition that every neighbourhood of x_o contains x_o: in other words x_o is a point of each of its neighbourhoods. If N is a neighbourhood of x_o and $N \subset M$ then M is also a neighbourhood of x_o.

3.6 Uniform Continuity

Let (S, d) and (Y, ρ) be metric spaces and $f : S \to Y$ a mapping. We say that f is uniformly continuous if to every positive real number ε there corresponds a positive real number δ such that $\rho(f(x), f(v)) < \varepsilon$ for every pair $x, v \in \mathbb{Z}$ such that $d(x, v) < \delta$.

Example 3.6.1

Let (S, d) be a metric space. Given $a \in S$ define $f : S \to \mathbb{R}$ by $f(x) = d(a, x)$. Then f is uniformly continuous.

Proof

For every pair $x, y \in S$

$f(x) = d(a,x) \leq d(a,y) + d(x,y) = f(y) + d(x,y)$

$f(y) = d(a,y) \leq d(a,x) + d(x,y) = f(x) + d(x,y)$

Hence $\mid f(x) - f(y) \mid \leq d(x,y)$

Given a positive real number ε let $\delta = \varepsilon$ then

$\mid f(x) - f(y) \mid < \varepsilon$ for every pair $x, y \in S$ such that $d(x,y) < \delta$.

Theorem 3.6.1

A mapping T of a metric space X into a metric space Y is continuous if and only if the inverse image of any open subset of Y is an open subset of X.

Proof

Suppose T is continuous.

Let $S \subset Y$ be open and S_o the inverse image of S.

If $S_o = \varnothing$ it is open.

Let $S_o \neq \varnothing$ for any $x_o \in S_0$. Let $y_0 = Tx_o$.

Since S is open, it contains an $\varepsilon-$neighbourhood of N of y_o. Since T is continuous, x_o has a $\delta-$neighbourhood N_o which is mapped into N. Since $N \subset S$, we have $N_o \subset S_o$, so that S_o is open because $x_o \in S_o$ was arbitrary.

Suppose the inverse image of every open set in Y is an open set in X. Then for every $x_o \in X$ and any $\varepsilon-$neighbourhood N of Tx_o, the inverse image N_o of N, since N is open and $\| f(x) - f(a) \| < \varepsilon$ for every $x \in X$ such that $\| x - a \| < \delta$ contains x_o. Hence N_o also contains a $\delta-$neighbourhood of x_o by definition. Since $x_o \in X$ was arbitrary, T is continuous.

3.7 Definition(Convergence of a sequence)

A sequence x_n in a metric space (X,d) is said to converge or to be convergent if there is an $x \in X$ such that $\lim_{n\to\infty} d(x_n,x) = 0$, and x is called the limit of x_n and we write $\lim_{n\to\infty} x_n = x$.

We see that d yields the sequence of real numbers $a_n = d(x_n,x)$ whose convergence defines that of x_n. Hence if $x_n \to x$, given $\varepsilon > 0$ there is $N = N(\varepsilon)$ such that x_n with $n > N$ lie in the $\varepsilon-$neighbourhood $B(x,\varepsilon)$ of x.

Lemma 3.7.1

Let (X,d) be a metric space such that if $x_n \to x$ and $y_n \to y$ then $d(x_n,y_n) \to d(x,y)$

Proof

By the generalised triangle inequality

$d(x_n,y_n) \leq d(x_n,x)+d(x,y)+d(y,y_n)$. Hence we have

$d(x_n,y_n)-d(x,y) \leq d(x_n,x)+d(y,y_n)$. That implies

$\mid d(x_n,y_n)-d(x,y) \mid \leq d(x_n,x)+d(y,y_n) \to 0$ as $n \to \infty$.

Definition 3.7.1 (Cauchy sequnece)

A sequence x_n in a metric space (X,d) is said to be Cauchy if for every $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that $d(x_m,x_n) < \varepsilon$ for every $m,n > N$.

Theorem 3.7.2

Every convergent sequence in a metric space is a Cauchy sequence.

Proof

If $x_n \to x$, then for every $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that

$d(x_m,x) < \varepsilon$ for every $n > N$.

Hence by the generalized triangle inequality we obtain for every $m,n > N$,

$d(x_m,x_n) \leq d(x_m,x)+d(x,x_n) < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$.

This shows that x_n is Cauchy.

Notation 3.7.1

A Cauchy sequence in a metric space (X,d) need not to be convergent in (X,d).

Example 3.7.1

Let $X =]0,1[$. Define $a_n = \frac{1}{2n}$ for all $n \geq 1$.

Then a_n is convergent in the Euclidean line $\mathbb{R}$.

Hence a_n is a Cauchy sequence in $\mathbb{R}$ when the metric on $\mathbb{R}$ is defined by

$d(x,z) = \mid x-z \mid$ for all $x,z \in$

Thus a_n is a Cauchy sequence in X.

In $\mathbb{R}$ we have $a_n \to 0$ as $n \to \infty$.

Assume that a_n is convergent in X.

Let u be the unique element of X such that $a_n \to u$ as $n \to \infty$.

There is a contradiction $u = 0$ by the uniqueness in $\mathbb{R}$ and $0 < u < 1$.

Definition 3.7.2

Let X be a Hausdorff space.

1. *A subset A of X is said to* Compact *if for every nonempty collection $\{G_\gamma \mid \gamma \in \Gamma\}$ of open sets in X such that $A \subset \bigcup_{\gamma\in\Gamma} G_\gamma$ there exist finitely many elements $\gamma_1, \ldots, \gamma_n$ of Γ such that $A \subset \cup_{j=1}^{n} G_{\gamma_j}$.*

2. *A subset A of X is said to have the* finite intersection property *if for every nonempty collection $\{F_\gamma \mid \gamma \in \Gamma\}$ of closed sets in X such that $A \bigcap (\bigcap_{\gamma\in\Gamma} F_\gamma) = \varnothing$ there exist finitely many elements $\gamma_1, \ldots, \gamma_n$ of Γ such that*
   $A \bigcap (\bigcap_{j=1}^{n} F_{\gamma_j}) = \varnothing$.

Theorem 3.7.3

These two statements about a subset A of a Hausdorff space X are equivalent:

1. *A is compact.*

2. *A has a finite intersection property.*

Proof

Suppose 1 is true

Let $\{F_\gamma \mid \gamma \in \Gamma\}$ be a nonempty collection of closed sets in X such that

$A \bigcap (\bigcap_{j=1}^{n} F_{\gamma_j}) = \varnothing$.

Then $A \subset X - (\bigcap_{\gamma\in\Gamma} F_\gamma) = \bigcup_{\gamma\in\Gamma}(X - F_\gamma)$.

For every $\gamma \in \Gamma$ we have $X - F_\gamma$ is open.

Using condition 1 choose finitely many elements $\gamma_1, \ldots, \gamma_n$ of Γ such that

$A \subset \bigcup_{j=1}^{n}(X - F_{\gamma_j})$.

Then $A \subset X - \bigcap_{j=1}^{n} F_\gamma$ by De Morgan's Rule.

And so $A \bigcap (\bigcap_{j=1}^{n} F_{\gamma_j}) = \varnothing$.

Hence $1 \Rightarrow 2$.

Suppose 2 is true.

Let $\{G_\gamma \mid \gamma \in \Gamma\}$ be a nonempty collection of open sets in X such that $A \subset \bigcup_{\gamma\in\Gamma}$.

Then $(X - G_\gamma)$ is closed in X for all $\gamma \in \Gamma$ and

$A \cap (\bigcap_{\gamma\in\Gamma}(X - G_\gamma)) = A \bigcap (X - \bigcup_{\gamma\in\Gamma} G_\gamma) = \varnothing$

Using condition 2 choose finitely many elements such that $A \bigcap (\bigcap_{j=1}^{n}(X - G_{\gamma_j})) = \varnothing$

Then $A \bigcap [X - \bigcup_{j=1}^{n} G_{\gamma_j}] = \varnothing$ by De Morgan's Rule.

It follows that $A \subset \bigcup_{j=1}^{n} G_{\gamma_j}$

Hence $2 \Rightarrow 1$.

Theorem 3.7.4

If X is a Hausdorff space and $a \in X$, then $\{a\}$ is closed.

Proof

Let $b \in X - \{a\}$

Then $a \neq b$ and so there exist open sets D, G in X such that $a \in D$ and $b \in G$,

That implies $b \in G \subset X - D \subset X - \{a\}$.

$\therefore X - \{a\}$ is open and so $\{a\}$ is closed.

Theorem 3.7.5

If (Y,d) is a metric space and A is a compact subset of Y then A is bounded in (Y,d).

Proof

Choose $a \in Y$

For every positive integer n let $D(a,n) = \{y \in Y \mid d(a,y) < n\}$ the open ball with its centre a and radius n.

Then $A \subset Y = \bigcup_{j=1}^{\infty} D(a,n)$

Choose a positive integer k such that $A \subset \bigcup_{j=1}^{k} D(a,n) = D(a,n)$.

Then $A \subset D(a,k)$. Hence A is bounded.

Example 3.7.2

$S = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ *is compact because it is closed and bounded.*

$G = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 < 9\}$ *is not compact because it is open and bounded.*

$H = \{(x,y) \in \mathbb{R}^2 \mid y > 0\}$ *is not compact because it is open and not bounded.*

$U = \{(x,y) \in \mathbb{R}^2 \mid y \geq 0\}$ *is not compact because it is closed but not bounded.*

$L = \{(x,y) \in \mathbb{R}^2 \mid x = y\}$ *its not compact because it is not bounded.*

3.8 Theorem(Cauchy Schwarz's Inequality)

Let $(V, \langle . \rangle)$ be an inner product space over the field F. Then $\mid \langle x,y \rangle \mid \leq \sqrt{\langle x,x \rangle}\sqrt{\langle y,y \rangle}$ for all $x,y \in V$. Equality holds if and only if x and y are linearly dependent over the field $\mathbb{R}$.

Proof

If $x = 0$ then $\mid \langle x,y \rangle \mid = 0 = \sqrt{\langle x,x \rangle}\sqrt{\langle y,y \rangle}$

Assuming $x \neq 0$

Given $x,y \in V$ let $A = \langle x,x \rangle, B = \langle x,y \rangle$ and $D = \langle y,y \rangle$

Then it suffices to show that $\mid B \mid \leq \sqrt{A}\sqrt{D}$.

