

Class Notes

Math 556 - Applied Functional Analysis
Professor Liliana Borcea
University of Michigan



Jinyan Miao
and his friends from UMich Honors Math

Fall 2022

Contents

	Page
1 Measure Theory	3
1.1 σ -Algebra	3
1.2 Measure Spaces	5
1.3 Lebesgue Measure	6
1.4 Measure Zero	9
1.5 Lebesgue Measurable Functions	10
1.6 Lebesgue Integrals	12
1.7 Convergence Theorems	15
2 The Algebra and Topology	19
2.8 Vector Spaces	19
2.9 Metric Spaces	21
2.10 Completeness	24
2.11 Compactness	30
3 Banach Spaces	33
3.12 Normed Vector Spaces	33
3.13 The Little l^p Spaces	38
3.14 The Lebesgue L^p Spaces	42
4 Continuous Functions	45
4.15 Uniform Continuity	45
4.16 Equicontinuity	47
4.17 Lipchitz Functions	49
5 Bounded Linear Operators	53
5.18 The Bounded Operators	53
5.19 Uniform Boundedness and Convergence	57
5.19.1 Application in Numerical Quadrature	59
5.20 Approximation Schemes	62
5.21 Operator Power Series	63
6 Hilbert Spaces	65
6.22 Inner Product Spaces	65
6.23 Orthogonality	68
6.23.1 Linear Least Squares	72
6.24 Orthonormal Sets	73
6.24.1 Unordered Sums	73
6.25 Orthonormal Bases	78

6.25.1	The Gram-Schmidt Orthogonalization	79
7	Bounded Linear Operators on Hilbert Spaces	81
7.26	Duality of Hilbert Space	81
7.27	Weak Convergence	83
7.28	Adjoint Operator	88
7.29	The Fredholm Theory	90
8	Spectrum of Operators	93
8.30	Operators on Finite Dimensional Spaces	93
8.31	Operators on Infinite Dimensional Spaces	95
8.32	Spectrum of Compact Operators	98
8.33	The Spectral Theorem	101
9	Linear Differential Operators	105
9.34	Adjoints of Unbounded Linear Operators	105
9.34.1	Self-adjoint Unbounded Operators	106
9.34.2	Adjoint of Linear Differential Operator	107
9.34.3	Green's Function	109

Courses Instructor	Liliana Borcea
Notes Transcriber	Jinyan Miao
Content Editor	Jinyan Miao
Art Designers	Wenyu Chen Jinyan Miao

This text is prepared using the \TeX typesetting language.

Materials credit to the Department of Mathematics at the University of Michigan.
This work is licensed under a Creative Commons By-NC-ND 4.0 International License.



This text is intended as a supplement class notes for students who is taking, or has taken, Math 556 at the University of Michigan. We assume that the reader has taken the Honors Math course sequence, Math 295-396, or equivalent sequence of rigorous mathematical analysis courses at the University of Michigan. We will make use of some usual notation that were introduced in those courses. The Honors Math sequence covers topics related to sets, groups, topology, series and sequences, differentiation and integration of functions of single variable, linear algebra, differentiation and integration of functions of multivariables, differentiable manifolds and exterior calculus, Riemannian metrics, complex analysis, and introduction to measure theory.

The class number for 2020 Math 295/296 is 57, but in this text we will use 19 as our prime number, notice that we have $19 \cdot 3 = 57$.

This text is edited by Jinyan Miao. The course 2022 Math 556 is taught by Professor Liliana Borcea at the University of Michigan - Ann Arbor. Except as permitted by both Jinyan and Professor Borcea, no part of this text is allowed to be distributed. This text contains information obtained from authentic sources, but the editors cannot assume responsibility for the validity of all materials in this text or the consequences of their use. The editors have attempted to trace the copyright holders of all material reproduced in this text and apologize to copyright holders if permission to share in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know. If you have any questions or concerns regarding this text, or if you find any typos in this text, please contact Jinyan through jmiu@umich.edu.

1 | Measure Theory

For a given set X , say $X \subseteq \mathbb{R}^n$, and $A \subseteq X$, we want to associate measure of A such that $m(A) \geq 0$ gives the generalization of volume of A , but perhaps not all sets are measurable. To have useful framework, we need measurable sets be closed under countable unions and intersections.

σ -Algebra

Definition 1.0.0.0.1

Let X be a set, a family \mathcal{F} of subsets of X is a σ -algebra on X provided that we have:

1. $\emptyset \in \mathcal{F}$, and $X \in \mathcal{F}$.
2. If $A \in \mathcal{F}$, then $X - A \in \mathcal{F}$.
3. If $A_k \in \mathcal{F}$ for all $k \in \mathbb{N}$, then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.

Lemma 1.0.1

Let X be a set, let \mathcal{F} be a σ -algebra on X . If $A_k \in \mathcal{F}$ for all $k \in \mathbb{N}$, then $\bigcap_{k=1}^{\infty} A_k \in \mathcal{F}$

Proof. Note that for all $k \in \mathbb{N}$, we have $A_k^c = X - A_k \in \mathcal{F}$, hence we have:

$$\bigcup_{k=1}^{\infty} A_k^c \in \mathcal{F}$$

By De Morgan Law, we can write:

$$X - \left(\bigcap_{k=1}^{\infty} A_k \right) = \bigcup_{k=1}^{\infty} A_k^c \in \mathcal{F} \quad \Rightarrow \quad \bigcap_{k=1}^{\infty} A_k \in \mathcal{F}$$

This completes the proof. □

Example: Let X be a set, the smallest σ -algebra on X is $\mathcal{F} = \{\emptyset, X\}$.¹

Example: Let X be a set, the largest σ -algebra on X is $\mathcal{F} = \mathcal{P}(X)$.²

Example: Let $X = \{e_1, e_2, \dots, e_6\}$ be the set of possible outcomes in die rolling, where e_j is the event where outcome is j . Then $\mathcal{F} = \mathcal{P}(X)$ has 2^6 elements.

Definition 1.0.1.0.1

Let X be a set, let $\mathcal{F} \subseteq \mathcal{P}(X)$ be a family of subsets of X , a σ -algebra generated by \mathcal{F} is the smallest σ -algebra that contains \mathcal{F} .

¹The term smallest σ -algebra on X means that such σ -algebra is contained in all σ -algebra on X .

²The term largest σ -algebra on X means that such σ -algebra contains all other σ -algebra on X .

Example: Let $X = \mathbb{R}^n$, let \mathcal{F} be the family of the open sets in \mathbb{R}^n , the σ -algebra generated by \mathcal{F} is the Borel σ -algebra of \mathbb{R}^n , denoted as $\mathcal{B}(X)$. $\mathcal{B}(X)$ is closed under complements so it contains all closed sets in \mathbb{R}^n , and the family of all closed sets in X also generate $\mathcal{B}(X)$.

Exercise: $\mathcal{B}(\mathbb{R})$ is generated by the collection of intervals $\{(a, b] \mid a < b\}$.

Measure Spaces

Definition 2.0.0.0.1

Let X be a set, let \mathcal{F} be a σ -algebra on X . A measure on \mathcal{F} is a function $\mu : \mathcal{F} \rightarrow [0, \infty]$ that satisfies the followings:

1. There exists $A \in \mathcal{F}$ such that $\mu(A) < \infty$
2. μ is countable additive. That is, if $\{E_j \mid j \in \mathbb{N}\}$ is a set of disjoint sets, then:

$$\mu \left(\bigcup_{j=1}^{\infty} E_j \right) = \sum_{j=1}^{\infty} \mu(E_j) \quad (1.1)$$

Note: LHS of equation (1.1) might be finite while the RHS of that is an infinite sum.

Definition 2.0.0.0.2

Let X be a set, let \mathcal{F} be a σ -algebra on X , and let $\mu : \mathcal{F} \rightarrow [0, \infty]$ be a measure on \mathcal{F} , (X, \mathcal{F}, μ) is called a measure space. All sets $A \in \mathcal{F}$ are said to be measurable.

Definition 2.0.0.0.3

Let X be a set, and let \mathcal{F} be a σ -algebra on X . A measure on \mathcal{F} is said to be finite provided that $\mu(X) < \infty$. A measure on \mathcal{F} is said to be σ -finite provided that there exists a countable family $\{A_j \in \mathcal{F} \mid j \in \mathbb{N}\}$ of measurable sets such that $\mu(A_i) < \infty$ and we have $X = \bigcup_{i=1}^{\infty} A_i$.

Theorem 2.1

Let (X, \mathcal{F}, μ) be a measure space, the followings hold:

1. $\mu(\emptyset) = 0$.
2. For $E, F \in \mathcal{F}$ with $E \subseteq F$, we have $\mu(E) \leq \mu(F)$.
3. Let J be an index set, let $E_j \in \mathcal{F}$ for $j \in J$, we have:

$$\mu \left(\bigcup_{j \in J} E_j \right) \leq \sum_{j \in J} \mu(E_j)$$

4. $\forall j \in \mathbb{N}$, let $E_j \in \mathcal{F}$ that satisfies $E_j \subseteq E_{j+1}$, then we have:

$$\mu \left(\bigcup_{j=1}^{\infty} E_j \right) = \lim_{j \rightarrow \infty} \mu(E_j)$$

5. $\forall j \in \mathbb{N}$, let $E_j \in \mathcal{F}$ that satisfies $E_j \supseteq E_{j+1}$, if $\exists m \in \mathbb{N}$ s.t. $\mu(E_m) < \infty$, then:

$$\mu \left(\bigcap_{j=1}^{\infty} E_j \right) = \lim_{j \rightarrow \infty} \mu(E_j)$$

Lebesgue Measure

The Borel σ -algebra on \mathbb{R}^n , denoted as $\mathcal{B}(\mathbb{R}^n)$ can be generated by the family of cubes C defined by:

$$C = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$$

where $a_i, b_i \in \mathbb{R}$ with $a_i < b_i$ for all $1 \leq i \leq n$. The Lebesgue measure on $\mathcal{B}(\mathbb{R}^n)$, denoted as $\lambda : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, \infty]$, is defined such that we have:

$$\lambda(C) = \text{volume}(C) = V(C) = \prod_{i=1}^n (b_i - a_i)$$

Note that the Lebesgue measure is a σ -finite measure as we have:

$$\mathbb{R}^n = \bigcup_{i=1}^{\infty} (-i, i)^n$$

At this time, the existence of the Lebesgue measure is yet to be explained. In the following discussion, we will see how we can construct the Lebesgue measure.

Construction of the Lebesgue measure

First we consider partitioning \mathbb{R}^n into cubes of side length 2^{-k} for $k \in \mathbb{N}$, such partition is given as the following set of cubes:

$$C_k := \left\{ \left[\frac{\gamma_1}{2^k}, \frac{\gamma_1 + 1}{2^k} \right) \times \left[\frac{\gamma_2}{2^k}, \frac{\gamma_2 + 1}{2^k} \right) \times \cdots \times \left[\frac{\gamma_n}{2^k}, \frac{\gamma_n + 1}{2^k} \right) \mid \gamma_1, \gamma_2, \dots, \gamma_n \in \mathbb{Z} \right\}$$

We define the radius of such partition as:

$$\delta_k := 2^{-k} \sqrt{n}$$

Let $G \subseteq \mathbb{R}^n$ be an open set. Given $k \in \mathbb{N}$, we define the partition of G as:

$$S_k(G) := \{C \in C_k \mid \bar{C} \subseteq G\}$$

And we define the following set:

$$\mathcal{S}_k(G) := \bigcup_{C \in S_k(G)} C$$

Lemma 3.0.1

Let G be an open subset of \mathbb{R}^n , then we have $\mathcal{S}_k(G) \subseteq \mathcal{S}_{k+1}(G)$ for all $k \in \mathbb{N}$, and that:

$$G = \bigcup_{k=1}^{\infty} \mathcal{S}_k(G)$$

For such open set G , now we can define:

$$m_k(G) := V(\mathcal{S}_k(G)) = \frac{\#(S_k(G))}{2^{kn}}$$

Here it is easy to see that $(m_k(G))$ is a monotone increasing sequence, and hence it has a limit, so we can define: ³

$$m(G) := \lim_{k \rightarrow \infty} m_k(G)$$

m defined in this way behaves like a measure as it satisfies property (4) in Theorem 2.1:

$$\mathcal{S}_k(G) \subseteq \mathcal{S}_{k+1}(G) \quad \Rightarrow \quad m\left(G = \bigcup_{k=1}^{\infty} \mathcal{S}_k(G)\right) = \lim_{k \rightarrow \infty} m_k(G)$$

Theorem 3.1

Let J be finite or infinite index set, for $j \in J$, let G_j be open subsets of \mathbb{R}^n , we have:

1. If $G_1 \subseteq G_2$, then $m(G_1) \leq m(G_2)$.
2. If $G_1 \cap G_2 = \emptyset$, then $m(G_1 \cup G_2) = m(G_1) + m(G_2)$.
3. $m(G_1 \times G_2) = m(G_1) \cdot m(G_2)$.
4. $m(\bigcup_{j \in J} G_j) \leq \sum_{j \in J} m(G_j)$

Theorem 3.1 also ensures that m behaves like a measure, but m is yet a measure because we have only define it on the set of open sets, which is not a σ -algebra.

Theorem 3.2

Let G be an open subset of \mathbb{R}^n , then we have:

$$m(G) = \sup\{m(H) \mid H \text{ is open, } \bar{H} \subseteq G\}$$

Theorem 3.2 implies that we can get $m(G)$ by approximating G from inside with closed sets.

Definition 3.2.0.0.1

Let E be an arbitrary subset of \mathbb{R}^n ,

$$m^*(E) := \inf \{m(G) \mid G \text{ is an open superset of } E\}$$

In Definition 3.2.0.0.1, we take E and enclose it by open sets for which we have m .

Theorem 3.3

Let J be an indexing set, and for $j \in J$, let E_j be a subset of \mathbb{R}^n , then we have the followings hold:

1. $m^*(E_1) = m(E_1)$ if E_1 is open.
2. $m^*(E_1) \leq m^*(E_2)$ if $E_1 \subseteq E_2$
3. $m^*(\bigcup_{j \in J} E_j) \leq \sum_{j \in J} m^*(E_j)$

³Here we use the usual convention that the limit approaches ∞ when $(m_k(G))$ is not bounded above.

Theorem 3.4

The following families coincide:

1.

$$\mathcal{L}_1(\mathbb{R}^n) := \left\{ E \subseteq \mathbb{R}^n \left| \inf_{\substack{E \subseteq G \subseteq \mathbb{R}^n, \\ G \text{ is open in } \mathbb{R}^n}} m^*(G - E) = 0 \right. \right\}$$

2.

$$\mathcal{L}_2(\mathbb{R}^n) := \left\{ E \subseteq \mathbb{R}^n \left| \inf_{\substack{F \subseteq E \subseteq G \subseteq \mathbb{R}^n, \\ G \text{ is open in } \mathbb{R}^n, \\ F \text{ is closed in } \mathbb{R}^n}} m^*(G - F) = 0 \right. \right\}$$

3.

$$\mathcal{L}_3(\mathbb{R}^n) := \left\{ E \subseteq \mathbb{R}^n \left| \inf_{\substack{F \subseteq E \subseteq \mathbb{R}^n, \\ F \text{ is closed in } \mathbb{R}^n}} m^*(E - F) = 0 \right. \right\}$$

Theorem 3.4 shows that the Lebesgue measure of a set E can be approximated by open sets from outside of E , and by closed sets from inside of E .

Definition 3.4.0.0.1

$$\mathcal{L}(\mathbb{R}^n) := \left\{ E \subseteq \mathbb{R}^n \left| \inf_{\substack{F \subseteq E \subseteq \mathbb{R}^n, \\ F \text{ is closed in } \mathbb{R}^n}} m^*(E - F) = 0 \right. \right\}$$

is called the family of Lebesgue measurable sets in \mathbb{R}^n .

Theorem 3.5

In \mathbb{R}^n , $\mathcal{L}(\mathbb{R}^n)$ is a σ -algebra that contains the Borel σ -algebra, that is $\mathcal{B}(\mathbb{R}^n) \subseteq \mathcal{L}(\mathbb{R}^n)$.

Theorem 3.6

In \mathbb{R}^n , $\lambda := m^*|_{\mathcal{L}(\mathbb{R}^n)}$ defines a measure on $\mathcal{L}(\mathbb{R}^n)$.

Definition 3.6.0.0.1

In \mathbb{R}^n , $\lambda := m^*|_{\mathcal{L}(\mathbb{R}^n)}$ is called the Lebesgue measure on $\mathcal{L}(\mathbb{R}^n)$.

Measure Zero

In this section, the measure is referred to the Lebesgue measure and is denoted as λ unless mentioned otherwise, and \mathbb{R}^n is considered to be equipped with the σ -algebra $\mathcal{L}(\mathbb{R}^n)$ and the Lebesgue measure to form a measure space.

Definition 4.0.0.0.1

In the metric space $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n), \lambda)$, a set $K \in \mathcal{L}(\mathbb{R}^n)$ is said to have measure zero provided that we have $\lambda(K) = 0$.

Theorem 4.1

Let $E \subseteq \mathbb{R}^n$. If $m^*(E) = 0$, then $E \in \mathcal{L}(\mathbb{R}^n)$.

Proof. Here we have:

$$0 = m^*(E) = \inf\{m(G) \mid G \text{ is open, } E \subseteq G\}$$

Take arbitrary $\epsilon > 0$ and by the Characterization of Infimum, there exists open $G \subseteq \mathbb{R}^n$ such that $E \subseteq G$ and $0 \leq m(G) \leq \epsilon$. Note that $G - E \subseteq G$, then by Theorem 3.3, we have $m^*(G - E) \leq m(G) \leq \epsilon$. Since ϵ is arbitrary, we see that:

$$\inf_{\substack{E \subseteq G \subseteq \mathbb{R}^n, \\ G \text{ is open}}} \{m^*(G - E)\} = 0$$

then $E \in \mathcal{L}(\mathbb{R}^n)$ by Theorem 3.4. □

Example: For all $x \in \mathbb{R}$, $\{x\} \in \mathcal{L}(\mathbb{R})$ and $\lambda(\{x\}) = 0$. Since we have $\mathcal{E} = (x - \epsilon, x + \epsilon)$ being open sets containing $\{x\}$, and $m(\mathcal{E}) \rightarrow 0$ as $\epsilon \rightarrow 0$.

Example: Every countable set $\{x_j \mid j \in \mathbb{N}\}$ has measure 0 by countable additivity of measure:

$$\lambda(\{x_j \mid j \in \mathbb{N}\}) = \sum_{j=1}^{\infty} m(\{x_j\}) = 0$$

Example: Let $E \subseteq \mathbb{R}^n$ and $F \subseteq \mathbb{R}^m$ with $\lambda(E) = 0$, then $E \times F \in \mathcal{L}(\mathbb{R}^{n+m})$ and has measure zero. Notice that F does not necessary belong to $\mathcal{L}(\mathbb{R}^m)$.

Definition 4.1.0.0.1

Given a measure space (X, \mathcal{F}, μ) , a property $p(x)$ for $x \in X$ is said to hold almost everywhere on X provided that it fails only on set $S \subseteq X$ with $\mu(S) = 0$.

Definition 4.1.0.0.2

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the essential supremum of f is defined by:

$$\text{ess sup } (f) := \inf\{\alpha \in \mathbb{R} \mid f(x) < \alpha \text{ almost everywhere on } \mathbb{R}^n\}$$

Lebesgue Measurable Functions

From now on, \mathbb{R}^n is considered to be equipped with the σ -algebra $\mathcal{L}(\mathbb{R}^n)$ and the Lebesgue measure λ to form a measure space if not mentioned otherwise.

For notation, we denote the extended reals as $\bar{\mathbb{R}} := [-\infty, \infty]$. We make the following definitions of algebraic operations involving $x \in \mathbb{R}$ and ∞ :

$$\begin{aligned} x + \infty &= \infty & x - \infty &= -\infty \\ x \cdot \infty &= \infty & x \cdot (-\infty) &= -\infty & (\text{if } x > 0) \\ x \cdot \infty &= -\infty & x \cdot (-\infty) &= \infty & (\text{if } x < 0) \\ |\infty| &= |-\infty| = \infty \\ 0 \cdot \infty &= 0 \cdot (-\infty) = 0 \end{aligned}$$

When concerned, the usual topology equipped on the extended reals is the collection of sets (a, b) , $[-\infty, a)$, $(a, \infty]$ and any union of these types. The order equipped on the extended reals satisfies $-\infty < a < \infty$ for all $a \in \mathbb{R}$. Note that any monotone sequence (x_n) of points in $\bar{\mathbb{R}}$ has a limit. We also, when concerned, equip $\bar{\mathbb{R}}$ with the Borel σ -algebra $\mathcal{B}(\bar{\mathbb{R}})$, which is a σ -algebra generated by the family of semi-infinite intervals, given by the following:

$$\mathcal{F} = \{[-\infty, c) \mid c \in \mathbb{R}\}$$

Definition 5.0.0.0.1

Let $X \subseteq \mathbb{R}^n$, a function $f : X \rightarrow \mathbb{R}$ is said to be Lebesgue measurable provided that the preimage of any open subset of \mathbb{R} is Lebesgue measurable. That is, f is said to be Lebesgue measurable provided that for all open subset U of \mathbb{R} , $f^{-1}(U) \in \mathcal{L}(\mathbb{R}^n)$.

More generally, one can consider the following definition:

Definition 5.0.0.0.2

Let X be a subset of \mathbb{R}^n , a function $f : X \rightarrow \mathbb{R}$ is Lebesgue measurable provided for all S in a family that generates $\mathcal{B}(\mathbb{R})$, we have $f^{-1}(S) \in \mathcal{L}(\mathbb{R}^n)$.

Definition 5.0.0.0.3

Let X be a subset of \mathbb{R}^n , a function $f : X \rightarrow \mathbb{R}$ is Lebesgue measurable provided for all S in a family that generates $\mathcal{B}(\bar{\mathbb{R}})$, we have $f^{-1}(S) \in \mathcal{L}(\mathbb{R}^n)$.

Definition 5.0.0.0.4

Let (X, \mathcal{F}, μ) and (Y, \mathcal{G}, ν) be measure spaces.

A measurable function is a function $f : X \rightarrow Y$ such that $f^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{G}$.

From Definition 5.0.0.0.4, we see that the measurability of a function f depends only on the σ -algebra on X and Y , and not on what measure. In the case of Lebesgue measurable function, the codomain of the function of interest must be equipped with the family of Lebesgue measurable sets as its σ -algebra.

Note: If a function $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is continuous, then it must be Lebesgue measurable, because by definition of continuity, if U is open in \mathbb{R} , then $f^{-1}(U)$ is an open set and thus belongs to $\mathcal{L}(\mathbb{R}^n)$.

Note: Let $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be a Lebesgue measurable function, we get a σ -algebra:

$$\{S \subseteq \bar{\mathbb{R}} \mid f^{-1}(S) \in \mathcal{L}(\mathbb{R}^n)\}$$

Theorem 5.1

Let J be an index set, for $j \in J$, let $\phi_j : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be a Lebesgue measurable function. The following functions are Lebesgue measurable:

1. $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}} \quad x \mapsto \sup_{j \in J} \phi_j(x)$
2. $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}} \quad x \mapsto \inf_{j \in J} \phi_j(x)$
3. $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}} \quad x \mapsto \limsup \phi_j(x)$
4. $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}} \quad x \mapsto \liminf \phi_j(x)$

Lebesgue Integrals

Definition 6.0.0.0.1

For a set $A \subseteq \mathbb{R}^n$, we define:

$$\chi_A : A \rightarrow \{0, 1\} \quad x \mapsto \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}$$

χ_A is called the characteristic function of A .

Definition 6.0.0.0.2

Let (X, \mathcal{F}, μ) be a measure space, a simple function defined on X is a function of the form given by the following:

$$\phi : X \rightarrow \mathbb{R} \quad x \mapsto \sum_{j=1}^N c_j \chi_{A_j}(x) \quad (1.2)$$

where A_j are disjoint sets in \mathcal{F} and $c_j \in \mathbb{R}$.

Definition 6.0.0.0.3

Let f be a simple function defined on a measure space (X, \mathcal{F}, μ) .

The integral of f is given by:

$$\int_X f d\mu = \sum_{a \in f(X)} a \cdot \mu(f^{-1}(a))$$

For the interest of this text, we focus on Lebesgue measure space, in which case we would use the following:

Definition 6.0.0.0.4

Let $X \in \mathcal{L}(\mathbb{R}^n)$, a simple function on X is of the form:

$$\phi : X \rightarrow \mathbb{R} \quad x \mapsto \sum_{j=1}^N c_j \chi_{A_j}(x)$$

where A_j are disjoint sets in $\mathcal{L}(\mathbb{R}^n)$ and $c_j \in \mathbb{R}$. The Lebesgue integral of ϕ is given by:

$$\int_X \phi d\lambda := \sum_{j=1}^N c_j \lambda(A_j)$$

Definition 6.0.0.0.5

Let $X \in \mathcal{L}(\mathbb{R}^n)$, let $f : X \rightarrow \bar{\mathbb{R}}$ be a non-negative Lebesgue measurable function, the Lebesgue integral of f is given by the following:

$$\int_X f dx := \sup \left\{ \int_X \phi d\lambda \mid \phi \text{ is simple and } \phi(x) \leq f(x) \text{ for all } x \in X \right\}$$

Theorem 6.1

Let $E \subseteq \mathbb{R}^n$ such that $f : E \rightarrow [0, \infty]$ is a measurable function. There exists a monotone increasing sequence of simple functions (f_j) that converges pointwise to f from below.

Proof. For $k \in \mathbb{N}$, we divide the codomain of f into intervals:

$$I_{k,2^k+1} := [2^k, \infty], \quad I_{k,j} := \left[\frac{j-1}{2^k}, \frac{j}{2^k} \right) \quad \text{for } j \in \{1, 2, \dots, 2^k\} \quad (1.3)$$

The family of intervals defined by equation (1.3) are members of the Borel σ -algebra $\mathcal{B}(\mathbb{R})$. Since f is Lebesgue measurable function, then we can write:

$$A_{k,j} := f^{-1}(I_{k,j}) = \{x \in X \mid f(x) \in I_{k,j}\} \in \mathcal{L}(\mathbb{R}^n)$$

Here we define a sequence of simple functions (f_k) :

$$f_k : E \rightarrow [0, \infty] \quad x \mapsto \sum_{j=1}^{2^k+1} \frac{j-1}{2^k} \chi_{A_{k,j}}(x)$$

Here by definitions, it is immediate that we have $f(x) \geq f_k(x)$ for all $k \in \mathbb{N}$. As the increase of k , we refine the intervals $I_{k,j}$ of the codomain of f , and we obtain a monotone sequence of simple function:

$$f_1(x) \leq f_2(x) \leq \dots \leq f_k(x) \leq f_{k+1}(x) \leq \dots \leq f(x) \quad \forall x \in E$$

Now we will show that (f_k) converges to f pointwise. Take $x \in E$ and let $\epsilon > 0$ be given. Let $k \in \mathbb{N}$ such that we have:

$$\begin{cases} \frac{1}{2^k} < \epsilon \text{ and } f(x) < 2^{k-1} & f(x) \neq \infty \\ k \text{ is sufficient large} & f(x) = \infty \end{cases}$$

and let $j \in \mathbb{N}$ such that $x \in A_{k,j} = f^{-1}(I_{k,j})$. Here we consider two cases: (1) $1 \leq j \leq 2^k$, and (2) $j = 2^k + 1$. For case (1), we have $f(x) \in I_{k,j} = [\frac{j-1}{2^k}, \frac{j}{2^k})$, and $f_k(x) = \frac{j-1}{2^k}$, which implies we have:

$$|f(x) - f_k(x)| < \frac{j-1}{2^k} - \frac{j}{2^k} = \frac{1}{2^k} < \epsilon$$

For case (2), we have $f(x) \in I_{k,j} = [2^k, \infty]$, and in particular, $f(x) = \infty$, so $f(x)$ is in such interval $[2^k, \infty]$, for all k . On the other hand, we also have $f_k(x) = 2^k$, where $(2^k) \rightarrow \infty$. Hence in both cases, we see that (f_k) converges to f pointwise from below. This completes the proof. \square

We have defined the Lebesgue integral for non-negative Lebesgue measurable functions. For other measurable functions, we consider the following definitions for Lebesgue measurable function $f : E \rightarrow \bar{\mathbb{R}}$:

$$\begin{aligned} f_+ : E &\rightarrow \bar{\mathbb{R}} & x &\mapsto \max\{0, f(x)\} \\ f_- : E &\rightarrow \bar{\mathbb{R}} & x &\mapsto \min\{0, f(x)\} \end{aligned}$$

We see that $f_+(x) \geq 0$ and $f_-(x) \leq 0$ for all $x \in E$.

Definition 6.1.0.0.1

Let $f : E \rightarrow \bar{\mathbb{R}}$ be a Lebesgue measurable function such that one of $\int_E f_+ d\lambda$ and $\int_E f_- d\lambda$ is finite, we define the Lebesgue integral of f over E to be the following: ⁴

$$\int_E f d\lambda = \int_E f_+ d\lambda - \int_E f_- d\lambda$$

Definition 6.1.0.0.2

A measurable function $f : E \rightarrow \bar{\mathbb{R}}$ is said to be summable, or integrable, provided that:

$$\int_E |f| d\lambda < \infty$$

Note: For measurable function $f : E \rightarrow \bar{\mathbb{R}}$, we have $\int_E |f| d\lambda = \int_E f_+ d\lambda + \int_E f_- d\lambda$.