If $A = 0$, then $x = 0$ and so $B = 0$. In this case we have $\mid B \mid = 0 = \sqrt{A}\sqrt{A}$

If $A > 0$, let $\lambda = \frac{B}{A}$ then

$$\begin{aligned}
0 \leq &\quad \langle \lambda x - y, \lambda x - y \rangle \\
0 = &\quad \langle \lambda x, \lambda x \rangle - \langle \lambda x, y \rangle - \langle y, \lambda x \rangle + \langle y,y \rangle \\
0 = &\quad \mid \lambda \mid^2 A - \lambda \langle x,y \rangle - \bar{\lambda} \langle y,x \rangle + D \\
0 = &\quad \mid \lambda \mid^2 A - \frac{B\bar{B}}{A} - \frac{B\bar{B}}{A} + D \\
0 \leq &\quad D - \frac{\mid B \mid^2}{A} \\
\mid B \mid^2 \leq &\quad AD \\
\mid B \mid \leq &\quad \sqrt{A}\sqrt{D}
\end{aligned}$$

That concludes that $| \langle x,y \rangle | \leq \sqrt{\langle x,x \rangle}\sqrt{\langle y,y \rangle}$

Schwarz's inequality is equivalent to $| \langle x,y \rangle | \leq \| x \| \| y \|$

Suppose x, y are linearly dependent.

Let $y = cx$ where $c \in \mathbb{C}$. Then

$$\begin{aligned} | \langle x,y \rangle | &= | \langle x,cx \rangle | \\ &= | \bar{c} || \langle x,x \rangle | \\ &= | c | \| x \| \| x \| \\ &= \| x \| \| cx \| \\ &= \| x \| \| y \| \end{aligned}$$

Now let x and y be vectors such that $\langle x,y \rangle \langle y,x \rangle = \langle x,x \rangle \langle y,y \rangle$

x and y are linearly dependent if we can show that $\langle y,y \rangle x - \langle x,y \rangle y = 0$

$\langle \langle y,y \rangle x - \langle x,y \rangle y \rangle = \langle y,y \rangle^2 \langle x,x \rangle - \langle y,y \rangle \langle y,x \rangle \langle x,y \rangle - \langle x,y \rangle \langle y,y \rangle \langle y,x \rangle +$
$\langle x,y \rangle \langle y,y \rangle \langle y,x \rangle = 0$ completing the proof.

3.8.1 Pythagora's Theorem

If $(V, \langle . \rangle)$ is an inner product space $x, y \in V$ and $\langle x,y \rangle = 0$ then
$\| x+y \|^2 = \| x-y \|^2 = \| x \|^2 + \| y \|^2$

Proof

$$\begin{aligned} \| x+y \|^2 &= \langle x+y, x+y \rangle \\ &= \langle x,x \rangle + \langle x,y \rangle + \langle y,x \rangle + \langle y,y \rangle \\ &= \| x \|^2 + \| y \|^2 \end{aligned}$$

and

$$\begin{aligned}\| x-y \|^2 &= \langle x-y, x-y\rangle \\ &= \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle \\ &= \| x \|^2 + \| y \|^2\end{aligned}$$

Theorem 3.8.1 (Parallelogram Law)

Let $(V,\langle . \rangle)$ be an inner product space, then for arbitrary $x,y \in V$

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

Proof

By Pythagora's Theorem $\| x+y \|^2 = \| x-y \|^2 = \| x \|^2 + \| y \|^2$

Then $\| x+y \|^2 + \| x-y \|^2 = \| x \|^2 + \| y \|^2 + \| x \|^2 + \| y \|^2$

Hence $\| x+y \|^2 + \| x-y \|^2 = 2(\| x \|^2 + \| y \|^2)$

Theorem 3.8.2 (Appolonius)

Let $(V,\langle . \rangle)$ be an inner product space, then for arbitrary $a,b \in V$

$\| a \|^2 + \| b \|^2 = 2(\| \frac{a+b}{2} \|^2 + \| \frac{a-b}{2} \|^2)$

Proof

$$\begin{aligned}\| \frac{a+b}{2} \|^2 + \| \frac{a-b}{2} \|^2 &= \langle \frac{a+b}{2}, \frac{a+b}{2} \rangle + \langle \frac{a-b}{2}, \frac{a-b}{2} \rangle \\ &= \frac{1}{2} \bullet \frac{1}{2} \langle a+b, a+b \rangle + \frac{1}{2} \bullet \frac{1}{2} \langle a-b, a-b \rangle \\ &= \frac{1}{2} \{ \| a \|^2 + \| b \|^2 \}\end{aligned}$$

That implies that $\| a \|^2 + \| b \|^2 = 2(\| \frac{a+b}{2} \|^2 + \| \frac{a-b}{2} \|^2)$.

Theorem 3.8.3 (Polarization Identity)

If $(V, \langle . \rangle)$ be an inner product space and arbitrary $x, y \in V$ then

$\langle x,y \rangle = \frac{1}{4}\{\| x+y \|^2 - \| x-y \|^2 + i \| x+iy \|^2 - i \| x-iy \|^2\}$

Proof

$\| x+y \|^2 = \langle x+y, x+y \rangle = \| x \|^2 + \| y \|^2 + \langle x,y \rangle + \langle y,x \rangle$

$\| x-y \|^2 = \langle x-y, x-y \rangle = \| x \|^2 + \| y \|^2 - \langle x,y \rangle - \langle y,x \rangle$

$\| x+iy \|^2 = \langle x+iy, x+iy \rangle = \| x \|^2 - \| y \|^2 + \langle x,iy \rangle + \langle iy,x \rangle$

$\| x-iy \|^2 = \langle x-iy, x-iy \rangle = \| x \|^2 - \| y \|^2 - \langle x,iy \rangle - \langle iy,x \rangle$

Substituting these into the right hand side of the expression we get

$\frac{1}{4}\{\| x+y \|^2 - \| x-y \|^2 + i \| x+iy \|^2 - i \| x-iy \|^2\} = \frac{1}{4}\{2\langle x,y \rangle + 2\langle y,x \rangle + 2i\langle x,iy \rangle + 2i\langle iy,x \rangle\}$

$\frac{1}{4}\{\| x+y \|^2 - \| x-y \|^2 + i \| x+iy \|^2 - i \| x-iy \|^2\} = \frac{1}{4}\{2\langle x,y \rangle + 2\langle y,x \rangle + 2\langle x,y \rangle - 2\langle y,x \rangle\}$

$\frac{1}{4}\{\| x+y \|^2 - \| x-y \|^2 + i \| x+iy \|^2 - i \| x-iy \|^2\} = \frac{1}{4}\{4\langle x,y \rangle\} = \langle x,y \rangle$

3.8.2 Orthogonal and Orthonormal Sets

In a vector space V a basis means a set of linearly independent vectors such that any vector in V can be written as a linear combination of the element in the basis.

Since an inner product space permits to establish when two vectors are orthogonal, in an inner product space we introduce the concept of orthonormal basis, where the condition of linearly independence will be replaced by the orthogonality condition.

The nice feature of this approach is represented by the fact we able to give explicit representation of orthogonal basis L_2. Before we go can give a defintion of orthonormal basis, we first introduce the concept of orthogonal system.

1. *A* set S in an inner product space E is called *orthogonal set* if $\langle x,y\rangle = 0$ for each $x,y \in S$ and $x \neq y$.

2. A non empty subset A of E is said to be *orthogonal* if $\langle a,b\rangle = 0$ for each $a,b \in A$ and $a \neq b$.

3. The set S is called *orthonormal set* if it is orthogonal set and $\| x \| = 1$ for each $x \in S$.

Remark.

If x is orthogonal to each $x_1,\ldots,x_n \in S$, then x is orthogonal to every linear combination y of vectors $x_1,\ldots,x_n \in S$.

Check:

Let $y = \sum_{i=1}^{n} \alpha_i x_i$ and $\langle y,x_i\rangle = 0$ for all $i = 1,2,\ldots,n$ then

$\langle x,y\rangle = \langle x, y = \sum_{i=1}^{n} \alpha_i x_i\rangle = \sum_{i=1}^{n} \bar{\alpha}_i \langle x,x_i\rangle = 0$

Example 3.8.1

1. In $\mathbb{R}^3$ let $i = (1,0,0)$, $j = (0,1,0)$ and $k = (0,0,1)$ then $\{i,j,k\}$ is an orthonormal set in $\mathbb{R}^3$.

2. $S = \{f_m(x) = \frac{e^{imx}}{\sqrt{2}} \mid m \in \mathbb{Z}\}$ is an orthonormal system for $L_2([-\pi,\pi])$ equipped with an inner product $\langle f,g\rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx$ for all $f,g \in L_2([-\pi,\pi])$.

Proof

$\langle f_m, f_n\rangle = \int_{-\pi}^{\pi} \mid f_m(x) \mid^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}e^{-imx}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} dx = 1 < \infty.$

This shows that $\| f_m \| = 1$ for all $n \in \mathbb{Z}$.

For orthogonality we take $f_m, f_n \in S$ with $n \neq m$ then

$$
\begin{aligned}
\langle f_m, f_n \rangle &= \int_{-\pi}^{\pi} f_m(x)\overline{f_n(x)}dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imx}e^{-inx}dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)x}dx \\
&= \frac{1}{2\pi i(m-n)}e^{i(m-n)x}|_{-\pi}^{\pi} \\
&= \frac{1}{2\pi i(m-n)}\left\{e^{i(m-n)\pi} - e^{-i(m-n)\pi}\right\} \\
&= \frac{1}{2\pi i(m-n)}\{\cos(m-n)\pi + i\sin(m-n)\pi - (\cos(m-n)\pi - i\sin(m-n)\pi)\} \\
&= \frac{\sin(m-n)\pi}{\pi(m-n)} = 0
\end{aligned}
$$

Since $(m-n)$ is an integer and sine varnishes at any integer multiple of π.

We present here some useful examples for the study of trigonometric functions. Let $a_o(x) = \frac{1}{\sqrt{2\pi}}$ for every $x \in [-\pi, \pi]$. If n is a positive integer let $a_n(x) = \frac{1}{\sqrt{\pi}}\cos nx$ and $b_n = \frac{1}{\sqrt{\pi}}\sin nx$. Finally let V be the set of all continuous real-valued functions on $[-\pi, \pi]$. Then V is a vector space over $\mathbb{R}$ and an inner product $\langle , \rangle$ is defined on V by the formula $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$.

Notation 3.8.1 *It is important to note that $\{a_o, a_1, a_2, \dots\} \bigcup \{b_1, b_2, b_3, \dots\}$ is an orthonormal set in V.*

Proof

$\langle a_o, a_o \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} dx = 1$

$\langle a_n, a_n \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2 nxdx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\{1 + \cos 2nx\}dx = 1$ for every positive integer n.