Example: Consider bounded function $f : [a, b] \rightarrow \mathbb{R}$ that is continuous except at countably many points in $[a, b]$, then f is Riemann integrable, Darboux integrable, and Lebesgue integrable, and the three integrals coincide. ⁵

Example: $L^p(E)$, where E is a Lebesgue measurable set, consists of functions whose p^{th} power is summable, with $1 \leq p < \infty$:

$$\int_E |f|^p d\lambda < \infty$$

⁴If we have $\int_E f_+ d\lambda = \infty$ and $\int_E f_- d\lambda = \infty$, we cannot define $\int_E f d\lambda$ because $\infty - \infty$ is resolutely refused to be defined.

⁵For details about integrating a multivariate function, check out *Chapter 3 - Multivariate Integration* in the text *Class Note: Math 395/396 - Honors Analysis*, transcribed by Jinyan Miao in Winter 2022. The course Math 395/396 at the University of Michigan in Winter 2022 was taught by Professor David Barrett

Convergence Theorems

Theorem 7.1 (Monotone Convergence Theorem)

Let $f_j : E \rightarrow \mathbb{R}$ be Lebesgue measurable non-negative functions defined on the Lebesgue measure space $(E, \mathcal{L}(E), \lambda)$ such that (f_j) converges to a function $f : E \rightarrow \mathbb{R}$ pointwise monotonically from below. In other words, for all $x \in E$, $(f_j(x))$ is monotone increasing such that $(f_j(x)) \rightarrow f(x)$. Then f is measurable and we have the following holds:

$$\int_E f \, d\lambda = \int_E \lim_{j \rightarrow \infty} f_j \, d\lambda = \lim_{j \rightarrow \infty} \int_E f_j \, d\lambda \quad (1.4)$$

Proof. By Theorem 5.1, we know that f is measurable because for all $x \in E$, we have:

$$f(x) = \lim_{j \rightarrow \infty} f_j(x) = \limsup_{j \rightarrow \infty} f_j(x) = \liminf_{j \rightarrow \infty} f_j(x)$$

Since $f_j(x) \geq 0$ for all $x \in E$, and $\lim_{j \rightarrow \infty} f_j(x) = f(x) \geq 0$, then we see that f is measurable and non-negative, so the Lebesgue integral of f over E exists. It remains to show that equation (1.4) holds. First we will show that $\int_E f \, d\lambda \geq \lim_{j \rightarrow \infty} \int_E f_j \, d\lambda$. Consider the sequence (z_j) where:

$$z_j := \int_E f_j \, d\lambda$$

note that (z_j) is monotone increasing and hence it converges to some $z \in \bar{\mathbb{R}}$, that is, we write $\lim_{j \rightarrow \infty} z_j = z$. Since $f(x) \geq f_j(x)$ for all $x \in E$ and all $j \in \mathbb{N}$, then we have:

$$\int_E f \, d\lambda \geq \int_E f_j \, d\lambda = z_j$$

hence in the limit, we can write:

$$\int_E f \, d\lambda \geq \lim_{j \rightarrow \infty} \int_E f_j \, d\lambda = z$$

Now we will show that $\int_E f \, d\lambda \leq \lim_{j \rightarrow \infty} \int_E f_j \, d\lambda$. By definition we have:

$$\int_E f \, d\lambda = \sup \left\{ \int_E \phi \, d\lambda \mid \phi \text{ is simple and } 0 \leq \phi(x) \leq f(x) \text{ for all } x \in E \right\} \quad (*)$$

For all such $\phi : E \rightarrow \mathbb{R}$ defined in (*), and for $\alpha \in (0, 1)$, we define:

$$E_j := \{x \in E \mid f_j(x) \geq \alpha \phi(x)\} \subseteq E \quad (i)$$

here $f_j - \alpha \phi$ is measurable, and we have:

$$E_j := (f_j - \alpha \phi)^{-1}([0, \infty]) \in \mathcal{L}(E) \quad (ii)$$

Note that (f_j) is monotone increasing, and hence:

$$E_j \subseteq E_{j+1} \quad (iii)$$

Since (f_j) converges to f from below, and $f \geq \phi$, then we know that for all $x \in E$, there exists $j \in \mathbb{N}$ such that $x \in E_j$, that is $f_j(x) \geq \alpha \phi(x)$. then by (1), we can write:

$$E = \bigcup_{j \in \mathbb{N}} E_j \quad (iv)$$

By Theorem 2.1 and (ii), (iii), (iv), we can write:

$$\lambda(E) = \lambda\left(\bigcup_{j \in \mathbb{N}} E_j\right) = \lim_{j \rightarrow \infty} \lambda(E_j) \quad (\text{v})$$

Since $f_j(x) \geq 0$ for all $x \in E$ and $E_j \subseteq E$, then we get:

$$\int_E f_j d\lambda \geq \int_{E_j} f_j d\lambda \geq \alpha \int_{E_j} \phi d\lambda \quad (\text{vi})$$

take the large limit of (vi), by (v) and simple property of ϕ we get:

$$\lim_{j \rightarrow \infty} \int_E f_j d\lambda \geq \alpha \lim_{j \rightarrow \infty} \int_{E_j} \phi d\lambda = \alpha \int_E \phi d\lambda \quad (\text{vii})$$

here (vii) holds for all $\alpha \in (0, 1)$ and for all ϕ , define the set:

$$K := \left\{ \int_E \phi d\lambda \mid \phi \text{ is simple and } 0 \leq \phi(x) \leq f(x) \text{ for all } x \in E \right\}$$

Here we see that K is bounded above by $\lim_{j \rightarrow \infty} \int_E f_j d\lambda$, which is then greater than the supremum of K , and hence:

$$\int_E f d\lambda \leq \lim_{j \rightarrow \infty} \int_E f_j d\lambda$$

This completes the proof. □

Remark: The statement of Theorem 7.1 can be modified as the following:

Theorem 7.2

Let $f_j : E \rightarrow \bar{\mathbb{R}}$ be Lebesgue measurable non-negative functions defined on the Lebesgue measure space $(E, \mathcal{L}(E), \lambda)$ such that (f_j) converges to a function $f : E \rightarrow \bar{\mathbb{R}}$ pointwise monotonically from below almost everywhere on E . In other words, almost everywhere on E , $(f_j(x))$ is monotone increasing such that $(f_j(x)) \rightarrow f(x)$. Then f is measurable and we have the following holds:

$$\int_E f d\lambda = \int_E \lim_{j \rightarrow \infty} f_j d\lambda = \lim_{j \rightarrow \infty} \int_E f_j d\lambda$$

Theorem 7.3 (Fatou's Lemma)

Let (f_j) be a sequence of Lebesgue measurable non-negative functions defined on a Lebesgue measure space $(E, \mathcal{L}(E), \lambda)$, then we have:

$$\int_E \liminf f_j d\lambda \leq \liminf \int_E f_j d\lambda$$

Proof. For all $k \in \mathbb{N}$, we define:

$$h_k : E \rightarrow \bar{\mathbb{R}} \quad x \mapsto \inf_{n \geq k} f_n(x)$$

note that $h_k(x) \leq f_j(x)$ for all $j \geq k$ and all $x \in E$. Here h_k is measurable by Theorem 5.1, and $h_k \geq 0$, so $\int_E h_k d\lambda$ exists. Now we can write:

$$\int_E h_k d\lambda \leq \int_E f_j d\lambda \quad \forall j \geq k$$

which implies:

$$\int_E h_k d\lambda \leq \inf \left\{ \int_E f_j d\lambda \mid j \geq k \right\} \quad \forall k \in \mathbb{N}$$

thus we have:

$$\lim_{k \rightarrow \infty} \int_E h_k d\lambda \leq \liminf \int_E f_j d\lambda$$

Since (h_k) is increasing, then by Monotone Convergence Theorem, we know that:

$$\lim_{k \rightarrow \infty} \int_E h_k d\lambda = \int_E \lim_{k \rightarrow \infty} h_k d\lambda = \int_E \liminf f_j d\lambda$$

This completes the proof. □

Example: Consider the sequence (f_j) of simple functions defined by the following:

$$f_j : [0, 1] \rightarrow \bar{\mathbb{R}} \quad x \mapsto \begin{cases} j & 0 < x < 1/j \\ 0 & \text{otherwise} \end{cases}$$

note that (f_j) converges to $f : [0, 1] \rightarrow \{0\}$ $x \mapsto 0$ pointwise. Here we also have:

$$0 = \int_0^1 \liminf f_j d\lambda \leq \liminf \int_0^1 f_j d\lambda = 1$$

Theorem 7.4 (Dominated Convergence Theorem)

Let $(E, \mathcal{L}(E), \lambda)$ be a Lebesgue measure space, let (f_j) be a sequence of Lebesgue measurable and integrable functions that converge pointwise to $f : E \rightarrow \bar{\mathbb{R}}$ on E . If there exists an integrable non-negative function $g : E \rightarrow \bar{\mathbb{R}}$ such that $|f_j(x)| \leq g(x)$ for all $x \in E$ and for all $j \in \mathbb{N}$, then f is integrable with:

$$\int_E f d\lambda = \lim_{j \rightarrow \infty} \int_E f_j d\lambda \quad (1.5)$$

Proof. By Theorem 5.1, since f is the pointwise limit of measurable functions (f_j) , then f must be measurable, hence $\int_E |f| d\lambda$ is well defined. Since $|f_j(x)| \leq g(x)$ for all $x \in E$ and for all $j \in \mathbb{N}$, then we have:

$$|f(x)| = \lim_{j \rightarrow \infty} |f_j(x)| \leq g(x) \quad \Rightarrow \quad \int_E |f| d\lambda \leq \int_E g d\lambda < \infty$$

which shows that f is integrable over E . We will use Fatou's Lemma to show that equation (1.5) holds. Note that for all $x \in E$ and all $j \in \mathbb{N}$, we have $|f_j(x)| \leq g$, so $-f_j(x) \leq g(x)$, and hence $f_j(x) + g(x) \geq 0$. By Fatou's Lemma, we can write:

$$\int_E (f + g) d\lambda = \int_E \liminf (f_j + g) d\lambda \leq \liminf \int_E (f_j + g) d\lambda$$

That is, we can write:

$$\int_E f d\lambda + \int_E g d\lambda \leq \liminf \int_E f_j d\lambda + \int_E g d\lambda$$

which implies:

$$\int_E f d\lambda \leq \liminf \int_E f_j d\lambda \quad (\#)$$

Similarly, for all $x \in E$ and all $j \in \mathbb{N}$, we have $|f_j(x)| \leq g$, so $f_j(x) \leq g(x)$, and hence $g(x) - f_j(x) \geq 0$. By Fatou's Lemma, we can write:

$$\int_E \liminf (g - f_j) d\lambda \leq \liminf \int_E (g - f_j) d\lambda$$

which implies:

$$-\int_E f d\lambda \leq \liminf \int_E (-f_j) d\lambda = -\limsup \int_E f_j d\lambda \quad (\#\#)$$

Combining $(\#)$ and $(\#\#)$ we get:

$$\limsup \int_E f_j d\lambda \leq \int_E f d\lambda \leq \liminf \int_E f_j d\lambda$$

but $\liminf k_j \leq \limsup k_j$ for sequence (k_j) of points in $\bar{\mathbb{R}}$, then we must have:

$$\int_E f d\lambda = \lim_{j \rightarrow \infty} \int_E f_j d\lambda$$

This completes the proof. \square

2 | The Algebra and Topology

Vector Spaces

In this text, a field is usually denoted as \mathbb{K} , which is usually considered as \mathbb{R} or \mathbb{C} .

Definition 8.0.0.0.1

A vector space X over a field \mathbb{K} is a set with elements called vectors, on which we have addition $+$ and multiplication \cdot with scalars satisfying the followings:

1. $\forall x, y, z \in X, \exists x + y \in X, x + y = y + x, x + (y + z) = (x + y) + z.$
2. $\exists 0 \in X, \text{ s.t. } \forall x \in X, \text{ we have } x + 0 = x.$
3. $\forall x \in X, \exists -x \in X \text{ s.t. } -x + x = 0.$
4. $\forall \alpha \in \mathbb{K}, \forall x \in X, \alpha x \in X.$
5. $1x = x$ where 1 is the unitary element in $\mathbb{K}.$
6. $\forall \alpha, \beta \in \mathbb{K} \text{ and } x, y \in X, (\alpha\beta)x = \alpha(\beta x), \alpha(x + y) = \alpha x + \alpha y, (\alpha + \beta)x = \alpha x + \beta x.$

Definition 8.0.0.0.2

Let X be a vector space over the field $\mathbb{K}.$

A subspace of X is a set S that is itself a vector space over the field $\mathbb{K}.$

Definition 8.0.0.0.3

Let X be a vector space over the field $\mathbb{K}.$

Consider $A = \{\theta_1, \theta_2, \dots, \theta_N\} \subseteq X$, the span of A is defined by:

$$\text{span}(A) := \left\{ x \in X \mid x = \sum_{j=1}^N \alpha_j \theta_j, \alpha_j \in \mathbb{K} \right\}$$

Consider $A \subseteq X$ is an infinite set, then the span of A is defined by:

$$\text{span}(A) := \left\{ x \in X \mid x = \sum_{j \in J} \alpha_j \theta_j, \alpha_j \in \mathbb{K}, \theta_j \in A, J \text{ is a finite set of indices} \right\}$$

Definition 8.0.0.0.4

Hamel basis of a vector space X over a field \mathbb{K} is a family \mathbb{H} of vectors in X satisfying:

1. Any finite subfamily of \mathbb{H} , denoted as $\{x_j \mid j \in J, J \text{ is a finite set of indices}\}$, is linearly independent, that is, one has $\sum_{j \in J} \alpha_j x_j = \vec{0}$ provided that $\alpha_j = 0$ for all $j \in J$. In such case, the vectors in \mathbb{H} are said to be linearly independent.
2. For all $x \in X$, x is a linear combination of finitely many vectors in $\mathbb{H}.$

Definition 8.0.0.5

A partially ordered set S is a set with order relation \leq that satisfies the followings:

1. Reflexive, that is $x \leq x$ for all $x \in S$.
2. Antisymmetric, that is if we have $x \leq y$ and $y \leq x$, then $x = y$.
3. Transitive, that is if we have $x \leq y$ and $y \leq z$, then $x \leq z$.

Definition 8.0.0.6

Let X be a partially ordered set with order relation \leq , a subset Y of X is said to be totally ordered provided that $\forall x, y \in Y$, we have either $x \leq y$ or $y \leq x$.

Definition 8.0.0.7

Let S be a nonempty partially ordered set that contains a partially ordered set Y , with order relation \leq , an upper bound of Y is a element $m \in S$ such that $y \leq m$ for all $y \in Y$.

Lemma 8.0.1 (Zorn's Lemma)

Consider a nonempty set S that is only partially ordered, such tat every totally ordered subset has an upper bound, then S has at least one maximal element.

The proof of Lemma 8.0.1 involves the Axiom of Choice, here we omit the proof.

Theorem 8.1

Let X be a nontrivial vector space over the field \mathbb{K} , then we have the followings:

1. There exists a Hamel basis for X .
2. All Hamel bases of X have some cardinality which defines $\dim(X)$.

Proof. Let S be a set of all linearly independent families in X , note here for all $u \in X$, if $u \neq 0$, then $\{u\} \in S$. Here we can define a partial order on S : For $E \in S$, and $G \in S$, if E, G can be compared, we say $E \leq G$ if E is contained in G . Take any totally ordered subset Y of S , we define:

$$\mathcal{F} := \bigcup_{G \in Y} G$$

One can check that \mathcal{F} is an upper bound for Y and $\mathcal{F} \in S$. Now by Zorn's Lemma, there exists a maximal element $H \in S$, and we claim that $\text{span}(H) = X$, if not, then there exists $\zeta \in X$, such that $\zeta \notin \text{span}(H)$, hence $H \cup \{\zeta\}$ contradicts that H is a maximal element of S . \square

Metric Spaces

Definition 9.0.0.0.1

Let X be a set, let \mathcal{O} be a family of subsets of X . (X, \mathcal{O}) is said to be a topological space provided that we have the followings hold:

1. $\emptyset \in \mathcal{O}$, $X \in \mathcal{O}$.
2. The intersection of finitely many sets in \mathcal{O} belongs to \mathcal{O} .
3. Let I be an indexing set, finite or infinite, let $\{A_i \mid i \in I\}$ be a subfamily of \mathcal{O} , then we have:

$$\bigcup_{i \in I} A_i \in \mathcal{O}$$

Definition 9.0.0.0.2

Let (X, \mathcal{O}) be a topological space.

$K \subseteq X$ is said to be open provided that $K \in \mathcal{O}$.

$C \subseteq X$ is said to be closed provided that $X - C$ is open.

Example: Let X be a set, $\mathcal{O} = \{\emptyset, X\}$ is a trivial topology on X .

Definition 9.0.0.0.3

Let (X, \mathcal{O}_X) and (Y, \mathcal{O}_Y) be topological spaces, $f : X \rightarrow Y$ is said to be a continuous function provided that the inverse image of f of any open set in \mathcal{O}_Y is open in \mathcal{O}_X .

Definition 9.0.0.0.4

A metric space is a set X equipped with a metric $d : X \times X \rightarrow [0, \infty)$ that satisfies the following properties:

1. $d(x, y) = 0$ if and only if $x = y$.
2. $d(x, y) = d(y, x)$ for all $x, y \in X$.
3. $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$

such metric space is denoted as (X, d) .

Definition 9.0.0.0.5

Let (X, d) be a metric space, for $x \in X$, let $r > 0$, $B_r(x) := \{y \in X, d(x, y) < r\}$ is called an open ball centered at x of radius r .

Definition 9.0.0.0.6

Let (X, d) be a metric space, a set $S \subseteq X$ is said to be open provided that for all $x \in S$, $\exists r > 0$ s.t. $B_r(x) \subseteq S$. $K \subseteq X$ is said to be open provided that $X - K$ is open.

Note: Let (X, d) be a metric space, the collection of open sets \mathcal{O} constitutes a topology on X , and (X, \mathcal{O}) then forms a metrized topological space, and \mathcal{O} is called a topology induced by d , or the metric topology.

Definition 9.0.0.0.7

Let (X, \mathcal{O}) be a topological space, for $x \in X$, a neighborhood of x , denoted as $\mathcal{N}(x)$ is a open set that contains x .

Definition 9.0.0.0.8

Let (X, \mathcal{O}) be a topological space, $S \subseteq X$ is said to be bounded provided that S is contained in some open ball. That is, S is bounded provided that $\exists x \in X$ with $r > 0$ s.t. we have $S \subseteq B_r(x) \subseteq X$.

Definition 9.0.0.0.9

Let (X, \mathcal{O}) be a topological space, (X, \mathcal{O}) is said to be Hausdorff provided that, for all $x, y \in X$ with $x \neq y$, there exists disjoint neighborhood $\mathcal{N}(x)$ and $\mathcal{N}(y)$.

Theorem 9.1

Let (X, \mathcal{O}) be a metrized topological space with metric d , then (X, \mathcal{O}) is Hausdorff and normed.

Proof. For $x, y \in X$ that $x \neq y$, it is trivial that $d(x, y) > 0$, define $\rho = d(x, y)/3$, consider $\mathcal{N}(x) = B_\rho(x)$ and $\mathcal{N}(y) = B_\rho(y)$. For all $z \in \mathcal{N}(x)$, we have:

$$d(x, y) \leq d(x, z) + d(x, y) \quad \Rightarrow \quad d(z, y) \geq 2\rho$$

Hence $z \notin \mathcal{N}(y)$. With similar argument, one can show that $\mathcal{N}(x), \mathcal{N}(y)$ are disjoint. \square

Definition 9.1.0.0.1

Let (X, \mathcal{O}) be a topological space, (X, \mathcal{O}) is said to be normal provided that, for all disjoint closed sets, say $C_1, C_2 \subseteq X$, there exists disjoint open sets $K_1, K_2 \subseteq X$ such that $C_1 \subseteq K_1$ and $C_2 \subseteq K_2$.

Definition 9.1.0.0.2

Let (X, d) be a metric space, for all $S \subseteq X$, and for all $x \in X$,

$$\text{dist}(x, S) := \inf\{d(x, y) \mid y \in S\}$$

Proposition 9.1.1

For normal metrizable topological space equipped with metric d , consider disjoint closed sets $C_1, C_2 \subseteq X$, there exists $\delta > 0$ such that for all $x \in C_2$, $\text{dist}(x, C_1) > \delta$.

Proof. We proceed by contradiction, suppose for all $\delta > 0$, say $\delta = \frac{1}{n}$, there exists $y_n \in C_1$ such that $d(x, y_n) \leq \frac{1}{n}$. We obtain a sequence (y_n) in C_1 which converges to x . Since C_1 is closed, then $x \in C_1$, which contradicts that $x \in C_2$ as C_1, C_2 are disjoint. \square

Lemma 9.1.2

For normal metrizable topological space equipped with metric d , consider disjoint closed sets $C_1, C_2 \subseteq X$, then there exists some $\delta > 0$ such that we have:

$$\text{dist}(C_1, C_2) := \inf\{d(x, C_1) \mid x \in C_2\} \geq \delta$$

Proof. We first define:

$$K_1 := \bigcup_{x_1 \in C_1} B_{d(x_1, C_2)/2}(x_1), \text{ which is open in } C_1$$

$$K_2 := \bigcup_{x_2 \in C_2} B_{d(x_2, C_1)/2}(x_2), \text{ which is open in } C_2$$

We will show that $K_1 \cap K_2 = \emptyset$, we proceed by contradiction, suppose there exists $y \in K_1 \cap K_2$, then $\exists x_1 \in C_1$ s.t. $y \in B_{d(x_1, C_2)/2}(x_1)$, and $\exists x_2 \in C_2$ s.t. $y \in B_{d(x_2, C_1)/2}(x_2)$, here we can write:

$$\begin{aligned} d(x_1, x_2) &= d(x_2, x_1) \leq d(x_1, y) + d(x_2, y) \\ &< \frac{1}{2} (d(x_1, C_2) + d(x_2, C_1)) \leq \max\{d(x_1, C_2), d(x_2, C_1)\} \end{aligned}$$

Considering the cases $d(x_1, C_2) > d(x_2, C_1)$ and $d(x_1, C_2) < d(x_2, C_1)$ separately will get us the desired contradiction, which completes the proof. \square

Definition 9.1.2.0.1

Let (x_n) be a sequence in metric space (X, d) . (x_n) is convergent provided that $\exists x \in X$ s.t. $\forall \mathcal{N}(x), \exists N \in \mathbb{N}$ s.t. $x_m \in \mathcal{N}(x) \forall m \geq N$.

Theorem 9.2

In Hausdorff space, the limit of a sequence is unique.

Proof. Suppose X is a Hausdorff space and let (x_n) be a sequence in X that converges to x and x' . We proceed by contradiction, suppose $x \neq x'$, then there exist disjoint $\mathcal{N}(x)$ and $\mathcal{N}(x')$, which in such case (x_n) converges to only one of $\mathcal{N}(x)$ and $\mathcal{N}(x')$, one will reach a contradiction. \square

Note: In a metrized space, a set C is closed if it contains all of its limit points.

Definition 9.2.0.0.1

Let S be a subset of a topological space X , the interior of S , denoted as $\text{Int}(S)$, is defined to be the union of all open sets, open in X , that is contained in S .

Definition 9.2.0.0.2

Let S be a subset of a topological space X , the closure of S , denoted as \bar{S} , is the smallest closed set that contains S . Equivalently, the closure of S is the intersection of all closed sets, closed in X , that is contained in S . Equivalently, the closure of S is the union of S with the set of the limit points of S .

Note: Let S be a subset of a metrizable space X , for all $x \in \bar{S}$, there exists (x_n) in S such that (x_n) converges to x .

Definition 9.2.0.0.3

Let X be a set, let $f : X \rightarrow \mathbb{C}$,

$$\text{supp}(f) = \{x \in X \mid f(x) \neq 0\}$$

Definition 9.2.0.0.4

Let X be a topological space, let $S \subseteq X$, S is said to be dense in X provided that $\bar{S} = X$.

Definition 9.2.0.0.5

A metric space is separable provided that it has a countable dense subset.

Example: \mathbb{R} equipped with the Euclidean metric is separable as $\bar{\mathbb{Q}} = \mathbb{R}$.

Completeness

Definition 10.0.0.0.1

Let (X, d_X) and (Y, d_Y) be metric spaces, and let $i : X \rightarrow Y$ be a function. i is said to be an isometry provided that for all $x, z \in X$, we have:

$$d_Y(i(x), i(z)) = d_X(x, z)$$

Note: Isometries are injective and continuous. Let (X, d_X) and (Y, d_Y) be metric spaces, let $i : X \rightarrow Y$ be an isometry, then $\forall x, z \in X$, we have $i(z) = i(x)$ if and only if $d_Y(i(z), i(x)) = 0 \Rightarrow d_X(x, z) = 0 \Rightarrow x = z$. One can show the continuity of i similarly.

Definition 10.0.0.0.2

An bijective isometry is called an isomorphism between metric spaces. A metric space (X, d_X) is said to be isomorphic to a metric space (Y, d_Y) provided that there exists an isomorphism $i : X \rightarrow Y$.

Definition 10.0.0.0.3

Let (X, d) be a metric space, let $A \subseteq X$, then in the settings of extended reals, we define the diameter of A :

$$\text{diam}(A) := \sup\{d(x_n, x_m) \mid x_n, x_m \in A\} \in \bar{\mathbb{R}}$$

Definition 10.0.0.0.4

Let (X, d) be a metric space, a sequence (x_n) of points in X is said to be Cauchy provided that the sequence $(\text{diam}(\bigcup_{i \geq n} \{x_i\}))$ converges to 0.

Lemma 10.0.1

Cauchy sequences are bounded.

Proof. Let (X, d) be a metric space, let (x_n) be a Cauchy sequence in X , then by definition of Cauchy sequence, there exists $N \in \mathbb{N}$ such that we have:

$$\sup\{d(x_m, x_n) \mid n, m \geq N\} < 1$$

then for all $n \geq N$, we have $x_n \in B_1(x_N)$. Now define:

$$\rho := \max\{d(x_n, x_N) \mid 1 \leq n \leq N\}$$

Now for all $n \geq 1$, we have $x_n \in B_r(x_N)$ where $r := \max\{1, \rho\}$. □

Note: Convergent sequences in a metric space are Cauchy.

Lemma 10.0.2

Let (x_n) be a Cauchy sequence in a metric space, let $(x_{\sigma(k)})$ be a subsequence of (x_n) where $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ is an injection. If $(x_{\sigma(k)})$ converges, then (x_n) converges.

Definition 10.0.2.0.1

A metric space is said to be complete provided that every Cauchy sequence in such space converges.

Theorem 10.1

Let (X, d) be a metric space, let $S \subseteq X$. $(S, d|_S)$ is a complete metric space if and only if the set S is closed in X .

Example: The set of real numbers equipped with the Euclidean metric is complete. The set $[0, \infty)$ is also complete if equipped with the Euclidean metric. However, $[0, \infty)$ equipped with the metric d_k defined by the following is not complete:

$$d_k : [0, \infty) \times [0, \infty) \rightarrow [0, \infty) \quad (x, y) \mapsto \left| \frac{x}{1+x} - \frac{y}{1+y} \right|$$

Proposition 10.1.1

$[0, \infty)$ equipped with the metric d_k defined by the following is not complete:

$$d_k : [0, \infty) \times [0, \infty) \rightarrow [0, \infty) \quad (x, y) \mapsto \left| \frac{x}{1+x} - \frac{y}{1+y} \right|$$

Proof. Consider the sequence (x_n) defined by $x_n = n$ for all $n \in \mathbb{N}$. Notice that we have:

$$d_k(x_n, 0) = \frac{n}{n+1} < 1$$

Moreover, we can write:

$$d_k(x_n, x_m) = \left| \frac{n}{n+1} - \frac{m}{m+1} \right| \leq \left| \frac{n}{n+1} - 1 \right| + \left| \frac{m}{m+1} - 1 \right| = \frac{1}{n+1} + \frac{1}{m+1}$$

For all $\epsilon > 0$, we let $N = \lfloor \frac{2}{\epsilon} \rfloor + 19$, then for all $n, m \geq N$, we can write:

$$d(x_n, x_m) \leq \frac{1}{n+1} + \frac{1}{m+1} \leq \frac{2}{N+1} < \epsilon$$

However, if $\exists x \in S$ such that $(x_n) \rightarrow x$, then we can write:

$$d(x_n, x) = \left| \frac{n}{n+1} - \frac{x}{x+1} \right|$$

here $d(x_n, x)$ tends to $\frac{1}{1+x}$ as n tends to ∞ . Hence (x_n) does not converge in X . \square

Definition 10.1.1.0.1

Let (X, d) be a metric space, (\bar{X}, \bar{d}) is said to be a completion of (X, d) provided that:

1. There exists an isometric embedding $i : X \rightarrow \bar{X}$ and $i(X)$ is dense in \bar{X} .
2. (\bar{X}, \bar{d}) is a complete metric space.

Definition 10.1.1.0.2

Let S be a set, a relation between elements in S , denoted as \sim , is said to be equivalent provided that we have:

1. $\forall a \in S, a \sim a$.
2. $\forall a, b \in S, a \sim b$ implies $b \sim a$.
3. $\forall a, b, c \in S, a \sim b$ and $b \sim c$ imply $a \sim c$.