$\langle b_n, b_n \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2 nxdx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\{1 - \cos 2nx\}dx = 1$ for every positive integer n.

$\langle a_o, a_n \rangle = \frac{1}{\sqrt{2\pi}} \frac{1}{\pi} \int_{-\pi}^{\pi} \cos nx dx = 0$ for every positive integer n.

$\langle a_o, b_n \rangle = \frac{1}{\sqrt{2\pi}} \frac{1}{\pi} \int_{-\pi}^{\pi} \sin nx dx = 0$ for every positive integer n.

$\langle a_n, b_n \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos nx \sin nx dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin 2nx dx = 0$ for every positive integer n.

$\langle a_m, b_n \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos mx \sin nx dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\sin(m+n)x + \sin(m-n)x\} dx = 0$ for every pair of positive integers m,n such that $m \neq n$.

$\langle a_m, a_n \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos mx \cos nx dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\cos(m+n)x + \cos(m-n)x\} dx = 0$ for every pair of positive integers m,n such that $m \neq n$.

$\langle b_m, b_n \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin mx \sin nx dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\cos(m-n)x - \cos(m+n)x\} dx = 0$ for every pair of positive integers m,n such that $m \neq n$.

Definition 3.8.1 *L_2 is the Hilbert space of all Lebesgue measurable functions f on the closed interval $[-\pi, \pi]$ such that $\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$*
For every pair $f, g \in L_2$ we have $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$ and $\|f\| = \sqrt{\langle f, f \rangle}$
Let H be a Hilbert space and $\{a_n | n = 1, 2, 3, \ldots\}$ an orthonormal set in H.
If $f \in H$ then the numbers $\{\langle f, a_n \rangle | n = 1, 2, \ldots\}$ are called the FOURIER COEFFICIENTS of f with respect to the orthonormal set $\{a_n | n = 1, 2, 3, \ldots\}$
$\sum_{n=1}^{\infty} \langle f, a_n \rangle a_n \in H$ then $\langle f, a_o \rangle a_o + \sum_{n=1}^{\infty} \langle f, a_n \rangle a_n$ is called FOURIER SERIES on f with respect to the orthonormal set $\{a_n | n = 1, 2, 3, \ldots\}$

Theorem 3.8.4 *(Parseval's identity)*
The Parseval's identity is $\|f\|^2 = |\langle f, a_o \rangle|^2 + \sum_{n=1}^{\infty} \left\{ |\langle f, a_n \rangle|^2 + |\langle f, b_n \rangle|^2 \right\}$ where $a_o(x) = \frac{1}{\sqrt{2\pi}}$ for every $x \in [-\pi, \pi]$ and for every positive integer n $a_n(x) = \frac{1}{\sqrt{\pi}} \cos nx$ while $b_n = \frac{1}{\sqrt{\pi}} \sin nx$.

Example 3.8.2 *If $f(x) = x$, calculate $\langle a_o, f\rangle, \langle b_n, f\rangle, \langle a_n, f\rangle, \langle f, f\rangle$ and hence prove that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$*

Proof

$\langle f, f\rangle = \|f\|^2 = \int_{-\pi}^{\pi} x \bullet x dx = 2\left[\frac{x^3}{3}\right]_0^{\pi} = \frac{2\pi^3}{3}$

$\langle a_o, f\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x dx = 0$ since x is odd

$\langle a_n, f\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x \cos nx dx = 0$ since $x \cos nx$ is odd

$\langle b_n, f\rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x \sin nx dx = \frac{2}{\sqrt{\pi}} \int_0^{\pi} x \sin nx dx = \frac{2}{\sqrt{\pi}} \left[\frac{-x}{n} \cos nx + \frac{1}{n^2} \sin nx\right]_0^{\pi} = \frac{-2\pi}{\sqrt{\pi} n}(-1)^n$

Using the Parseval's identity
$\|f\|^2 = |\langle f, a_o\rangle|^2 + \sum_{n=1}^{\infty} \left\{|\langle f, a_n\rangle|^2 + |\langle f, b_n\rangle|^2\right\}$

$\frac{2\pi^3}{3} = |0|^2 + \sum_{n=1}^{\infty} \left\{|0|^2 + |\frac{-2\pi}{\sqrt{\pi} n}(-1)^n|^2\right\}$

$\frac{2\pi^3}{3} = \sum_{n=1}^{\infty} \frac{4\pi^2}{\pi n^2}$ dividing through by 4π

$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2}$ Hence the proof.

Definition 3.8.2

An orthonormal set M in an inner product space $(V, \langle . \rangle)$ is called a maximal orthonormal set *if $A = M$ for every orthonormal set A in V such that $M \subset A$.*

Theorem 3.8.5

These two statements about an orthonormal set M in an inner product space $(V, \langle . \rangle)$ *are equivalent:*

1. *M is a maximal orthonormal set in* $(V, \langle . \rangle)$.

2. $w = 0$ *for every* $w \in V$ *such that* $\langle u, w \rangle = 0$ *for all* $u \in M$.

Proof

Given that 1 is true assume that 2 is false.

Choose $z \in V$ such that $\langle u, z \rangle = 0$ for all $u \in M$ and $z \neq 0$.

Let $b = \frac{1}{\|z\|} z$

Then $\langle b, b \rangle = \langle \frac{1}{\|z\|} z, \frac{1}{\|z\|} z \rangle = 1$

Then $M \bigcup \{b\}$ is an orthonormal set in V, $M \subset M \bigcup \{b\}$ and $M \neq M \bigcup \{b\}$.

This gives a contradiction $M = M \bigcup \{b\}$ and $M \neq M \bigcup \{b\}$. The assumption is false.

Hence $1 \Rightarrow 2$.

Given that 2 is true assume that 1 is false.

Choose an orthonormal set L in V such that $M \subset L$ and $L \neq M$.

Choose $f \in L - M$ then the inner product $\langle f, u \rangle = 0$ for all $u \in M$.

That implies a contradiction $f = 0$ and $f \neq 0$ because $\| f \| = 1$.

The assumption is false. Hence $2 \Rightarrow 1$.

Notation 3.8.2

In $\mathbb{R}^3$ *we have three unit vectors* e_1, e_2, e_3 *and they form a* basis *for* $\mathbb{R}^2$ *so for every* $x \in \mathbb{R}^2$ *has a unique representation* $x = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$.

There is a great deal of advantage of orthogonality. Given x, *we can find the unknown coefficients* $\alpha_1, \alpha_2, \alpha_3$.

For instance α_1 *is obtained by* $\alpha_1 = \langle x, e_1 \rangle$ *since:*

$\langle x, e_1 \rangle = \alpha_1 \langle e_1, e_1 \rangle + \alpha_2 \langle e_2, e_1 \rangle + \alpha_3 \langle e_3, e_1 \rangle$

Theorem 3.8.6

Orthonormal systems are linearly independent.

Proof

Let S be an orthogonal system.

Suppose $\sum_{k=1}^{n} \alpha_k x_k = 0$ for some $x_1, \ldots, x_n$ scalars $\alpha_1, \ldots, \alpha_n$.Then

$$\begin{aligned} 0 &= \sum_{m=1}^{n} \langle 0, \alpha_m x_m \rangle \\ &= \sum_{m=1}^{n} \langle \sum_{k=1}^{n} \alpha_k x_k, \alpha_m x_m \rangle \\ &= \sum_{m=1}^{n} \mid \alpha_m \mid^2 \| x_m \|^2 \end{aligned}$$

This implies that $\alpha_m = 0$ for each $m \in \mathbb{N}$. Thus $x_1, \ldots, x_n$ are linearly independent.

3.9 Properties of Orthonormal Systems

Lemma 3.9.1

If $\{x_1, \ldots, x_\alpha\}$ *is an orthogonal set then*

$\| x_1 + \ldots, x_\alpha \|^2 = \| x_1 \|^2 + \cdots + \| x_\alpha \|^2$.

Consequently, $\langle x_i, x_j \rangle = 0$*, if* $i \neq j$ *then*

$$\begin{aligned} \left\| \sum_{j=1}^{\alpha} x_j \right\|^2 &= \left\langle \sum_{j=1}^{\alpha} x_j, \sum_{k=1}^{\alpha} x_k \right\rangle \\ &= \sum_{j=1}^{\alpha} \sum_{k=1}^{\alpha} \langle x_j, x_k \rangle \\ &= \sum_{j=1}^{\alpha} \langle x_j, x_j \rangle \\ &= \sum_{j=1}^{\alpha} \| x_j \|^2 \end{aligned}$$

3.9.1 Theorem(Pythagorean Formula)

If $\{x_1, \ldots, x_\alpha\}$ is an orthogonal set in an inner product space then,

$\| \sum_{k=1}^{n} x_k \|^2 = \sum_{k=1}^{n-1} \| x_k \|^2$

Proof

If $x_1 \perp x_2$, then $\| x_1 + x_2 \|^2 = \| x_1 \|^2 + \| x_2 \|^2$

Thus, the theorem is true for $n = 2$.

Assume now that the theorem hold for $n-1$, that is

$\| \sum_{k=1}^{n} x_k \|^2 = \sum_{k=1}^{n} \| x_k \|^2$

Set $x = \sum_{k=1}^{n-1} x_k$ and $y = x_n$

Since $x \perp x$, we have

$$\begin{aligned}
\| \sum_{k=1}^{n} x_k \|^2 &= \| x + y \|^2 \\
&= \| x \|^2 + \| y \|^2 \\
&= \sum_{k=1}^{n-1} \| x_k \|^2 + \| x_n \|^2 \\
&= \sum_{k=1}^{n} \| x_k \|^2
\end{aligned}$$

This completes the proof.

Lemma 3.9.2

If A is an orthonormal set in an inner product space V, $f \in V$ and $a_1, \ldots, a_k$ are finitely many elements of A, then $\sum_{n=1}^{k} | \langle f, a_n \rangle |^2 \leq \| f \|^2$

Proof

Let $x = \sum_{n=1}^{k} \langle f, a_n \rangle a_n$ and $y = f - x = f - \sum_{n=1}^{k} \langle f, a_n \rangle a_n$. Then

$$\begin{aligned}
\langle x, y \rangle &= \langle \sum_{n=1}^{k} \langle f, a_n \rangle a_n, f - \sum_{n=1}^{k} \langle f, a_n \rangle a_n \rangle \\
&= \sum_{n=1}^{k} \langle f, a_n \rangle \langle a_n, f \rangle - \sum_{n=1}^{k} \langle f, a_n \rangle \overline{\langle f, a_n \rangle} \\
&= \sum_{n=1}^{k} | \langle f, a_n \rangle |^2 - \sum_{n=1}^{k} | \langle f, a_n \rangle |^2 \\
&= 0
\end{aligned}$$

Hence $\sum_{n=1}^{k} \mid \langle f, a_n \rangle \mid^2 = \| x \| \leq \| x \|^2 + \| y \|^2 = \| f \|^2$ (by Pythagora's Theorem)

Hence the prove.