Definition 10.1.1.0.3

Let S be a set, and let \sim be an equivalent relation on S , for $a \in A$, $[a] := \{y \in S \mid y \sim a\}$ is called an equivalent class of elements in A that are equivalent to a . Elements in $[a]$ are called the representers.

Lemma 10.1.2

Let (X, d) be a metric space, the function d is continuous.

Theorem 10.2

Every metric space has a completion that is unique up to an isomorphism.

Proof. Let (X, d) be a metric space, and let (\bar{X}_1, \bar{d}_1) , (\bar{X}_2, \bar{d}_2) be completions of (X, d) . One needs to construct an isomorphism $\bar{i} : \bar{X}_1 \rightarrow \bar{X}_2$. By definition of completion, there exists isometric embeddings $i_1 : X \rightarrow \bar{X}_1$ and $i_2 : X \rightarrow \bar{X}_2$ such that $i_1(X)$ is dense in \bar{X}_1 and $i_2(X)$ is dense in \bar{X}_2 . Consider $x \in \bar{X}_1$, there exists a sequence (x_n) in X such that $(i_1(x_n))$ converges to x by the density of $i_1(X)$ in \bar{X}_1 . First we will show that (x_n) is Cauchy in X . Note that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n, m \in \mathbb{N}$ with $n, m \geq N$, we have $\bar{d}_1(i_1(x_n), i_1(x_m)) < \epsilon$, since i_1 is an isometry, hence $\bar{d}_1(i_1(x_n), i_1(x_m)) = d(x_n, x_m) < \epsilon$, so (x_n) is Cauchy in X . Now consider the sequence $(i_2(x_n))$, which is a sequence in \bar{X}_2 , which is also Cauchy in \bar{X}_2 because i_2 is an isometry, so $\bar{d}_2(i_2(x_n), i_2(x_m)) < \epsilon$ for all $n, m \in \mathbb{N}$ with $n, m > N$. Since \bar{X}_2 is complete, then $\bar{i}(x) := \lim_{n \rightarrow \infty} i_2(x_n) \in \bar{X}_2$. Now suppose that (x_n) and (z_n) are sequences in X such that $(i_1(x_n)) \rightarrow x$ and $(i_1(z_n)) \rightarrow x$, here we see that:

$$\bar{d}_2(i_2(z_n), i_2(x_n)) = d(z_n, x_n) = \bar{d}_1(i_1(x_n), i_1(z_n))$$

and we have $\bar{d}_1(i_1(x_n), i_1(z_n))$ tends to 0 as n tends to 0 because both (x_n) and (z_n) converges to x . Since metric function is continuous, then one can conclude that $(i_2(z_n))$ and $(i_2(x_n))$ have the same limit $\bar{i}(x)$. In such way, we can define a function $\bar{i} : \bar{X}_1 \rightarrow \bar{X}_2$ well defined for all $x \in \bar{X}_1$, where we write,

$$\forall x \in \bar{X}_1, \exists (x_n) \text{ of points in } X, \text{ s.t. } i_1(x_n) \rightarrow x, \text{ and } \bar{i}(x) := \lim_{n \rightarrow \infty} i_2(x_n) \quad (*)$$

We will show that we have the following holds for all $x, z \in \bar{X}_1$:

$$\bar{d}_2(\bar{i}(x), \bar{i}(z)) = \lim_{n \rightarrow \infty} \bar{d}_2(i_2(x_n), i_2(z_n))$$

where (z_n) and (x_n) are defined using (*). Here we can write the following for all $n \in \mathbb{N}$:

$$\begin{aligned} \bar{d}_2(\bar{i}(x), \bar{i}(z)) &\leq \bar{d}_2(\bar{i}(x), i_2(x_n)) + \bar{d}_2(i_2(x_n), \bar{i}(z)) \\ &\leq \bar{d}_2(\bar{i}(x), i_2(x_n)) + \bar{d}_2(i_2(x_n), i_2(z_n)) + \bar{d}_2(i_2(z_n), \bar{i}(z)) \end{aligned}$$

$$\bar{d}_2(i_2(x_n), i_2(z_n)) \leq \bar{d}_2(\bar{i}(x), \bar{i}(z)) + \bar{d}_2(\bar{i}(x), i_2(x_n)) + \bar{d}_2(\bar{i}(z), i_2(z_n))$$

which implies:

$$|\bar{d}_2(\bar{i}(x), \bar{i}(z)) - \bar{d}_2(i_2(x_n), i_2(z_n))| \leq \bar{d}_2(\bar{i}(x), i_2(x_n)) + \bar{d}_2(\bar{i}(z), i_2(z_n))$$

Now we can write:

$$\begin{aligned} \bar{d}_2(\bar{i}(x), \bar{i}(z)) &= \lim_{n \rightarrow \infty} \bar{d}_2(i_2(x_n), i_2(z_n)) \\ &= \lim_{n \rightarrow \infty} d(x_n, z_n) = \lim_{n \rightarrow \infty} \bar{d}_1(i_1(x_n), i_1(z_n)) = \bar{d}_1(x, z) \end{aligned}$$

Hence we have shown that \bar{i} is an isometry. Now we want to show that $\bar{i}(\bar{X}_1) = \bar{X}_2$, that is, we want to show $\forall x \in \bar{X}_2$, there exists $z \in \bar{X}_1$ such that $x = \bar{i}(z)$. Note that there

exists (x_n) in X such that $(i_2(x_n))$ converges to x as \bar{X}_2 is a completion of X . (x_n) is Cauchy in X , so $(i_1(x_n))$ is Cauchy in \bar{X}_1 , then there exists $z \in \bar{X}_1$ such that $(i_1(x_n))$ converges to z . Then we can write:

$$\bar{i}(z) = \lim_{n \rightarrow \infty} i_2(x_n) = x$$

This completes the proof that any completion of (X, d) is isomorphic. Now we will construct a completion for (X, d) . Let $S := \{\text{the set of all Cauchy sequence in } X\}$, and here we will define a equivalence relation on S , for $x, y \in S$, where $x = (x_n)$ and $y = (y_n)$, then we write $x \sim y$ provided that $\lim_{n \rightarrow \infty} d(x_n, y_n) = 0$. Hence we can define $\bar{X} := \{\text{the set of equivalence classes on } S\}$, and the metric \bar{d} , for $\zeta, \eta \in \bar{X}$, let (x_n) denote a representer of ζ and (t_n) be a representer of η , then we know that $(d(x_n, t_n))$ is a Cauchy sequence in \mathbb{R} . Using the triangle inequality, we can write:

$$\begin{aligned} d(x_n, t_n) &\leq d(x_g, t_g) + d(x_g, t_n) + d(t_g, t_n) \\ d(x_g, t_g) &\leq d(x_n, t_n) + d(x_g, x_n) + d(t_g, t_n) \end{aligned}$$

combining we have:

$$|d(x_n, t_n) - d(x_g, t_g)| \leq d(x_g, x_n) + d(t_g, t_n)$$

Now we define \bar{d} , with the notation that (x_n) is a representer for ζ and (t_n) is a representer for η :

$$\bar{d} : \bar{X} \times \bar{X} \rightarrow \mathbb{R} \quad (\zeta, \eta) \mapsto \lim_{n \rightarrow \infty} d(x_n, t_n)$$

Here we will show that \bar{d} is well defined, suppose (x'_n) is also a representer for ζ , and (t'_n) is also a representer for η . Employing similar argument above, we can write:

$$|d(x_n, t_n) - d(x'_n, t'_n)| \leq d(x'_n, x_n) + d(t'_n, t_n)$$

taking large n limit, and using the continuity of metric function, we can write:

$$\lim_{n \rightarrow \infty} d(x_n, t_n) = d(x'_n, t'_n) = \bar{d}(\zeta, \eta)$$

We now will check that \bar{d} is a metric. Notice that $\bar{d}(\zeta, \eta) \geq 0$ for all $\zeta, \eta \in \bar{X}$ as d is bounded below by 0. When $\bar{d}(\zeta, \eta) = 0$, it is easy to check that $\zeta = \eta$ by the equivalence relation and the converse holds true. Also, for $\zeta, \eta \in \bar{X}$, we get $\bar{d}(\zeta, \eta) = \bar{d}(\eta, \zeta)$. For the triangle inequality, we write the following for $\zeta, \eta, \xi \in \bar{X}$, with (z_n) be representer for ζ , (x_n) be representer for η , and (t_n) be representer for ξ :

$$d(x_n, t_n) \leq d(x_n, z_n) + d(z_n, t_n) \quad (**)$$

taking large n limit of $(**)$ we get the desired triangle inequality for \bar{d} . Now we want to define an isometry $i : X \rightarrow \bar{X}$. Take $x \in X$, we define:

$$i : X \rightarrow \bar{X} \quad x \mapsto \{y \in S \mid y \sim (x_i) \text{ where } x_i := x\}$$

Notice that, let (y_n) be a representer for y defined in the definition of i , we have:

$$\lim_{n \rightarrow \infty} d(y_n, x) = 0$$

It remains to show that i is an isometry. Let $x, t \in X$, we want to show that we have:

$$\bar{d}(i(x), i(t)) = d(x, t) \quad (***)$$

Let (x_n) be a sequence that $x_i = x$, and let (t_n) be a sequence that $t_i = t$, hence we get:

$$\bar{d}(i(x), i(t)) = \lim_{n \rightarrow \infty} d(x_n, t_n) = d(x, t)$$

as we see that $(***)$ follows. Now, we will show that $\overline{i(X)} = \bar{X}$. For $\zeta \in \bar{X}$, and let $\epsilon > 0$ be given, we would like to find $t \in X$ such that $\bar{d}(\zeta, i(t)) < \epsilon$. Take any representer (x_n) for ζ , which is Cauchy by definition, so there exists $N \in \mathbb{N}$ such that for all $n, m \in \mathbb{N}$ with $n, m > N$, we have $d(x_n, x_m) < \epsilon$, define $t := x_N$, so $d(x_n, t) < \epsilon$ for all $n \geq N$. Here we can write:

$$\bar{d}(\zeta, i(t)) = \lim_{n \rightarrow \infty} d(x_n, x_N) < \epsilon$$

the result that $\overline{i(X)} = \bar{X}$ follows. Finally, we will show that (\bar{X}, \bar{d}) is complete. Take any Cauchy sequence (ξ_n) in \bar{X} . Since $i(X)$ is dense in \bar{X} , then for all $k \geq 1$, there exists $x_k \in X$ such that $\bar{d}(\xi_k, i(x_k)) < 1/k$. Here one gets a sequence (x_k) of points in X . Let x_n and x_k denote elements in the sequence (x_n) , here we write:

$$d(x_k, x_n) = \bar{d}(i(x_k), i(x_n)) \leq \bar{d}(i(x_k), \xi_k) + \bar{d}(i(x_n), \xi_n) + \bar{d}(\xi_k, \xi_n) \leq \frac{1}{k} + \frac{1}{n} + \bar{d}(\xi_k, \xi_n)$$

in which case we can conclude that (x_k) is Cauchy as (ξ_n) is a Cauchy sequence. Denote ξ be the equivalence class of (x_n) , here $\xi \in \bar{X}$, we will show that $\xi = \lim_{n \rightarrow \infty} \xi_n$. Here we can write with x_n being an element in the sequence (x_k) :

$$\bar{d}(\xi_n, \xi) \leq \bar{d}(\xi_n, i(x_n)) + \bar{d}(i(x_n), \xi) = \bar{d}(\xi_n, i(x_n)) + \lim_{g \rightarrow \infty} d(x_n, x_g) \quad (\dagger)$$

Since (x_k) is Cauchy, then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x_g) < \epsilon$ for all $n, g \in \mathbb{N}$ with $n, g \geq N$. Then take $n \geq N$, (\dagger) becomes:

$$\bar{d}(\xi_n, \xi) < \frac{1}{n} + \epsilon$$

which shows that (ξ_n) converges to ξ . This completes the proof that (\bar{X}, \bar{d}) is complete. Hence completes the proof of the theorem. \square

Example:

Consider now the sequence of functions (f_k) defined by:

$$f_k : [-1, 1] \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} 0 & x \leq 0 \\ kx & x \in (0, 1/n) \\ 1 & x \geq 1/n \end{cases}$$

Here $f_k \in C([-1, 1])$. Suppose $C([-1, 1])$ is equipped with the metric:

$$d : C([-1, 1]) \times C([-1, 1]) \quad (f, g) \mapsto \left(\int_{-1}^1 |f - g|^2 \right)^{1/2}$$

One can show that (f_k) converges to the function ϕ defined by the following pointwise:

$$\phi : [-1, 1] \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} 1 & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Note that $C([-1, 1])$ is not a complete space, as we extend the definition of domain of d to include a set that contains the following function:

$$\phi_\alpha : [-1, 1] \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} 1 & x > 0 \\ 0 & x < 0 \\ \alpha & x = 0 \end{cases}$$

and it is easy to show that $(d(f_n, \phi_\alpha)) \rightarrow 0$ as $n \rightarrow \infty$.

Theorem 10.3 (Riesz-Fischer Theorem)

Let (f_n) be a Cauchy sequence of functions in a metric space of functions.

There exists a subsequence of (f_n) that converges pointwise.

Compactness

Definition 11.0.0.0.1

Let (X, \mathcal{O}) be a topological space, let $K \subseteq X$, a cover of K is a set $\mathcal{C} = \{G_\alpha \mid \alpha \in J\}$ where J is an index set such that we have:

$$K \subseteq \bigcup_{\alpha \in J} G_\alpha \quad (2.1)$$

And \mathcal{C} is said to be an open cover of K provided that (2.1) holds and G_α is open for all $\alpha \in J$. \mathcal{C} is said to be a finite cover of K provided that (2.1) holds and J is a finite set.

Definition 11.0.0.0.2

Let (X, \mathcal{O}) be a topological space, $K \subseteq X$ is said to be compact provided that every open cover of K admits a finite subcover.

Definition 11.0.0.0.3

Let (X, d) be a metric space, $K \subseteq X$ is said to be sequentially compact provided that every sequence of points in K has a subsequence that converges in K .

Theorem 11.1 (Theorem 1.62 on Hunter and Nachtergaele)

In a metric space, a set is compact if and only if the set is sequentially compact.

Theorem 11.2 (Heine Borel Theorem)

In \mathbb{R}^n , a set $K \subseteq \mathbb{R}^n$ is compact if and only if K is closed and bounded.

Definition 11.2.0.0.1

Let (X, d) be a metric space, let $K \subseteq X$, let $\epsilon > 0$, and let J be an index set. $\mathcal{N} := \{x_\alpha \in X \mid \alpha \in J\}$ is called an ϵ -net of K provided that we have:

$$K \subseteq \bigcup_{\alpha \in J} B_\epsilon(x_\alpha) \quad (2.2)$$

And \mathcal{N} is called a finite ϵ -net provided that (2.2) holds and J is finite.