Theorem 3.9.3

If $\{a_n \mid n = 1, 2, \dots\}$ is a denumerable set in an inner product space V and $f \in V$ then $\sum_{n=1}^{\infty} \mid \langle f, a_n \rangle \mid^2 \leq \| f \|^2$

Proof

By the lemma above $\sum_{n=1}^{k} \mid \langle f, a_n \rangle \mid^2 \leq \| f \|^2$ for every positive integer k.

Let $k \to \infty$. Then

$\sum_{n=1}^{\infty} \mid \langle f, a_n \rangle \mid^2 \leq \| f \|^2$

Example 3.9.1

Consider the Hilbert space $L_2([-\pi, \pi])$ and the orthonormal system

$S = \{f_n(x) = \frac{\sin(nx)}{\sqrt{\pi}} \mid n \in \mathbb{N}\}$ is indeed an orthonormal system.

Proof

$$
\begin{aligned}
1. \int_{-\pi}^{\pi} f_n^2(x)dx &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin^2(nx)dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} [1 - cos(2nx)]dx \\
&= \frac{1}{2\pi}\{x \mid_{-\pi}^{\pi} - \frac{1}{2n}\sin(2nx) \mid_{-\pi}^{\pi}\} \\
&= \frac{1}{2\pi} 2\pi \\
&= 1 < \infty
\end{aligned}
$$

This shows that $\| f_n \| = 1$ for all $n \in \mathbb{N}$

2. Let $m, n \in \mathbb{N}$ with $m \neq n$ then

$$
\begin{aligned}
\langle f_n, f_m \rangle &= \frac{1}{\pi}\int_{-\pi}^{\pi} \sin(nx)\sin(mx)dx \\
&= \frac{2}{\pi}\int_{0}^{\pi} \sin(nx)\sin(mx)dx \\
&= \frac{2}{\pi}\bullet\frac{1}{2}\int_{0}^{\pi} [\cos(n-m)x - \cos(n+m)x]dx \\
&= \frac{2}{\pi}\bullet\frac{1}{2}\left\{\int_{0}^{\pi} \cos(n-m)xdx - \int_{0}^{\pi} \cos(n+m)xdx\right\} \\
&= \frac{1}{\pi}\left[\frac{1}{n-m}\sin(n-m)x\,|_0^{\pi} - \frac{1}{n+m}\sin(n+m)x|_0^{\pi}\right] \\
&= 0
\end{aligned}
$$

In particular we want to show that in general $f \neq \sum_{n=1}^{\infty}\langle f, f_n\rangle f_n$ Let us take for instance $f(x) = \cos(x)$. Then

$$
\begin{aligned}
\langle f, f_n \rangle &= \int_{-\pi}^{\pi} f(x) f_n(x)dx \\
&= \frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} \sin(nx)\cos(x)dx \\
&= 0
\end{aligned}
$$

Hence

$$
\begin{aligned}
\sum_{n=1}^{\infty}\langle f, f_n\rangle f_n &= \sum_{n=1}^{\infty} 0\bullet\frac{\sin(nx)}{\sqrt{\pi}} \\
&= 0 \neq \cos(x)
\end{aligned}
$$

3.9.2 Complete Orthonormal Sequence

An orthonormal sequence x_n in an inner product space E is said to be complete if for every $x \in E$ we have $x = \sum_{n=1}^{\infty}\langle x, x_n\rangle x_n$

3.9.3 Orthonormal Basis

An orthonormal set A in an inner product space V is called an *orthonormal basis* if for every $x \in V$ has a unique representation $x = \sum_{n=1}^{\infty} \alpha_n x_n$.
Where $\alpha_n \in \mathbb{R}$ and x_n's are distinct elements of A.

Theorem 3.9.4

An orthonormal sequence x_n in a Hilbert space H is complete if and only if $\langle x, x_n \rangle = 0$ for all $n \in \mathbb{N}$. which implies that $x = 0$

Proof

Suppose x_n is a complete orthonormal sequence in H. Then every $x \in H$ has a representation $x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$.
Thus, if $\langle x, x_n \rangle = 0$ for all $n \in \mathbb{N}$ then $x = 0$.
Conversely, suppose $\langle x, x_n \rangle = 0$ for all $n \in \mathbb{N}$ implies $x = 0$
Let x be an element of H. Define $y = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$ since for every $n \in \mathbb{N}$,

$$\begin{aligned}
\langle x - y, x_n \rangle &= \langle x, x_n \rangle - \langle \sum_{k=1}^{\infty} \langle x, x_k \rangle x_k, x_n \rangle \\
&= \langle x, x_n \rangle - \sum_{k=1}^{\infty} \langle x, x_k \rangle \langle x_k, x_n \rangle \\
&= \langle x, x_n \rangle - \langle x, x_n \rangle \\
&= 0
\end{aligned}$$

We have $x - y = 0$ and hence $x = \sum_{n=1}^{\infty} \langle x, x_n \rangle x_n$.

Theorem 3.9.5

Let $(V, \langle . \rangle)$ be an inner product space over $\mathbb{R}$. When we are seeking examples of bounded linear functional on V, we choose any $a \in V$ and define $L : V \to \mathbb{R}$ by $L(x) = \langle x, a \rangle$. Then L is a linear functional and bounded.

Proof

For all $x, y \in V$ and $\lambda \in \mathbb{R}$ then :

$$\begin{aligned} L(x+y) &= \langle x+y, a\rangle \\ &= \langle x, a\rangle + \langle y+a\rangle \\ &= L(x) + L(y) \\ L(\lambda x) &= \langle \lambda x, a\rangle \\ &= \lambda\langle x, a\rangle \\ &= \lambda L(x) \end{aligned}$$

This proves that L is a linear function.

$$\begin{aligned} L(x,a) &= \langle x, a\rangle \\ &= \| x \| \| a \| \\ by \quad Schwarz's \quad inequality & \\ &= k \| x \| \end{aligned}$$

This proves that L is bounded.

Lemma 3.9.6

If a and b are non negative real numbers and α, β are positive real numbers such that $\alpha + \beta = 1$ then $a^{\alpha} b^{\beta} \leq \alpha a + \beta b$.

Proof

If $a = 0$ or $b = 0$, then $a^{\alpha} b^{\beta} = 0 \leq \alpha a + \beta b$

Suppose $a > 0$ and $b > 0$

Let $h(t) = \beta t^{\alpha} + \alpha t^{-\beta}$ for every positive real number t.

Then the derivative $h'(t) = \alpha\beta t^{\alpha-1} - \alpha\beta t^{-\beta-1}$

That implies that $h'(t) = \alpha\beta t^{-\beta}\{1 - \frac{1}{t}\}$

Thus $h(1) = \alpha + \beta = 1$

For every $t > 0$, $h'(t) > 0$

If $t > 1$

Then by the First mean-value theorem $h(t) - h(1) = (t-1)h'(\zeta)$, and $1 < \zeta < t$

Hence $h(t) > h(1) = 1$, for all $t > 1$.

If $0 < t < 1$

$$\begin{aligned} h(t) - h(1) &= -\{h(1) - h(t)\} \\ &= -(1-t)h'(\zeta) \quad for \quad all \quad t < \zeta < 1 \end{aligned}$$

That implies that $h'(\zeta) < 0$, therefore $h(t) - h(1) > 0$

That implies $h(t) > h(1)$

Thus for every real number $t > 0$ we have $h(t) \geq h(1) = 1$.

Let $t = \frac{b}{a}$ then

$\beta\frac{b^{\alpha}}{a^{\alpha}} + \alpha\frac{b^{-\beta}}{a^{-\beta}} = \beta\frac{b^{\alpha}}{a^{\alpha}} + \alpha\frac{a^{\beta}}{b^{\beta}} \geq 1$

Multiplying through by $a^{\alpha}b^{\beta}$ we get

$\alpha a^{\alpha+\beta} + \beta b^{\alpha+\beta} \geq a^{\alpha}b^{\beta}$

Hence $a^{\alpha}b^{\beta} \leq \alpha a + \beta b$ proved.

Definition 3.9.1

For every positive real number such that $p > 1$ let ℓ_p be the set of all mappings $\psi : \mathbb{N} \to F$ such that $\sum_{n=1}^{\infty} \mid \psi(n) \mid^p$ is convergent.

3.9.4 Holder's Inequality

Given $p > 1$ let $q = \frac{p}{p-1}$ so that $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

If $f \in \ell_p$ and $h \in \ell_p$ then

$\sum_{n=1}^{\infty} \mid f(n) \mid\mid h(n) \mid \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid h(n) \mid^q)^{\frac{1}{q}}$

Proof

If $\sum_{n=1}^{\infty} \mid f(n) \mid^{p} = 0$ then $f(n) = 0$ for all $n \geq 1$ and so

$\sum_{n=1}^{\infty} \mid f(n) \mid\mid h(n) \mid = 0 \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^{p})^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid h(n) \mid^{q})^{\frac{1}{q}}$

Similarly, if $\sum_{n=1}^{\infty} \mid h(n) \mid^{q} = 0$ then we have

$\sum_{n=1}^{\infty} \mid f(n) \mid\mid h(n) \mid = 0 \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^{p})^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid h(n) \mid^{q})^{\frac{1}{q}}$

If $\sum_{n=1}^{\infty} \mid f(n) \mid^{p} > 0$ and $\sum_{n=1}^{\infty} \mid h(n) \mid^{q} > 0$ let

$a_n = \frac{|f(n)|^p}{\sum_{n=1}^{\infty} |f(n)|^p}$ and $b_n = \frac{|h(n)|^q}{\sum_{n=1}^{\infty} |h(n)|^q}$

Applying the preceding lemma with $\alpha = \frac{1}{p}, \beta = \frac{1}{q}, a = a_n$ and $b = b_n$ then we get

$$\frac{|f(m)|}{(\sum_{m=1}^{\infty} |f(m)|^p)^{\frac{1}{p}}} \times \frac{|h(m)|}{(\sum_{m=1}^{\infty} |h(m)|^q)^{\frac{1}{q}}} \leq \frac{1}{p} \times \frac{|f(m)|^p}{\sum_{m=1}^{\infty} |f(m)|^p} + \frac{1}{q} \times \frac{|h(m)|^q}{\sum_{m=1}^{\infty} |h(m)|^q}$$

Summing up we get

$$\frac{\sum_{m=1}^{\infty} \mid f(m) \mid\mid h(m) \mid}{(\sum_{m=1}^{\infty} \mid f(m) \mid^{p})^{\frac{1}{p}} (\sum_{m=1}^{\infty} \mid h(m) \mid^{q})^{\frac{1}{q}}} \leq \frac{1}{p} + \frac{1}{q} = 1$$

Hence $\sum_{n=1}^{\infty} \mid f(n) \mid\mid h(n) \mid \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^{p})^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid h(n) \mid^{q})^{\frac{1}{q}}$

When $p = q = 2$ the result is the Cauchy Schwarz's inequality.

3.9.5 Minkowski's Inequality

If $p > 1$ and $f, g \in \ell_p$ then $f + g \in \ell_p$ and

$(\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^{p})^{\frac{1}{p}} \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^{p})^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid g(n) \mid^{p})^{\frac{1}{p}}$

Proof

Let $N_1 = \{k \in \mathbb{N} \mid \mid f(k) \mid \leq \mid g(k) \mid\}$

$N_2 = \{k \in \mathbb{N} \mid \mid f(k) \mid > \mid g(k) \mid\}$

Then $N_1 \bigcap N_2 = \varnothing$ and $N_1 \bigcup N_2 = \mathbb{N}$

If $k \in N_1$ then $\mid f(k) + g(k) \mid^p \leq \{2 \mid g(k) \mid\}^p = 2^p \mid g(k) \mid^p$ and

If $k \in N_2$ then $\mid f(k) + g(k) \mid^p \leq \{2 \mid f(k) \mid\}^p = 2^p \mid f(k) \mid^p$. Hence

$\sum_{n=1}^{\infty} \mid f(k) + g(k) \mid^p \leq 2^p \{\sum_{n=1}^{\infty} \mid f(n) \mid^p + \sum_{n=1}^{\infty} \mid g(n) \mid^p\} < \infty$

Next if $\sum_{n=1}^{\infty} \mid f(k) + g(k) \mid = 0$ then

$(\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p)^{\frac{1}{p}} = 0 \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid g(n) \mid^p)^{\frac{1}{p}}$

Next if $\sum_{n=1}^{\infty} \mid f(k) + g(k) \mid > 0$ then we write

$\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p = \sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^{p-1} \mid f(n) + g(n) \mid$

$\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p \leq \sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^{p-1} \mid f(n) \mid + \sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^{p-1} \mid g(n) \mid$

$\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p \leq \{\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^{(p-1)q}\}^{\frac{1}{q}} \{\sum_{n=1}^{\infty} \mid f(n) \mid^p\}^{\frac{1}{p}} + \{\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^{(p-1)q}\}^{\frac{1}{q}} \{\sum_{n=1}^{\infty} \mid g(n) \mid^p\}^{\frac{1}{p}}$

$\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p \leq (\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p)^{\frac{1}{q}} [(\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid g(n) \mid^p)^{\frac{1}{p}}]$

Dividing through by $\{\sum_{n=1}^{\infty} \mid f(n) + g(n) \mid^p\}^{\frac{1}{q}}$ we get

$(\sum_{n=1}^{\infty} \mid f(n)+g(n) \mid^p)^{\frac{1}{p}} \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid g(n) \mid^p)^{\frac{1}{p}}$

Notation 3.9.1

Suppose p is a real number such that $p > 1$ define a norm $\| \,.\, \|$ on ℓ_p by

$\| f \| = \{\sum_{n=1}^{\infty} \mid f(n) \mid^p\}^{\frac{1}{p}}$

(a). If $f, g \in \ell_p$ then $f+g \in \ell_p$ and

$(\sum_{n=1}^{\infty} \mid f(n)+g(n) \mid^p)^{\frac{1}{p}} \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid g(n) \mid^p)^{\frac{1}{p}}$ by the Minkowski's inequality.

(b). If $\lambda \in F$ and $f \in \ell_p$ then $(\sum_{n=1}^{\infty} \mid \lambda f(n) \mid^p)^{\frac{1}{p}} = \mid \lambda \mid (\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}}$

1. $\| f \| \geq 0$ since $\mid f(n) \mid \geq 0$ if and only if $\sum_{n=1}^{\infty} \mid f(n) \mid \geq 0$

2. $\| f \| = 0$ if and only if $\mid f \mid = 0$ if and only if $f(n) = 0$

3. If $\lambda \in F$ and $f \in \ell_p$ then

$\| \lambda f \| = (\sum_{n=1}^{\infty} \mid \lambda f(n) \mid^p)^{\frac{1}{p}} = \mid \lambda \mid \| f \|$

4. If $f, g \in \ell_p$ then

$\| f+g \| = (\sum_{n=1}^{\infty} \mid f(n)+g(n) \mid^p)^{\frac{1}{p}} \leq (\sum_{n=1}^{\infty} \mid f(n) \mid^p)^{\frac{1}{p}} (\sum_{n=1}^{\infty} \mid g(n) \mid^p)^{\frac{1}{p}} \leq \| f \| + \| g \|$

The conclusion is that $(\ell_p, \| \,.\, \|)$ is a normed vector space.

3.10 Linear Mapping

Let X and Y be a vector space and let T be a mapping from X to Y. If $y = T(x)$ then y is called the image of x. If A is a subset of X, then $T(A)$ denotes the image of the set A i.e. $T(A)$ is the set of all vectors in Y which are the images of elements of A. If B is a subset of Y, then $T^{-1}(B)$ denotes the inverse of B i.e. $T^{-1}(B)$ is the set of all vectors in X whose images are elements of B.

$T(A) = \{y \in Y \mid y = T(x)\}$ for some $x \in A$ and $\{T^{-1}(B) : T(x) \in B\}$.

To specify the domain of T, it is denoted by $D(T)$. The set $T(D(T))$ is called the range of T and denoted by $R(T)$ i.e. $R(T) = \{y \in Y : T(x) = y\}$ for some $x \in D(T)$. By the null space of T, denoted by $N(T)$, we mean the set of all vectors $x \in D(T)$, such that $T(x) = 0$.

3.10.1 Linear Mapping(Definition)

A mapping $T : X \to Y$ is called a linear mapping if : $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$ for all $x, y \in X$ and scalars $\alpha, \beta \in \mathbb{R}$.

If the linear space Y is replaced by the scalar field K, then the mapping T is the special case called a *Linear Functional on X.*

REMARKS: Since linear functionals are special forms of linear mappings, any result proved for linear mappings also holds for linear functionals.

Proposition 3.10.1

Let X and Y be two linear spaces over a scalar field K, and $T : X \to Y$ be a linear mapping. Then

(i) $T(0) = 0$

(ii) The range of T, $R(T) = \{y \in Y \mid T(x) = y\}$ for some $x \in X$ is a linear subspace of Y.

(iii) T is one-to-one if and only if $T(0) = 0$ implies that $x = 0$.

(iv) If T is one-to-one, then T^{-1} exist on $R(T)$ and $T^{-1} : R(T) \to X$ is a linear mapping.

Proof

(i) Since T is linear, we have $T(\alpha x) = \alpha T(x)$ for each $x \in X$ and each scalar α. If we take $\alpha = 0$ and (i) follows immediately.

(ii) We need to show that for $y_1, y_2 \in R(T)$ and α, β scalars and $\alpha y_1 + \beta y_2 \in R(T)$. Now $y_1, y_2 \in R(T)$ implies that there exist $x_1, x_2 \in X$ such that $T(x_1) = y_1$ $T(x_2) = y_2$. Moreover, $\alpha x + \beta y \in X$ since X is a linear space.Furthermore by the linearity of T,

$T(\alpha x + \beta y) = \alpha T(x_1) + \beta T(x_2) = \alpha y_1 + \beta y_2$.

Hence $\alpha y_1 + \beta y_2 \in R(T)$ and so $R(T)$ is a linear subspace of Y.

(iii) Assume that T is one-to-one. Clearly $T(x) = 0$ which implies that $T(x) = T(0)$ since T is linear and so $T(0) = 0$. But T is one-to-one , so $x = 0$.

Assume that whenever $T(u) = 0$, then u must be 0. We want to prove that T is one-to-one so let $T(x) = T(y)$. Then $T(x) - T(y) = 0$ and by linearity of T, $T(x-y) = 0$. By hypothesis, $x - y = 0$ which implies that $x = y$. Hence T is one-to-one.

(iv) Let $T : X \to Y$ be one-to-one. Then $T^{-1} : R(T) \to X$ exists. So we prove that T^{-1} is also linear.

Let $\alpha, \beta \in \mathbb{R}$ and $y_1, y_2 \in R(T)$. We know that $R(T)$ is a linear subspace of Y. Hence $\alpha y_1 + \beta y_2 \in R(T)$.

Let $x_1, x_2 \in X$ be such that $T(x_1) = y_1$ and $T(x_2) = y_2$ then $x_1 = T^{-1}(y_1)$ and $x_2 = T^{-1}(y_2)$. Moreover, $T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) = \alpha y_1 + \beta y_2$ so that
$T^{-1}(\alpha y_1 + \beta y_2) = \alpha x_1 + \beta x_2 = \alpha T^{-1}(y_1) + \beta T^{-1}(y_2)$. Thus T^{-1} is linear.

Definition 3.10.1

Continuous Mapping

Let X and Y be normed vector spaces. A mapping $F : X \to Y$ is said to be continuous at $x_o \in X$, if for any sequence x_n of elements of X converges to x_o, the sequence $(F(x_n))$ converges to $F(x_o)$

That is F is continuous at x_o if $\| x_n - x \| \to 0$ implies $\| F(x_n) - F(x_o) \| \to 0$ is continuous at every $x \in X$, then F is continuous.

Theorem 3.10.2

A linear mapping $T : X \to Y$ is continuous if and only if it is continuous at a point.

Proof

Assume that T is continuous at $x_o \in X$.

Let x be any arbitrary element of X and x_n be a convergent to x. Then the sequence $(x_n - x + x_o)$ converges to x_o and thus we have

$\| T(x_n) - T(x) \| = \| T(x_n - x + x_o) - T(x_o) \| \to 0.$

Hence the proof.