Definition 11.2.0.0.2

Let (X, d) be a metric space, $K \subseteq X$ is said to be totally bounded provided that it has a finite ϵ -net for all $\epsilon > 0$.

Note: Totally boundedness of a set implies the boundedness of that set.

Theorem 11.3

Let (X, d) be a metric space, let $K \subseteq X$.

The set K is compact if and only if K is complete and totally bounded.

Proof. For the \Leftarrow direction, suppose K is complete and totally bounded. Let (x_n) be a sequence in K . For $\epsilon = 1$, there exists a finite 1-net $\{\xi_j^{(1)} \mid j \in \mathbb{N}, j \leq N_1 \in \mathbb{N}\}$ such that we have:

$$K \subseteq \bigcup_{j=1}^{N_1} B_1(\xi_j^{(1)}) \quad (2.3)$$

Then there exists at least one of these balls in (2.3), denoted as $B^{(1)}$ that has infinitely many terms in the sequence. Pick $x_{\sigma(1)} \in B^{(1)}$, now we form a subsequence $(x_k^{(1)})$ of (x_n) by removing $x_{\sigma(1)}$ and all $x_n \notin B^{(1)}$. Next, $B^{(1)} \cap K$ is totally bounded and has finite $\frac{1}{2}$ -net $\{\xi_i^{(2)} \mid i \in \mathbb{N}, i \leq N_2 \in \mathbb{N}\}$ such that we can write:

$$B^{(1)} \cap K \subseteq \bigcup_{i=1}^{N_2} B_{\frac{1}{2}}(\xi_i^{(2)})$$

Hence there exists a ball $B^{(2)}$ that has infinitely many terms of $(x_k^{(1)})$, and we pick $x_{\sigma(2)} \in B^{(2)}$. Continue this process and we obtain a subsequence $(x_{\sigma(k)})$ of (x_n) with the property that:

$$x_{\sigma(j)} \in \bigcap_{k=1}^j B^{(k)} \quad \forall j \in \mathbb{N}$$

We will show that $(x_{\sigma(k)})$ is a Cauchy sequence. Consider $k \leq n$, we can write:

$$d(x_{\sigma(k)}, x_{\sigma(n)}) < \frac{2}{2^{k-1}}$$

hence one can easily verify that $(x_{\sigma(k)})$ is Cauchy. Since K is complete by assumption, then we know that $(x_{\sigma(k)})$ converges to some $x \in K$, this completes the proof of the \Leftarrow direction. Now for the \Rightarrow direction, we suppose K is compact, that is K is also sequentially compact by Theorem 11.1. Take Cauchy sequence (x_n) in K , which has a convergent subsequence $(x_{\sigma(k)})$ that converges to $x \in K$ by the sequentially compactness of K , but then the entire Cauchy sequence (x_n) converges to x , hence K is complete. To show that K is totally bounded, we proceed by contradiction, let $\epsilon > 0$ such that K has no finite ϵ -net. Take any finite subset $\{x_j \mid j \in \mathbb{N}, j \leq n \in \mathbb{N}\}$ of K , we know that:

$$K \not\subseteq \bigcup_{j=1}^n B_\epsilon(x_j)$$

then there exists $x_{n+1} \in K$ such that:

$$x_{n+1} \notin \bigcup_{j=1}^n B_\epsilon(x_j)$$

In such way, one can obtain a sequence (x_n) with the fact that $d(x_m, x_n) \geq \epsilon$ for all $m \neq n$, hence (x_n) has no convergent subsequence, which contradicts the compactness of K , this completes the proof. \square

Theorem 11.4

A compact metric space is separable.

Proof. Let (K, d) be a compact metric space. For all $n \geq 1$, there exists a finite $\frac{1}{n}$ -net A_n , then we know that K can be covered by balls of radius $\frac{1}{n}$ with centers in A_n . Now we define:

$$A := \bigcup_{n=1}^{\infty} A_n$$

Note that the countable union of finite set is countable, hence A is countable. Now $\forall x \in K$, and for all $\epsilon > 0$, by construction there exists some $\xi \in A$ such that $d(x, \xi) < \epsilon$. Here such ξ can be chose by letting $N \in \mathbb{N}$ satisfies $\frac{1}{N} < \epsilon$, then there exists $\xi \in A_N$ such that $d(\xi, x) < \epsilon$. This proves that K is separable as we have shown that A is a countable dense subset of K . \square

Definition 11.4.0.0.1

Let (X, d) be a metric space, and let $K \subseteq X$, K is said to be precompact provided that \bar{K} is compact.

3 | Banach Spaces

Normed Vector Spaces

Definition 12.0.0.0.1

Let X be a vector space over the field \mathbb{K} , a norm on X is a function $\|\cdot\| : X \rightarrow [0, \infty)$ which satisfies:

1. For $x \in X$, $\|x\| = 0$ if and only if $x = 0$.
2. For $x \in X$, and $\alpha \in \mathbb{K}$, we have $\|\alpha x\| = |\alpha| \|x\|$.
3. For $x, y \in X$, we have $\|x + y\| \leq \|x\| + \|y\|$.

Definition 12.0.0.0.2

A vector space X equipped with a norm $\|\cdot\|$ is called a normed space, denoted as $(X, \|\cdot\|)$.

Theorem 12.1

Let $(X, \|\cdot\|)$ be a normed space over the field \mathbb{K} , $d : X \times X \rightarrow \mathbb{R} \quad (x, y) \mapsto \|x - y\|$ is a metric on X with the following properties:

1. $d(x, y) = d(x + z, y + z)$ for all $x, y, z \in X$, called the translation invariant property.
2. $d(\alpha x, \alpha y) = |\alpha| d(x, y)$ for all $x, y \in X$, called the positively homogeneous property.
3. $B_r(x)$ is a convex set.

The topology \mathcal{O} induced by this metric d is called the strong, or norm, topology on X .

Proof. d is well defined by checking the definition of the norm $\|\cdot\|$ on X , here we note the triangle inequality of d :

$$\|x - y\| = \|(x - z) - (y - z)\| \leq \|x - z\| + \|y - z\|$$

hence we see immediately that $d(x, y) \leq d(x, z) + d(z, y)$. The translation invariant and positively homogeneous of d are immediate. We will show that $B_r(x)$ is convex, let $y, z \in B_r(x)$ for $x, y, z \in X$ and $r > 0$, consider $\alpha \in [0, 1]$, then we can write:

$$\begin{aligned} d(\alpha y + (1 - \alpha)z, x) &= \|\alpha y + (1 - \alpha)z - x\| \\ &= \|\alpha(y - x) + (1 - \alpha)(z - x)\| \\ &\leq \|\alpha(y - x)\| + \|(1 - \alpha)(z - x)\| \\ &= \alpha\|y - x\| + (1 - \alpha)\|z - x\| \\ &< r \end{aligned}$$

This completes the proof. □

Lemma 12.1.1

A norm $\|\cdot\| : X \rightarrow \mathbb{R}$ defined on a vector space X is a continuous function.

Proof. Let $x, y \in X$, we can write:

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

rearranging we have:

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

The result follows. □

Definition 12.1.1.0.1

A Banach space is a normed vector space that is complete in the norm topology.

Theorem 12.2 (Baire Category Theorem)

Let (X, d) be a complete metric space and let (V_n) be a sequence of open sets with the property that V_n is dense in X for all $n \in \mathbb{N}$. Let W be any non-empty open set in X , show that the intersection of W with $\bigcap_{n=1}^{\infty} V_n$ is not empty.

Proof. Note that the intersection of two open sets is open, hence we know that $V_n \cap W$ is open for all $n \in \mathbb{N}$. Given $n \in \mathbb{N}$, since V_n is dense in X , then $x \in W \subseteq X$ is a limit point of V_n by Corollary 4.0.1.1. Note that for $x \in W$, since W is open, then $\exists r > 0$ such that $B(x, r) \subseteq W$. On the other hand x is a limit point for V_n , then there exists a sequence (t_j) of points in V_n that converges to x , and hence there exists some $N \in \mathbb{N}$ such that for $i \in \mathbb{N}$ with $i \geq N$, we have $t_i \in B(x, r) \subseteq W$ and $t_i \in V_n$, here we relabel $m_n := t_i$ for some $i \geq N$, so $m_n \in W \cap V_n$. This shows that the intersection of V_n and any nonempty open set is open and nonempty, and in particular, $V_n \cap W$ is open and nonempty with $m_n \in V_n \cap W$, so we can find $r'_n > 0$ such that $m_n \in B(m_n, r'_n) \subseteq V_n \cap W$, then we can take:

$$r_n = r'_n/2 \tag{3.1}$$

such that we have:

$$K_n := \overline{B(m_n, r_n)} \subseteq B(m_n, 2r_n) = B(m_n, r'_n) \subseteq V_n \cap W \tag{3.2}$$

Now for $n + 1 \in \mathbb{N}$, since $B(m_n, r_n)$ is open and nonempty, then $V_{n+1} \cap B(m_n, r_n)$ is open and nonempty by the argument above, so there exists $r'_{n+1} > 0$ and $m_{n+1} \in V_{n+1} \cap B(m_n, r_n)$ such that $B(m_{n+1}, r'_{n+1}) \subseteq V_{n+1} \cap B(m_n, r_n)$. Here we take:

$$r_{n+1} := \min(r'_{n+1}, r_n)/2 \tag{3.3}$$

It is easy to see that:

$$K_{n+1} := \overline{B(m_{n+1}, r_{n+1})} \subseteq B(m_{n+1}, r'_{n+1}) \subseteq V_{n+1} \cap B(m_n, r_n) \subseteq B(m_n, r_n) \subseteq K_n \tag{3.4}$$

Since $n \in \mathbb{N}$ is arbitrary, we can set $n = 1$ to obtain K_1 and r_1 by (3.1) and (3.2), then set $n = k$ to obtain K_{k+1} and r_{k+1} using (3.3) and (3.4) for all $k \in \mathbb{N}$. Hence we obtain a sequence of nonempty closed sets (K_j) , and a sequence of the radius (r_j) of the closed sets with the following properties:

1. By (3.4), $K_{i+1} \subseteq K_i$ for all $i \in \mathbb{N}$
2. By (3.3), $\text{diam}(K_{i+1}) = 2r_{i+1} \leq r_i < 2r_i = \text{diam}(K_i)$ for all $i \in \mathbb{N}$

Using Property (2), denoting $r = \text{diam}(K_1)$, we have $\text{diam}(K_i) \leq r/2^i$ for all $i \in \mathbb{N}$. Since $r < \infty$ and the sequence (2^j) is not bounded above, then by the Archimedean Property, for $\epsilon > 0$, there exists some $k \in \mathbb{N}$ and $t \in \mathbb{N}$ such that $2^t > rk$ and $1/k < \epsilon$, so we have $\text{diam}(K_t) \leq r/2^t < 1/k < \epsilon$, which implies the sequence $(\text{diam}(K_j))$ converges to 0. By Q3 in Math 556 HW1, since (K_j) is a sequence of nested nonempty closed sets with $(\text{diam}(K_j)) \rightarrow 0$, and X is a complete metric space, then we know that $\bigcap_{j \in \mathbb{N}} K_j = \{x\}$ with some $x \in X$. Here notice that $x \in \bigcap_{j \in \mathbb{N}} K_j$ and hence $x \in K_i$ for all $i \in \mathbb{N}$, hence by (3.4) and (3.2), $x \in V_i$ for all $i \in \mathbb{N}$, and by (3.2), we also have $x \in W$. This shows that:

$$\{x\} \subseteq W \cap \left(\bigcap_{j \in \mathbb{N}} V_j \right)$$

which completes the proof. \square

Corollary 12.2.1

Let (X, d) be a metric space, and let (F_n) be a sequence of closed set such that:

$$X = \bigcup_{n=1}^{\infty} F_n$$

then we have either one of the two of the followings:

1. If $\text{Int}(F_n) = \emptyset$ for all $n \in \mathbb{N}$, then X is not complete.
2. If X is complete, then there exists at least one $n \in \mathbb{N}$ such that $\text{Int}(F_n) \neq \emptyset$.

Theorem 12.3

Let $(X, \|\cdot\|)$ be a infinite dimensional Banach space over the field \mathbb{F} , then X does not have a countable Hamel basis.

Proof. The proof of Theorem 12.2 follows from the Baire Category Theorem. We proceed by contradiction, suppose there exists a countable Hamel basis $\mathbb{H} = \{h_j \mid j \geq 1\}$. Let $\mathbb{F}_n := \text{span}(h_1, h_2, \dots, h_n)$, which is a closed set and $\text{Int}(\mathbb{F}_n) = \emptyset$ for all $n \in \mathbb{N}$ as we will show in the following. Suppose now $\text{Int}(\mathbb{F}_n) \neq \emptyset$, then there exists $x \in X$, $r > 0$ such that $B_r(x) \subseteq \mathbb{F}_n$, now take:

$$y = x + \frac{r}{2} \frac{h_{n+1}}{\|h_{n+1}\|}$$

then we can write $d(x, y) = \|x - y\| = r/2$, hence $y \in B_r(x) \subseteq \mathbb{F}_n$, and so $y - x = h_{n+1} \in \mathbb{F}_n$, but this leads to a contradiction. Hence $\text{Int}(\mathbb{F}_n) = \emptyset$ for all $n \in \mathbb{N}$, then by Corollary 12.2.1, we know that X is not complete, but X is a Banach space, another contradiction, which completes the proof. \square

Example: Let P be a set of polynomials on defined on $[-1, 1]$, equipped with the norm $\|p\| := \sup_{x \in [-1, 1]} |p(x)|$ for $p \in P$, P is not finite dimensional.