Definition 3.10.2

Let S be a set of real numbers. We say that S is bounded if there exist a positive integer M such that $| x | \leq M$.

3.11 Bounded Linear Mappings

Let X and Y be normed vector spaces over a scalar field,K, and let $T : X \to Y$ be a linear mapping. Then T is said to be bounded if there exists some constant$K \geq 0$ such that for each $x \in X$ we have $\| T(x) \| \leq K \| X \|$

Definition 3.11.1

If X and Y are normed vector spaces, then $B(X,Y)$ are the set of all bounded linear mappings from X to Y with norm defined by $\| T \| = sup_{\|x\|=1} \| T(x) \|$

Theorem 3.11.1

Let X and Y be normed vector space and $T : X \to Y$ be a linear mapping.Then these three statements are equivalent:

(i) T is continuous at 0

(ii) T is bounded

(iii) T is uniformly continuous

Proof

Suppose (i) is true

Choose a positive real number δ such that for every $x \in X$ we have $\| x - 0 \| < \delta$ which implies that $\| T(x) - T(0) \| < 1$.

That is $\| x \| < \delta$ implies $\| T(x) \| < 1$

Now suppose $z \in X$ and $z \neq 0$ then $\| \frac{\frac{1}{2}\delta z}{\|z\|} \| = \frac{1}{2}\delta\frac{\|z\|}{\|z\|} = \frac{1}{2}\delta < \delta$

And so $\| T(\frac{\frac{1}{2}\delta z}{\|z\|}) \| < 1$

$\Rightarrow \frac{\frac{1}{2}\delta z}{\|z\|} \| T(z) \| < 1$

$\Rightarrow \| T(z) \| < \frac{2\|z\|}{\delta}$

Let $K = \frac{2}{\delta}$ then for every $v \in X$ we have $\| T(v) \| \leq K \| v \|$

Therefore T is bounded and hence $(i) \Rightarrow (ii)$

Suppose (ii) is true

Given $\varepsilon > 0$, let $\delta = \frac{\varepsilon}{k}$ then for $z, v \in X$ we have $\| z - v \| < \frac{\varepsilon}{k}$

$\Rightarrow \| T(z) - T(v) \| = k \| T(z - v) \| < \varepsilon$

Therefore T is uniformly continuous. Hence $(ii) \Rightarrow (iii)$

$(iii) \Rightarrow (i)$ by definition.

Definition 3.11.2

A normed vector space $(V, \| \,.\, \|)$ is said to be complete if every Cauchy sequence in V is convergent in V. When V is complete then V is called a Banach space.

Remark

1. A metric space (X, d) is said to be complete if every Cauchy sequence in it is a convergent sequence.
2. A normed vector space $(X, \| \,.\, \|)$ is a Banach space if it is a complete metric space.
3. A Banach space $(X, \| \,.\, \|)$ is a Hilbert space if the norm $\| \,.\, \|$ is induced by an inner product space.

Theorem 3.11.2

Let a, b be real numbers such that $a < b$. Denote by $C[a, b]$ the set of all continuous real-valued functions on the closed interval $[a, b]$. For every $f \in C[a, b]$ let

$\| f \| = sup\{| f(x) | \,|\, a \leq x \leq b\}$. Then $C[a, b]$ is a Banach space.

Proof

Suppose f_n is a Cauchy sequence in $C[a, b]$.

Then for every pair of positive integers m, n and every $x \in [a, b]$

$| f_m(x) - f_n(x) | = | (f_m - f_n)(x) | \leq \| f_m - f_n \|$

Therefore, $f_n(x)$ is a convergent sequence of real numbers for all $x \in X$.

Let $f(x) = \lim_{n \to \infty} f_n(x)$ for all $x \in X$.

Then $f_n(x)$ is a convergent sequence of real numbers.

Thus $f_n \to f$ uniformly on $[a, b]$ as $n \to \infty$.

Hence f is continuous on $[a,b]$.

Thus $f \in C[a,b]$ and $f_n \to f$ as $n \to \infty$ in $C[a,b]$.

Thus, its a Banach space since every Cauchy sequence in $C[a,b]$ is a convergent sequence in $C[a,b]$.

Theorem 3.11.3

If p is a real number such that $p \geq 1$ then every Cauchy sequence in ℓ_p is a convergent sequence in ℓ_p.

Proof

The norm on ℓ_p is given by $\| f \| = (\sum_{r=1}^{\infty} | f(r) |^p)^{\frac{1}{p}}$

Let $f(n)$ be a Cauchy sequence in ℓ_p.

Given a positive real number η, let t be a positive integer such that $\| f_m - f_n \| < \eta$ for all $m,n \geq t$.

Then for every positive integer j, $| f_m(j) - f_n(j) | \leq \| f_m - f_n \| < \eta$ for all $m,n \geq t$

This proves that $f_n(j)$ is a Cauchy sequence of real numbers for every positive integer j.

Let $f(j) = \lim_{n \to \infty} f_n(j)$ for every positive integer j, then $f : N \to F$ is a mapping.

Choose a positive integer r such that $\| f_m - f_n \| < 1$ for all $m,n \geq r$.

Then in particular $\| f_m - f_r \| < 1$ for all $n \geq r$

$\Rightarrow \| f_n \| - \| f_r \| \leq \| f_m - f_r \| < 1$ for all $n \geq r$

and so $\| f_n \| < 1 + \| f_r \|$ for all $n \geq 1$

Hence for every positive integer T, in particular $p = 1$,

$\sum_{r=1}^{T} | f_n(r) | \leq \| f_n \| < 1 + \| f_r \|$

$\Rightarrow \sum_{r=1}^{T} | f_n(r) |^p \leq \| f_n \|^p < (1 + \| f_r \|)^p$

The conclusion is that if we let $n \to \infty$ we get $\sum_{r=1}^{T} | f_n(r) |^p \leq (1 + | f_r |)^p$ for every positive integer T.

Let $T \to \infty$ and so $\sum_{r=1}^{T} | f_n(r) |^p$ is convergent. Hence $f \in \ell_p$.

Given $\varepsilon > 0$ let k be a positive integer such that $\| f_m - f_n \| < \frac{\varepsilon}{2}$ for all $m,n \geq k$.

$\| f_m - f \| = (\sum_{r=1}^{\infty} | f_m(r) - f(r) |^p)^{\frac{1}{p}}$

$\| f_m - f \| = \lim_{n\to\infty} \lim_{T\to\infty} \{(\sum_{r=1}^{T} | f_m(r) - f(r) |^p)^{\frac{1}{p}}\} \leq (\sum_{r=1}^{\infty} | f_m(r) - f(r) |^p)^{\frac{1}{p}}$

$\| f_m - f \| \leq \frac{\varepsilon}{2} < \varepsilon$ for all $m \geq k$.

Hence $f_m \to f$ as $n \to \infty$ in ℓ_p.

Theorem 3.11.4

If X is a normed space and Y is Banach space, then $B(X,Y)$ defined by

$\| T \| = sup_{\|x\|=1} \| T(x) \|$ is a Banach space.

Proof

We need to show that $B(X,Y)$ is complete.

Let T_n be a Cauchy sequence in $B(X,Y)$, and let x be an arbitrary element of T.

Then $\| T_m(x) - T_n(x) \| \leq \| T_m - T_n \| \| x \| \to 0$ as $m,n \to \infty$ which shows that T_n is a Cauchy sequence in Y.

By the completeness of Y, there is a unique element $y \in Y$ such that $T_n \to y$ since x is an arbitrary element of X, this defines a mapping $T : X \to Y$ such that

$\lim_{n\to\infty} T_n(x) = T(x)$.

We will show that $T \in B(X,Y)$ and $\| T_n - T \| \to 0$.

Clearly, T is a linear mapping. Since Cauchy sequences are bounded, there exists a constant M such that $\| T_n \| \leq M$ for all $n \in \mathbb{Z}$.

Consequently,$\| T(x) \| = \| \lim_{n\to\infty} T_n(x) \| = \lim_{n\to\infty} \| T_n(x) \| \leq M \| x \|$

Therefore T is bounded and thus $T \in B(X,Y)$.

It remains to show that $\| T_n - T \| \to 0$.

Let $\varepsilon > 0$ and let k be such that $\| T_m - T_n \| < \varepsilon$ for every $m,n \geq k$

If $\| x \| = 1$ and $m,n \geq k$ then $\| T_m(x) - T_n(x) \| \leq \| T_m - T_n \| < \varepsilon$

By letting $n \to \infty$ with m remaining fixed we obtain

$\| T_m(x) - T(x) \| < \varepsilon$ for every $m \geq k$ and every $x \in X$.

With $\| x \| = 1$, means that $\| T_m - T_n \| < \varepsilon$ for all $m \geq k$ which completes the proof.

Definition 3.11.3

A topological space X is said to be Hausdorff *if for every pair of elements $x, y \in X$ such that $x \neq y$ there exist open sets D, G in X such that $x \in D, y \in G$ and $D \bigcap G = \varnothing$.*

Theorem 3.11.5

Every metric space is Hausdorff.

Proof

Let (X, d) be a metric space. Suppose $x, y \in X$ and $x \neq y$. Then $d(x, y)$ is a positive real number.

Let $D = \{z \in X \mid d(x,z) < \frac{1}{2}d(x,y)\}$ and $G = \{z \in X \mid d(y,z) < \frac{1}{2}d(x,y)\}$.

Then D and G are open sets in X and $x \in D, y \in G$ and $D \bigcap G = \varnothing$.

The conclusion is that (X, d) is Hausdorff.

Theorem 3.11.6

Let X be a Compact Hausdorff space and $\mathbb{C}(X)$ the vector space of all continuous real-valued functions on X. For every $f \in \mathbb{C}(x)$ let $\| f \| = sup\{| f(x) | \mid x \in X\}$ Then

(a) $\| \, . \, \|$ is a norm on $\mathbb{C}(x)$

(b) Every Cauchy sequence in $\mathbb{C}(x)$ is a convergent sequence in $\mathbb{C}(x)$ and so $(\mathbb{C}(x), \| \, . \, \|)$ is a Banach space over $\mathbb{R}$.

Proof

(i) Choose $z \in X$ then $\| f \| \geq | f(z) | \geq 0$ for all $f \in \mathbb{C}(x)$

(ii) for $h \in \mathbb{C}(x)$, $\| h \| = 0$ if and only if $0 \leq | h(x) | \leq \| h \| = 0$ for all $x \in X$.

Then $\| h \| = 0$ if and only if $h(x) = 0$ for all $x \in X$ if and only if h is the zero in $\mathbb{C}(x)$.