Theorem 12.4

Let X be a vector space of dimension n over the field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, let $\{b_1, b_2, \dots, b_n\}$ be a basis of X . Consider the function $\|\cdot\| : X \rightarrow \mathbb{R}$ defined by:

$$\|\cdot\| : X \rightarrow \mathbb{R} \quad x = \sum_{j=1}^n x_j b_j \mapsto \sum_{j=1}^n |x_j|$$

Then $(X, \|\cdot\|)$ constitutes a Banach space, and all norms on X are equivalent to $\|\cdot\|$. That is, for norms $\|\cdot\|_X$ and $\|\cdot\|'_X$ on X , $\exists c, C > 0$ such that for all $x \in X$, we have:

$$c\|x\|'_X \leq \|x\|_X \leq C\|x\|'_X$$

Proof. For all Cauchy sequence (x_k) in X , we can write:

$$x_k = \sum_{j=1}^n x_{j,k} b_j \quad \text{where } x_{j,k} \in \mathbb{F}$$

We will show that $(x_{j,k})$ is a Cauchy sequence in \mathbb{F} . Here we write:

$$\|x_k - x_g\| = \sum_{i=1}^n |x_{i,k} - x_{i,g}| \geq |x_{j,k} - x_{j,g}| \quad \text{holds for all } i \leq n \quad (3.5)$$

From (3.5), taking large limit of k, g , it is easy to see that $(x_{j,k})$ is a Cauchy sequence. Then by the completeness of \mathbb{F} there exists $\xi_j \in \mathbb{F}$ such that $\lim_{k \rightarrow \infty} x_{j,k} = \xi_j$, and hence we can define:

$$\xi := \sum_{j=1}^n \xi_j b_j$$

Now we write:

$$\|x_k - \xi\| = \sum_{j=1}^n |x_{j,k} - \xi_j| \quad (3.6)$$

and here it is easy to show that (3.6) tends to 0 as k tends to ∞ , that is (x_k) converges to ξ . This shows that $(X, \|\cdot\|)$ is in fact Banach as we have verified its completeness. Now we will show that all norms on X are equivalent. We will show that there exists $m, M > 0$ such that $m\|x\| \leq \|x\|_X \leq M\|x\|$ for all $x \in X$ and an arbitrary norm $\|\cdot\|_X$ on X . Take $x \in X$ such that $x \neq 0$, we define $\xi = x/\|x\|$. Let $S \subseteq \mathbb{F}^n$ be defined by:

$$S := \left\{ (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{F}^n \left| \sum_{j=1}^n |\xi_j| = 1 \right. \right\}$$

which is a unit sphere defined by the norm $\|\cdot\|$. Clearly S is bounded, and it can be easily shown that S is closed, which implies that S is compact by the Heine-Borel Theorem. Now consider $f : S \rightarrow X$ defined by:

$$f : S \rightarrow X \quad (\xi_1, \xi_2, \dots, \xi_n) \mapsto \sum_{j=1}^n \xi_j b_j$$

where f is continuous because, for $(\xi_1, \xi_2, \dots, \xi_n), (\eta_1, \eta_2, \dots, \eta_n) \in \mathbb{F}$, we get:

$$\begin{aligned} \|f(\xi_1, \xi_2, \dots, \xi_n) - f(\eta_1, \eta_2, \dots, \eta_n)\|_X &= \left\| \sum_{j=1}^n (\xi_j - \eta_j) b_j \right\|_X \leq \sum_{j=1}^n \|(\xi_j - \eta_j) b_j\|_X \\ &= \sum_{j=1}^n |\xi_j - \eta_j| \cdot \|b_j\|_X \\ &\leq \max_{j \in \{1, 2, \dots, n\}} \|b_j\|_X \cdot \sum_{j=1}^n |\xi_j - \eta_j| \end{aligned}$$

Now consider the function F defined by:

$$F : S \rightarrow \mathbb{R} \quad (\xi_1, \xi_2, \dots, \xi_n) \mapsto \|f(\xi_1, \xi_2, \dots, \xi_n)\|_X$$

where we see that F must be continuous on the compact set S , hence F admits a maximum and a minimum on S . Let M denote the maximum of F on S and let m denote the minimum of F on S . Then for all $x \in X$, if $x \neq 0$, we can define:

$$\xi := \frac{x}{\|x\|} = \sum_{j=1}^n \xi_j b_j$$

such that we can write:

$$\|\xi\| = \sum_{j=1}^n |\xi_j| = 1 \quad m \leq \|\xi\|_X = \left\| \frac{x}{\|x\|} \right\|_X = \frac{\|x\|_X}{\|x\|} = \|\xi\|_X \leq M$$

hence we conclude that we have:

$$m\|x\| \leq \|x\|_X \leq M\|x\| \quad \text{for all } x \in X \text{ with } x \neq 0 \quad (3.7)$$

Taker another norm $\|\cdot\|_X$ on X , we can similarly fin m' and M' such that we can write:

$$m'\|x\| \leq \|x\|'_X \leq M'\|x\| \quad \text{for all } x \in X \text{ with } x \neq 0 \quad (3.8)$$

Combining (3.7) and (3.8) we get the followings for all $x \in X$ with $x \neq 0$:

$$\|x\|'_X \geq m'\|x\| \geq \frac{m'}{M}\|x\|_X \quad \|x\|'_X \leq M'\|x\| \leq \frac{M'}{m}\|x\|_X$$

this completes the proof of the theorem. \square

Definition 12.4.0.0.1

Let $(X, \|\cdot\|)$ be a separable normed space, let (e_n) be a sequence of points in X , (e_n) is called a Schauder topological basis for X provided that, for all $x \in X$, there exists unique sequence of scalar (α_n) such that we have:

$$\lim_{N \rightarrow \infty} \left\| x - \sum_{n=1}^N \alpha_n e_n \right\| = 0$$

The Little l^p Spaces

Definition 13.0.0.1

Let \mathbb{F} denote the field \mathbb{R} or the field \mathbb{C} . For $p \in [1, \infty]$, l^p is the space of sequences of points in \mathbb{F} such that all $(x_n) \in l^p$ satisfies the followings:

$$\begin{cases} \sum_{n=1}^{\infty} |x_n|^p < \infty & 1 \leq p < \infty \\ \sup_{n \geq 1} |x_n| < \infty & p = \infty \end{cases}$$

Example: For l^p space with $1 \leq p < \infty$, define (e_n) by letting $e_n = (0, 0, \dots, 1, 0, 0, \dots)$ where the term 1 appears only at the n -th term in e_n , (e_n) then forms a Schauder topological basis for the separable normed space l^p .

Note: Consider the following function:

$$\|\cdot\|_p : l^p \rightarrow \mathbb{R} \quad x \mapsto \begin{cases} (\sum_{n=1}^{\infty} |x_n|^p)^{1/p} & 1 \leq p < \infty \\ \sup_{n \geq 1} |x_n| & p = \infty \end{cases}$$

Here $\|\cdot\|_p$ defines a norm on the l^p space. In the rest of this text, if not mentioned otherwise, l^p is equipped with the norm defined by $\|\cdot\|_p$.

Lemma 13.0.1 (Young's Inequality)

For all $\alpha, \beta \in \mathbb{R}^+$, and for all conjugate pair (p, g) , we have:

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^g}{g}$$

Proof. For all $r, s \in \mathbb{R}$, and for all $\theta \in [0, 1]$, since \exp is a convex function, we observe that we have:

$$\exp(\theta r + (1 - \theta)s) \leq \theta \exp(r) + (1 - \theta) \exp(s)$$

let $\theta = 1/p$ and hence $1 - \theta = 1 - 1/p = 1/g$, let $r = p \log(\alpha)$ and $s = g \log(\beta)$, then we can write the following:

$$\begin{aligned} \exp\left(\frac{1}{p}p \log(\alpha) + \frac{1}{g}g \log(\beta)\right) &\leq \frac{1}{p} \exp(p \log(\alpha)) + \frac{1}{g} \exp(g \log(\beta)) \\ \exp(\log(\alpha) + \log(\beta)) &\leq \frac{\alpha^p}{p} + \frac{\beta^g}{g} \\ \alpha\beta &\leq \frac{\alpha^p}{p} + \frac{\beta^g}{g} \end{aligned}$$

This completes the proof. □

Lemma 13.0.2 (Holders Inequality)

Let $p, g \in \mathbb{R}$ with $p > 1$ that satisfies $1/p + 1/g = 1$, that is (p, g) is a conjugate pair. Let $(x_n) \in l^p$ and $(y_n) \in l^g$, then $z = (z_n)$ defined by $z_n = x_n y_n$ satisfies $z \in l^1$ and $\|z\|_1 \leq \|x\|_p \|y\|_g$. Note that for the case where $p = g = 2$, the result reduces to the Cauchy-Schwarz Inequality.

Proof. Denote $x = (x_n)$ and $y = (y_n)$, and define $\alpha = |x_j|/||x||_p$, $\beta = |y_j|/||y||_g$, then by Young's Inequality, we can write the following:

$$\frac{|x_j y_j|}{||x||_p ||y||_g} \leq \frac{1}{p} \left(\frac{|x_j|}{||x||_p} \right)^p + \frac{1}{g} \left(\frac{|y_j|}{||y||_g} \right)^g \quad \text{holds for all } j \geq 1 \quad (3.9)$$

here summing both sides of (3.9) over j , by the definition of norm on l^p space, we can write the following:

$$\frac{||xy||_1}{||x||_p ||y||_g} \leq \frac{||x||_p^p}{p ||x||_p^p} + \frac{||y||_g^g}{g ||y||_g^g} = \frac{1}{p} + \frac{1}{g} = 1$$

The result follows. \square

Lemma 13.0.3 (Minkovsiky Inequality)

For $p \in [1, \infty]$, $x, y \in l^p$ satisfy the following:

$$||x + y||_p \leq ||x||_p + ||y||_p \quad (3.10)$$

Lemma 13.0.4

For all $z \in [0, 1]$, there exists a binary expansion for z :

$$z = \sum_{j=1}^{\infty} \frac{z_j}{2^j} \quad \text{with } z_j \in \{0, 1\} \quad (3.11)$$

Theorem 13.1

For all $p \in [1, \infty]$, $(l^p, ||\cdot||_p)$ is a Banach space, and $(l^p, ||\cdot||_p)$ is separable for $1 \leq p < \infty$, but not separable for $p = \infty$.

Proof. First we will show the separability for $p \in [1, \infty)$. Consider the following:

$$A := \bigcup_{n=1}^{\infty} \{(y_j) \in l^p \mid y_j \in \mathbb{Q} \text{ for } 1 \leq j \leq n, \text{ and } y_j = 0 \text{ for } j > n\}$$

Note that A is countable because A is the union of finite Cartesian product of countable sets. Now we will show that, given $x \in l^p$, for all $\epsilon > 0$, there exists $y \in A$ such that $||x - y||_p < \epsilon$. Given $x = (x_n) \in l^p$, then there exists $N \in \mathbb{N}$ such that $\sum_{j=N+1}^{\infty} |x_j|^p < \frac{\epsilon^p}{19}$, since \mathbb{Q} is dense in \mathbb{R} , then there exists $y_j \in \mathbb{Q}$ such that we have:

$$|x_j - y_j| < \frac{\epsilon}{(19N)^{1/p}}$$

let $y = (y_1, y_2, \dots, y_N, 0, 0, \dots) \in A$, then we can write:

$$||x - y||_p = \left(\sum_{j=1}^N |x_j - y_j|^p + \sum_{j=N+1}^{\infty} |x_j|^p \right)^{1/p} < \epsilon$$

this shows that A is countable and dense, which implies that l^p is separable.

Now we will show that l^∞ is not separable. Consider the following set:

$$A := \{(z_j) \mid z_j \in \{0, 1\}\} \subseteq l^\infty$$

Here we define the following function:

$$f : A \rightarrow [0, 1] \quad (z_j) \mapsto \sum_{j=1}^{\infty} \frac{z_j}{2^j}$$

here by Lemma 13.0.4 we see that f is a surjection, so we know that A is uncountable. Take any dense set S in l^∞ . For all $z \in A$, since S is dense, there exists $\phi(z) \in S$ such that $\|z - \phi(z)\|_\infty < \frac{1}{3}$, where ϕ is defined in such way as a function. We will show that $\phi : A \rightarrow S$ is in fact injective. Take $z, \hat{z} \in A$ such that $z \neq \hat{z}$, then we know that $\|z - \hat{z}\|_\infty = 1$, hence we can write:

$$\begin{aligned} \|\phi(z) - \phi(\hat{z})\|_\infty &= \|-z + \phi(z) + z - \hat{z} + \hat{z} - \phi(\hat{z})\|_\infty \\ &\geq \|z - \hat{z}\|_\infty - \|\phi(z) - z\|_\infty - \|\phi(\hat{z}) - \hat{z}\|_\infty > \frac{1}{3} \end{aligned}$$

so we know that $\phi(z) \neq \phi(\hat{z})$, and hence ϕ is an injection. Concluding we see that S is not countable, and hence l^p is not separable. Now we will show that $(l^p, \|\cdot\|_p)$ is a Banach space. Consider Cauchy sequence (x_n) of points $x_n = (x_{n,j})_{j \geq 1}$ in l^p , we will show that (x_n) converges. First we consider $a_n = (x_{j,n})_{j \geq 1} \in l^p$, we claim that a_n is also Cauchy, because we have the following:

$$|x_{n,j} - x_{q,j}| \leq \|x_n - x_q\|_p$$

hence we see that a_n is Cauchy in \mathbb{R} or \mathbb{C} , so for all $n \in \mathbb{N}$ there exists $\xi_n = \lim_{j \rightarrow \infty} x_{j,n}$, and we now obtain a sequence $\xi = (\xi_n)$. If p is finite, we define:

$$S_k := \sum_{j=1}^k |\xi_j|^p = \sum_{j=1}^k \left| \lim_{k \rightarrow \infty} x_{k,j} \right|^p = \lim_{k \rightarrow \infty} \sum_{j=1}^k |x_{k,j}|^p \leq \lim_{k \rightarrow \infty} \|x_k\|_p^p \leq M$$

hence we know that (S_k) is monotone increasing and bounded above, so we have:

$$\lim_{k \rightarrow \infty} S_k = \sum_{k=1}^{\infty} |\xi_k|^p < \infty$$

this shows that $\xi \in l^p$, for $1 \leq p < \infty$. For $p = \infty$, since (x_n) is Cauchy, we can write:

$$S_k := \sup_{j \in \{1, 2, \dots, k\}} |\xi_j| = \sup_{j \in \{1, 2, \dots, k\}} \left| \lim_{n \rightarrow \infty} x_{n,j} \right| \leq M$$

so we see with a similar argument that $\xi \in l^p$. Finally, we will show that (x_n) converges to ξ . For $p < \infty$, by assumption, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|x_n - x_q\|_p < \epsilon$ for all $n, q \geq N$. Take any k , and take $n, q \geq N$, we write:

$$\sum_{j=1}^k |\xi_{j,n} - x_{j,q}|^p \leq \|x_n - x_q\|_p^p < \epsilon^p$$

Take q approach infinity, we can write:

$$S_k := \sum_{j=1}^k |x_{j,n} - \lim_{q \rightarrow \infty} x_{j,q}|^p = \sum_{j=1}^k |x_{j,n} - \xi_j|^p < \epsilon^p$$

Here (S_k) is again bounded above and monotone increasing. Taking the limit of (S_k) , we see that for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that we have:

$$\|x_n - \xi\|_p^p = \sum_{j=1}^{\infty} |x_{j,n} - \xi_j|^p < \epsilon^p$$

This shows that $\xi = \lim_{n \rightarrow \infty} x_n$ when $p < \infty$. For $p = \infty$, one can proceed similarly and get the result. The result of the theorem follows. \square

Theorem 13.2 (Jensen's Inequality)

Let $p, q \in \mathbb{R}$ satisfying $1 \leq p < q$, let $x = (x_n) \in l^p$, then we have:

$$\|x\|_q \leq \|x\|_p$$

That is, we know that $l^p \subseteq l^q$.