(iii) if $\lambda \in \mathbb{R}$ and $f \in \mathbb{C}(x)$ choose $u \in X$ such that $| f(u) | = \| f \|$

Then for every $x \in X$, $| \lambda f(x) | = | \lambda || f(x) | \leq | \lambda | \| f \|$

That implies $| \lambda f(x) | \leq | \lambda || f(x) | \leq | \lambda || f |$

And so $| \lambda | \| f \| = sup\{| \lambda || f(x) | \mid x \in X\}$

Thus $\| \lambda f \| = | \lambda | \| f \|$

(iv) Let $f, g \in \mathbb{C}(x)$, then for all $x \in X$

$$| f(x) + g(x) | \leq | f(x) | + | g(x) | \leq \| f \| + \| g \|$$

Hence $\| f + g \| \leq \| f \| + \| g \|$. Thus $\| \, . \, \|$ is a norm on $\mathbb{C}(x)$

(b) Let f_n be a Cauchy sequence in $\mathbb{C}(x)$ for every positive real number ε.

Let q be a positive integer such that $\| f_m - f_n \| < \frac{\varepsilon}{2}$ for all $m, n \geq q$

Then $\| f_m(x) - f_n(x) \| \leq \| f_m - f_n \| < \frac{\varepsilon}{2}$ for all $x \in X$, and $m, n \geq q$

Hence by the Cauchy criterion, $f_n(x)$ converges uniformly on X.

Let $f(x) = \lim_{n \to \infty} f_n(x)$ for every $x \in X$. Hence f is a real-valued function on X and so $f \in \mathbb{C}(x)$

Given $\varepsilon > 0$, we get a positive integer q such that $\| f_m - f_n \| < \frac{\varepsilon}{2}$ for all $x \in X$ $, m, n \geq q$

and that implied $| f_m(x) - f_n(x) | < \frac{\varepsilon}{2}$ for all $x \in X$ $, m, n \geq q$

if $m \to \infty$ then

$| f(x) - f_n(x) | \leq \frac{\varepsilon}{2} < \varepsilon$ for all $x \in X$ $, m, n \geq q$

It follows that $\| f - f_n \| \leq \frac{\varepsilon}{2} < \varepsilon$ for all $n \geq q$

Hence $f_n \to f$ as $n \to \infty$ in the normed vector space $(\mathbb{C}(x), \| \, . \, \|)$

Thus $(\mathbb{C}(x), \| \, . \, \|)$ is a Banach space.

Theorem 3.11.7

A closed vector subspace of a Banach space is a Banach space itself.

Proof

Let $(V, \| \, . \, \|)$ be a Banach space and let G be a closed vector space of V. If x_n is a Cauchy sequence in G, then it is a Cauchy sequence in V and therefore there exists $x \in V$ such that $x_n \to x$. Since G is a closed subset of V, we have $x \in G$. Thus, every Cauchy sequence of G converges to an element of G.

3.12 Hilbert space

Given an inner product space $(V, \langle . \rangle)$, let $\| \, . \, \|$ be the norm induced on V by the inner product $\langle . \rangle$, then V is called a Hilbert space if and only if every Cauchy sequence in V is a convergent sequence in V.

Every Hilbert space is a Banach space.

Theorem 3.12.1

The space ℓ_p with $p \neq 2$ is not an inner product space, hence not a Hilbert space.

Proof

Let $x = (1, 1, 0, 0, \dots) \in \ell_p$ and $y = (1, -1, 0, 0, \dots) \in \ell_p$

$\| x \| = \| y \| = 2^{\frac{1}{p}}$ and $\| x + y \| = \| x - y \| = 2$

This does not satisfy the parallelogram equality. ℓ_p is complete. Hence ℓ_p with $p \neq 2$ is a Banach space which is not a Hilbert space.

3.13 Strong and Weak Convergence

Since every inner product space is a normed space, we have a natural notion of convergence defined by the norm. This convergence will be called strong convergence. We can also define a second kind of convergence involving the inner product. This convergence will be called weak convergence. We will see that strong convergence implies weak convergence but the converse does not need to be true in general.

3.13.1 Strong Convergence

A sequence x_n in a normed space X is said to be strongly convergent(or convergent in the norm) if there is an $x \in X$ such that

$\lim_{n\to\infty} \| x_n - x \| = 0$

This is written as $\lim_{n\to\infty} x_n = x$ or simply, $x_n \to x$. x is called the strong convergent limit of x_n and we say that x_n converges strongly to x.

3.13.2 Weak Convergence

A sequence x_n in an inner product space E is called weakly convergent to a vector x in E if

$\langle x_n, y\rangle \to \langle x, y\rangle$ as $n \to \infty$ for every $y \in E$.

The condition above can be stated as ; $\langle x_n - x, y\rangle \to 0$ as $n \to \infty$ for every $y \in E$.

It will be convenient to reserve $''x_n \to x''$ for the strong convergence and use $''x_n \rightharpoonup x''$ to denote weak convergence.

Chapter 4

Comparison

In this chapter we compare strong and weak convergence as well as pointwise and uniform convergence of sequence of functions treated in the previous section.

4.1 Theorem

A strongly convergent sequence is weakly convergent (to the same limit) i.e. $x_n \to x$ implies $x_n \rightharpoonup x$

Proof

Suppose the sequence x_n converges strongly to x. This means $\lim\limits_{n\to\infty} \| x_n - x \| = 0$
By the Schwarz's inequality we have
$| \langle x_n - x, y \rangle | \leq \| x_n - x \| \| y \| \to 0$ as $n \to \infty$
And thus $\langle x_n - x, y \rangle \to 0$ as $n \to \infty$ for every $y \in E$.

The converse of the above theorem is not true. This is shown in the example below.

Example 4.1.1 *Suppose the sequence x_n is an orthonormal sequence in the a Hilbert space H.*
Then by the Bessel's inequality $\sum_{n=1}^{\infty} | \langle x_n, y \rangle |^2 \leq \| y \|^2 \quad \forall y \in H$
Hence $\langle x_n, y \rangle \to 0 \quad$ *as* $\quad n \to \infty \quad$ *also,* $\langle 0, y \rangle = 0 \quad \forall y \in H$
Hence $\langle x_n, y \rangle \to \langle 0, y \rangle \quad$ *as* $\quad n \to \infty \quad \forall y \in H$

$\therefore x_n$ *is weakly convergent to zero in H.*

But x_n is not strongly convergent since x_n is not Cauchy.

Notation 4.1.1 *If X is finite dimensional vector space, then strong convergence is equivalent to weak convergence.*

Proof

$X = F^n$ under the Euclidean norm $\| \, . \, \|_2$

Suppose $x_n \to x$ weakly in F^n

Then for each standard basis vector e_k we have

$x_n \cdot e_k \to x \cdot e_k \quad$ *for $k = 1, 2, ..., n$*

This implies convergence in norm or strong convergence since

$\| x_n - x \|^2 = \sum_{k=1}^{n} | x_n \cdot e_k - x \cdot e_k |^2 \to 0 \quad$ *as $n \to \infty$*

Lemma 4.1.1

Let x_n be a weakly convergent sequence in a normed space X,

say,$x_n \rightharpoonup x$. Then

(a) The limit x of x_n is unique

(b) Every subsequence of x_n converges weakly to x.

Proof

(a) Suppose $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$. Then;

$f(x_n) \rightharpoonup f(x)$ as well as $f(x_n) \rightharpoonup f(y)$

Since $(f(x_n))$ is a sequence of numbers, its limit unique. Hence $f(x) = f(y)$

$f(x) - f(y) = f(x-y) = 0$ which implies that $x - y = 0$ for all $f \in X'$

(b) It follows from the fact that $(f(x_n))$ is a convergent sequence of numbers, so that every subsequence of $(f(x_n))$ converges and so has the same limit as the sequence.

4.1.1 Pointwise convergence and Uniform Convergence

Suppose f_n is a sequence of functions defined on a set X such that $f : s \to \mathbb{R}$. The condition $f_n \to f$ pointwise on S says that for every point $x \in S$ we have $f_n(x) \to f(x)$

as $n \to \infty$. This may be stated in the following form;

For every point $x \in S$ and every number $\varepsilon > 0$, there exists a number N such that whenever $n > N$ we have $| f_n(x) - f(x) | < \varepsilon$

This is the same as;

For every number $\varepsilon > 0$ and every point $x \in S$ there exists a number N such that whenever $n > N$ we have $| f_n(x) - f(x) | < \varepsilon$

On the other hand, in the definition of uniform convergence, no point x need be mentioned until after a value of N is specified; and therefore, in order to specify a value of N, we need only to know the value of ε . The conditions of point-wise convergence and uniform convergence are therefore quite different. Although every sequence f_n that converges uniformly to a function f will certainly converge pointwise to f.

The pointwise limit is the same as the uniform limit, so when we are asked to show uniform convergence we have to first show that it converges pointwise and then go on to show uniform convergence.

The only difference between them is that in the definition of pointwise convergence we are concerned only with one value of x at a time, the N we choose is thus allowed to depend not only on ε but also on the point x itself. In definition of uniform convergence, there must exist a single N which makes

$| f_n(x) - f(x) | < \varepsilon \quad for \quad all \quad x \in [a,b]$.

Thus uniform convergence is a stronger condition than pointwise convergence. The converse of this is not generally true.

Example 4.1.2 *The sequence f_n of functions where $f_n(x) = \frac{nx}{nx-n+x-1}$, is uniformly convergent in $[2,\infty[$ and converges pointwise as well.*

Proof

First, lets note the difference of two squares in the denominator, and so we can write our function as

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

First job is to show that this has a pointwise limit(i.e. take the limit as n goes to infinity).So,

$$\begin{aligned}
\lim_{n\to\infty} f_n(x) &= \lim_{n\to\infty} \frac{nx}{nx-n+x-1} \\
&= \frac{x}{x-1} \lim_{n\to\infty} \frac{n}{n+1} \\
&= \frac{x}{x+1}
\end{aligned}$$

Now for uniform convergence we need to find N in terms of ε. So we assume that $n \geq N$ (whatever N is) and manipulate the right side of the defintion above for uniform convergence until it is just in terms of N.

$$\begin{aligned}
\left| f_n(x) - \frac{x}{x-1} \right| &= \left| \frac{n}{n+1}\frac{x}{x-1} - \frac{x}{x-1} \right| \\
&= \left| \frac{x}{x-1}\left(\frac{n}{n+1} - 1\right) \right| \\
&= \left| \frac{x}{x-1} \right| \left| \frac{-1}{n+1} \right| \\
&= \frac{x}{x-1}\frac{1}{n+1} \\
&< \frac{x}{x-1}\frac{1}{n} \\
&< \frac{x}{x-1}\frac{1}{N} \\
&\leq \frac{2}{N} \\
&< \varepsilon
\end{aligned}$$

So we need to choose N such that $\frac{2}{N} < \varepsilon$. So choose $N > \frac{2}{\varepsilon}$, which all makes sense as $\varepsilon \neq 0$. And so we have shown uniform convergence.

The example below also indicates that every point-wise convergent sequence in not uniformly convergent.

Example 4.1.3 *It should, however, be remembered that every pointwise convergent sequence is not uniformly convergent as will proved in the this example where*
$f_n(x) = \frac{nx}{1+n^2x^2} \quad x \in \mathbb{R}$.

Proof

Here $\lim_{n\to\infty} f_n(x) = \frac{nx}{1+n^2x^2} = 0 \quad for \quad all \quad x \in \mathbb{R}$ showing that the sequence f_n, is point-wise convergent with point-wise limit f such that $f(x) = 0 \quad forall x \in \mathbb{R}$.

we shall now show that the convergence is *not uniform* in any interval $[a,b]$ with ,0, as an interior point.

Suppose that f_n is uniformly convergent in $[a,b]$, so that the point-wise limit f is also the uniform limit.

Let $\varepsilon > 0$ be given. Then there exists N such that for all $x \in [a,b]$ and for all $n \geq N$
$| \frac{nx}{1+n^2x^2} - 0 | < \varepsilon$.

We take $\varepsilon = \frac{1}{4}$. Now there exists an integer k such that $k \geq N$ and $\frac{1}{k} \in [a,b]$.

Taking $n = k$ and $x = \frac{1}{k}$, we have

$\frac{nx}{1+n^2x^2} = \frac{1}{2}$ which is not less than $\frac{1}{4}$.

Thus we arrive at a contradiction and such see that the sequence in *not uniformly convergent* in any interval $[a,b]$ with, 0, as an interior point even though it is point-wise convergent there.

Theorem 4.1.2

Let $(V, \langle . \rangle)$ be an inner product space. If $x_n \rightharpoonup x$ as $n \to \infty$ and $\| x_n \| \to \| x \|$ as $n \to \infty$ then x_n converges strongly to x. .

Proof

We have to show that $\| x_n - x \| \to 0$ as $n \to \infty$. Observe that

$$\begin{aligned}
\| x_n - x \|^2 &= \langle x_n - x, x_n - x \rangle \\
&= \| x \|^2 - \langle x_n, x \rangle - \langle x, x_n \rangle + \| x \|^2 \\
&= \| x \|^2 - 2\Re\langle x_n, x \rangle + \| x \|^2, n \to \infty \\
&= \| x \|^2 - 2\langle x, x \rangle + \| x \|^2 \\
&= \| x \|^2 - 2 \| x \|^2 + \| x \|^2 \\
&= 0
\end{aligned}$$

Chapter 5

Conclusion and Recommendation

5.1 Conclusion

We have been able to show in this work that:

1. Strong convergence imply weak convergence but the reverse is not generally true.

2. Uniform convergence imply pointwise but the reverse is not true.

5.2 Recommendation

For further developments we recommend that the following are treated ;

1. Convergence in distribution.

2. Convergence probability.

3. Almost sure convergence.

References

Ablowitz, M. J., & Fokas, A. S. (2003). *Complex variables: Introduction and applications* (P. D. G. Crighton, Ed.). Cambridge University Press.

Bartle, R. G. (1964). *The elements of real analysis* (F. Edition, Ed.). John Wiley and Sons, Inc.

Beals, R. (2004). *Analysis an introduction* (F. Edition, Ed.). Cambridge University Press.

Birkhorff, G., & Lane, S. M. (1977). *A survey of modern algebra* (Fourth, Ed.). Macmillan Publishing Company, Inc.

Blitzer, R. (2004). *College algebra* (E. Frank, Ed.). Prentice-Hall, Inc,.

Curtin, R. F., & Prutchard, A. J. (1977). *Functional analysis in modern applied mathematics* (R. Bellman, Ed.). Academic Press.

Desmond J. Higham, X. M., & Stuart, A. M. (2002). Strong convergence of euler-type methods for nonlinear stochastic differential equations. *Society of Industrial and Applied Mathematics.*, *40*.

Dugundji, J. (1966). *Topology* (I. Kaplansky, Ed.). Allyn and Bacon, Inc.

Folland, G. B. (1984). *Real analysis : Modern techniques and their applications*. Wiley, New York.

Frank Ayres, J., & Jaisingh, L. R. (1965). *Theory and problems of abstract algebra* (Second, Ed.). McGraw- Hill Companies, Inc.

Gonzalez-Velasco, E. A. (1992). Connections in mathematical analysis: the case of fourier series. *American Mathematical Monthly*, *90*, 427-441.

Gunawan, H. (2002). On convergence in n-inner product spaces. *Bulletin of the*

Malaysian Mathematical Sciences Society, *(second series)*, 11-16.

Hansen, B. E. (2008). uniform convergence rates for kernel estimation with dependent data. In *uniform convergence rates for kernel estimation with dependent data.*

Hardy, G. (1921). *A course of pure mathematics* (C. Clay, Ed.). Cambridge University Press.

Hassel, R. van. (2009). *Own lecture notes functional analysis*.

Hazewinkel, M. (2006). *Handbook of algebra* (M. Artin, Ed.). Elsevier B.V.

Herstein, I. N. (1996). *Abstract algebra* (J. Grant, Ed.). Prentice-Hall, Inc,.

Hoffman, K., & Kunze, R. (1971). *Linear algebra* (S. Edition, Ed.). Prentice-Hall, Inc,.

Knapp, A. W. (2005). *Basic real analysis* (C. L. Epstein, Ed.). Birkhauser Boston.

Knapp, A. W. (2007). *Advanced algebra* (C. L. Epstein, Ed.). Birkhauser Boston.

Kreyszig, E. (1978). *Introductory functional analysis with applications* (W. C. Library, Ed.). John Wiley and Sons, Inc.

Kristensen, D. (2008). Uniform convergence rates of kernel estimators with heterogeneous, dependent data. *Center for Research In Econometric Analysis of Time Series.*.

Kuelbs, J. (1976). A strong convergence theorem for banach space valued random variables. *The Annals of Probability*, *4*.

Lang, S. (1987). *Undergraduate algebra* (S. Axler, Ed.). Springer Science + Business Media, Inc.

Latakus, I. (1976). *Proofs and refutations*. Cambridge University Press, New York.

Lindgren, F. (2012). *On weak and strong convergence of numerical approximations of stochastic partial differential equations*. Unpublished doctoral dissertation, Chalmers University of Technology and University of Gothengburg.

Morash, R. P. (1987). *Bridge to abstract mathematics: Mathematical proofs and structures* (F. Edition, Ed.). Random House, Inc.

Narayan, S. (2000). *Elements of real analysis* (M. Raisinghania, Ed.). S. Chad & Company Ltd.

Newey, W. K. (1989). Uniform convergence in probability and stochastic equicontinuity. *Econometric Research Program*.

N.L.Carothers. (2000). *Real analysis*. Cambridge University Press.

Pederson, G. K. (1989). *Analysis now* (J. H. Ewing, Ed.). Springer-Verlag Inc.

Rosen, K. H. (2007). *Handbook of linear algebra* (L. Hogben, Ed.). Chapman & Hall/CRC.

Rubin, H. (1952). uniform convergence of random functions with applications to statistics. *The Annals of Statistics*.

Rudin, W. (1921). *Functional analysis* (J. L. Farnsworth, Ed.). Mc-Graw Hill Book Company.

Shilov, G. E. (1973). *Elementary real and complex analysis* (R. A. Silverman, Ed.). Dover Publications, Inc.

Takahashi, W., & Yao, J.-C. (2011). weak and strong convergence theorems for positively homogeneous nonexpansive mappings in banach spaces. *Taiwanese Journal of Mathematics*, *3*.

Wheeden, R. L., & Zygmund, A. (1977). *Measure and integral* (S. Kobayashi, Ed.). Marcel Dekker, Inc New York and Basel.

Wickless, W. J. (2004). *A graduate course in abstract algebra* (E. J. Taft, Ed.). Marcel Dekker, Inc New York and Basel.

Yang, Z. (2009). Pointwise convergence and ascoli theorems for nearness spaces. *Applied General Topology*, *10*.

Yuli Eidelman, V. M., & Tsolomitis, A. (2004). *Functional analysis - an introduction* (W. Craig, Ed.). American Mathematical Society.

Appendix A

pointwise convergence of fourier series

Theorem A.0.1 *(Féjer) Let f be integrable and periodic. Then*

$$\sigma_n(x) \to f(x)$$

at each point of continuity of f, and the convergence is uniform over every closed interval of continuity. In particular, $\sigma_n(x) \to f(x)$ uniformly everywhere if f is continuous everywhere.

Notation A.0.1 *(Trigonometric System)*
$e^{ikx} = \cos kx + \sin kx \quad (k = 0, \pm1, \pm2, \dots)$ *is called a trigonometric system. These functions are periodic, with period 2π. They form an orthogonal system over any interval $Q = (a, a+2\pi)$ of length 2π.*
$\Rightarrow \int_Q e^{ikx}\overline{e^{ikx}}dx = \left[\frac{e^{i(k-m)x}}{k-m}\right]_a^{a+2\pi} = 0$
Functions of the form $e^{ikx} = \cos kx + \sin kx \quad (k = 0, \pm1, \pm2, \dots)$ has Fourier series given by
$\sum_{-\infty}^{\infty} C_k e^{ikx}$ *where* $C_k = \frac{1}{2\pi}\int_Q f(t)e^{-ikt}dt$
The definition of a general orthogonal system presupposes that the functions in the system are of class L_2. This makes it possible to define Fourier coefficients for any $f \in L_2$. If f is not in L_2, it may be impossible to define its Fourier coefficients with respect to certain orthogonal systems.
In the case of the functions e^{ikx}, which are bounded, the coefficients C_k are defined for

any f which integrable over Q in particular for any $f \in L_p(Q), \quad 1 \leq p < \infty$
Thus, the trigonometric system is richer in properties than general orthogonal systems.

Theorem A.0.2 *The trigonometric system is complete. More precisely, if all the Fourier coefficients of an integrable f are zero, then $f = 0 \quad a.e$*