Proof. Consider $\xi = (\xi_n) \in l^p$ defined by $\xi_n = x_n / \|x\|_p$, hence we have $\|\xi\|_p = 1$. Here we have $|\xi_n| \leq 1$, since $p < q$, we can write:

$$\sum_{n=1}^{\infty} |\xi_n|^q \leq \sum_{n=1}^{\infty} |\xi_n|^p = \|\xi\|_p^p = 1$$

which implies that $\|\xi\|_q \leq 1$, hence we have:

$$\frac{\|x\|_q}{\|x\|_p} \leq 1$$

the results follow. \square

The Lebesgue L^p Spaces

Definition 14.0.0.1

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space where Ω is an open set in \mathbb{R}^n . $L^p(\Omega, \mu)$, or $L^p(\Omega)$, or L^p , is the space of equivalent classes of measurable functions $f : \Omega \rightarrow \mathbb{C}$ with equivalent relation defined by equality almost everywhere under measure μ , such that we have:¹

$$\begin{cases} \int_{\Omega} |f|^p d\mu < \infty & 1 \leq p < \infty \\ \exists M \geq 0 \text{ such that } |f(x)| \leq M \text{ almost everywhere} & p = \infty \end{cases}$$

Note: The following defines a norm on the space $L^p(\Omega, \mu)$:

$$\|\cdot\|_p : L^p(\Omega) \rightarrow \mathbb{R} \quad f \mapsto \begin{cases} \left(\int_{\Omega} |f|^p d\mu \right)^{1/p} & 1 \leq p < \infty \\ \inf\{M \geq 0 \mid |f(x)| \leq M \text{ almost everywhere}\} & p = \infty \end{cases}$$

Here $\|\cdot\|_p$ can also be denoted as $\|\cdot\|_{L^p}$ or $\|\cdot\|_{L^p(\Omega)}$.

Proposition 14.0.1 (Holder's Inequality)

Let (p, q) be a conjugate pair that satisfies $1/p + 1/q = 1$, let $f \in L^p$, and $g \in L^q$, then $h = fg \in L^1$, and we have $\|h\|_1 \leq \|f\|_p \|g\|_q$.

Proposition 14.0.2 (Minkovski's Inequality)

For $1 \leq p \leq \infty$, and all $f, g \in L^p$, we have the following holds:

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Theorem 14.1

For $1 \leq p \leq \infty$, $(L^p(\Omega), \|\cdot\|_p)$ is a Banach separable space if $1 \leq p < \infty$, and Banach but not separable space if $p = \infty$.

Remark: The proof of Proposition 15.0.1, Proposition 15.0.2, and Theorem 15.1 can be found on *Hunter and Nachtergaele*.

Lemma 14.1.1

Let $(X, \|\cdot\|)$ be a normed space. $(X, \|\cdot\|)$ is complete if and only if every absolutely convergent series converges. That is, X is complete if and only if $\sum_{n=1}^{\infty} x_n = x \in X$ for all sequence (x_n) of points in X that satisfies $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

Proof. For the \Rightarrow direction, we suppose X is complete. Consider (x_n) of points in X such that we have $\sum_{n=1}^{\infty} \|x_n\| < \infty$, here we define the sequence of partial sum (s_k) where $s_k := \sum_{n=1}^k x_n$. We will show that (s_k) is Cauchy. Here we can write the following, assuming that $k > q$:

$$\|s_k - s_q\| = \left\| \sum_{n=q+1}^k x_n \right\| \leq \sum_{n=q+1}^k \|x_n\| \quad (3.12)$$

but the last term can be made arbitrary small as we know that $\sum_{n=1}^{\infty} \|x_n\| < \infty$, so (s_k) converges to some $x \in X$ by the completeness of X . For the \Leftarrow direction, we suppose every absolutely convergent series converges. Take any Cauchy sequence (x_n) of points

¹The measure μ in such definition is usually referred to the Lebesgue measure on \mathbb{R}^n .

in X , one can find a subsequence $(x_{\sigma(k)})$ such that $\|x_{\sigma(k+1)} - x_{\sigma(k)}\| < 1/2^k$. Then we can write the following:

$$\sum_{k=1}^{\infty} \|x_{\sigma(k+1)} - x_{\sigma(k)}\| < 1$$

Hence by assumption, we have:

$$\lim_{N \rightarrow \infty} \left(\sum_{k=1}^N (x_{\sigma(k+1)} - x_{\sigma(k)}) \right) = \xi \in X$$

so the sequence $(x_{\sigma(N+1)} - x_{\sigma(1)})_{N \geq 1}$ converges to ξ , and hence we know that:

$$\lim_{k \rightarrow \infty} x_{\sigma(k)} = x_{\sigma(1)} + \xi$$

The subsequence of a Cauchy sequence converges only if the whole sequence converges, hence we conclude that (x_n) converges. This shows that X is complete. And that completes the proof. \square

Note: For the little l^p space, we have $l^p \subseteq l^q$ for $1 \leq p \leq q \leq \infty$, but in general this is not the case for L^p space. Suppose $\mu(\Omega) = \infty$, say $\Omega = (0, \infty)$ and μ is the Lebesgue measure, and let $1 \leq p < q \leq \infty$, let $b \in (1/q, 1/p)$, and define functions:

$$f : \Omega \rightarrow \mathbb{R} \quad x \mapsto x^{-b} \cdot \mathbb{I}_{(0,1)}(x) \quad \quad g : \Omega \rightarrow \mathbb{R} \quad x \mapsto x^{-b} \cdot \mathbb{I}_{(1,\infty)}(x)$$

Then we see that, $bp < 1$, so we have:

$$\int_0^{\infty} |f|^p = \int_0^1 x^{-bp} d\mu = \frac{x^{1-bp}}{1-bp} \Big|_{x=0}^{x=1} = \frac{1}{1-bp}$$

hence we see that $f \in L^p(\Omega)$, on the other hand, $bq > 1$, so we see that:

$$\int_0^{\infty} |f|^q = \infty$$

hence we see that $f \notin L^q(\Omega)$. This shows that $L^p(\Omega) \not\subseteq L^q(\Omega)$.

Conversely, we can write the following as we have $bq > 1$:

$$\int_0^{\infty} |g|^q = \int_1^{\infty} x^{-bq} = \frac{x^{1-bq}}{1-bq} \Big|_{x=1}^{x=\infty} = \frac{1}{bq-1}$$

and on the other hand, $bp < 1$, so we get:

$$\int_0^{\infty} |g|^p = \int_1^{\infty} x^{-bp} = \infty$$

hence we see that $L^q(\Omega) \not\subseteq L^p(\Omega)$.

Theorem 14.2

Consider the Lebesgue measure space $(\mathbb{R}^n, \mathcal{F}, \mu)$. Let Ω be open in \mathbb{R}^n with $\mu(\Omega) < \infty$. Then we have the followings hold:

1. $L^p(\Omega) \subseteq L^q(\Omega)$ for $1 \leq q < p \leq \infty$.
2. If $f \in L^\infty(\Omega)$, then $f \in L^p(\Omega)$ for all $1 \leq p < \infty$ and $\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p$.

Proof. For the proof of (1), we take $1 \leq q < p < \infty$. By Holder's Inequality, for $f \in L^q(\Omega)$, we can write the following:

$$\begin{aligned} \int_{\Omega} |f|^q \cdot 1 \, d\mu &\leq \left(\int_{\Omega} (|f|^q)^{p/q} \, d\mu \right)^{q/p} \cdot \left(\int_{\Omega} 1 \, d\mu \right)^{1-q/p} \\ &= \left(\int_{\Omega} |f|^p \, d\mu \right)^{q/p} (\mu(\Omega))^{1-q/p} \end{aligned} \quad (3.13)$$

Now take $1/q$ power for both sides of (3.13), we get:

$$\|f\|_q \leq \|f\|_p \cdot (\mu(\Omega))^{1/q-1/p}$$

Now consider $1 \leq 1 < p = \infty$. Note that we have $|f(x)| \leq \|f\|_\infty$ almost everywhere on Ω if $f \in L^\infty(\Omega)$, then we can write the following:

$$\int_{\Omega} |f|^q \, d\mu \leq \int_{\Omega} \|f\|_\infty^q \, d\mu = \mu(\Omega) \cdot \|f\|_\infty^q$$

hence it follows that we have:

$$\|f\|_q \leq (\mu(\Omega))^{1/q} \|f\|_\infty$$

Here we see that (1) follows. If $\|f\|_\infty = 0$, then the result of (2) follows.

Suppose $\|f\|_\infty > 0$, and let $\epsilon > 0$ be given, then we can define:

$$\Omega_\epsilon := \{x \in \Omega \mid |f(x)| \geq \|f\|_\infty - \epsilon\}$$

we see that $\mu(\Omega_\epsilon) > 0$, here we can write:

$$\int_{\Omega} |f|^q \, d\mu \geq \int_{\Omega_\epsilon} |f|^q \, d\mu \geq \int_{\Omega_\epsilon} (\|f\|_\infty - \epsilon)^q \, d\mu = \mu(\Omega_\epsilon) \cdot (\|f\|_\infty - \epsilon)^q$$

hence we get:

$$(\mu(\Omega_\epsilon))^{1/q} \cdot (\|f\|_\infty - \epsilon) \leq \|f\|_q \leq (\mu(\Omega))^{1/q} \|f\|_\infty$$

taking large limit of q , the result follows. □

4 | Continuous Functions

Uniform Continuity

Definition 15.0.0.0.1

Let (A, d) be a metric space. $C(A)$ is defined to be the space of continuous functions $f : A \rightarrow \mathbb{C}$. $BC(A) \subseteq C(A)$ is defined to be the normed space of bounded continuous function, whose norm is defined by the following:

$$\|\cdot\| : C(A) \rightarrow \mathbb{R} \quad f \mapsto \sup_{x \in A} |f(x)| \quad (4.1)$$

Theorem 15.1

Let (A, d) be a metric space. The space of bounded continuous functions $BC(A)$ equipped with the norm defined by (4.1) constitutes a Banach space.

Proof. It is not hard to check that $\|\cdot\|$ defined by (4.1) is a norm. For Cauchy sequence (f_n) in $BC(A)$, we will show that (f_n) converges to $f \in BC(A)$. $\forall \epsilon > 0$, there exists some $N \in \mathbb{N}$ such that for all $n, m \in \mathbb{N}$ with $n, m \geq N$, we have $\|f_n - f_m\| < \epsilon$. That is, for all $x \in A$, $(f_n(x))$ is a Cauchy sequence in \mathbb{C} because we can write:

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| < \epsilon \quad \forall n, m \geq N$$

By completeness of \mathbb{C} , we can write $\lim_{n \rightarrow \infty} f_n(x) = f(x)$, where we have defined:

$$f : A \rightarrow \mathbb{C} \quad x \mapsto \lim_{n \rightarrow \infty} f_n(x)$$

Now it suffices to show that the sequence $(\|f_n - f\|)$ converges to 0. Let $\epsilon > 0$ be given, note that, since (f_n) is Cauchy, then there exists some $N \in \mathbb{N}$ such that for $n, m \geq N$, and for arbitrary $x \in A$, we have:

$$|f_n(x) - f(x)| \leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| < \epsilon/19 + |f_m(x) - f(x)|$$

where the term $|f_m(x) - f(x)|$ can be minimized by using the pointwise convergence of $(f_m(x))$ to $f(x)$. That is, there exists $M(\epsilon, x) \in \mathbb{N}$ such that $\forall m \geq \max\{N, M(\epsilon, x)\}$, we have $|f_m(x) - f(x)| < \epsilon/19$. Then for all $n \geq N$, we conclude that $|f_n(x) - f(x)| < \epsilon$. Here ϵ is the upper bound on $\{|f_n(x) - f(x)| \mid x \in A, n \geq N\}$, hence we conclude that $\|f_n - f\| < \epsilon$ for all $n \geq N$. Now to show that $f \in BC(A)$, take $n = N$ and we obtain:

$$\|f\| \leq \|f_N\| + \|f - f_N\| < \epsilon + \|f_N\|$$

where we see that f is bounded. For all $x, y \in A$, for $\epsilon > 0$ we define N as above, we can write the following:

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f_N(y)| + |f(x) - f_N(x)| + |f_N(x) - f_N(y)| \\ &< 2\epsilon + |f_N(x) - f_N(y)| \end{aligned}$$

then by continuity of f_N , we see that f is continuous. Note that the convergence of (f_n) under the norm defined by (4.1) is called uniform convergence. \square

Theorem 15.2

Let (A, d) be a compact metric space. If (f_n) is a monotone sequence of functions $f_n : A \rightarrow \mathbb{R}$ in $C(A) = BC(A)$ converging pointwise to $f \in C(A)$, then $(\|f_n - f\|)_{n \geq 0}$ converges to 0.

Proof. WLOF, we assume that $(f_n(x))$ converges to $f(x)$ monotonically decreasing. Note that $g_n : A \rightarrow \mathbb{R} \quad x \mapsto f_n(x) - f(x)$ is a continuous positive function and $(g_n(x))$ converges to 0 monotonically decreasing. For all $n \in \mathbb{N}$, and for $\epsilon > 0$, we define $\mathcal{O}_n := g^{-1}((-\epsilon, \epsilon))$, which is an open set, and if $x \in \mathcal{O}_n$, $g_n(x) \in [0, \epsilon)$, and $0 \leq g_{n+1}(x) \leq g_n(x) \in [0, \epsilon)$, hence we have $\mathcal{O}_1 \subseteq \mathcal{O}_2 \subseteq \mathcal{O}_3 \subseteq \dots \subseteq A$. For all $x \in A$, since $(g_n(x))$ converges to 0 monotonically decreasing, then there exists $N \in \mathbb{N}$ such that $g_N(x) \in [0, \epsilon)$, then we have:

$$0 \leq g_n(x) \leq g_N(x) \in [0, \epsilon) \quad \forall n \geq N$$

that is for all $x \in A$, there exists N such that $x \in \mathcal{O}_N \subseteq \mathcal{O}_{N+1} \subseteq A$, here we get an open cover $\{\mathcal{O}_n \mid n \in \mathbb{N}\}$ of A . By the compactness of A , A admits a finite subcover, so:

$$A \subseteq \bigcup_{j \in J} \mathcal{O}_j \quad \text{where } J \text{ is a finite subset of } \mathbb{N}$$

take $M = \max\{J\}$, then we can write $A \subseteq \mathcal{O}_M$ as we know that (\mathcal{O}_n) is a sequence of nested sets. That is, we now can write $A = \mathcal{O}_M$ because $\mathcal{O}_M \subseteq A$ by definition. Then we know that $g_M(x) \in [0, \epsilon)$ for all $x \in A$, Then for all $n \geq M$, we have $0 \leq g_n(x) \leq g_M(x) < \epsilon$, that is we have $\|g_n\| < \epsilon$ for all $n \geq M$. The result of theorem follows. \square

Remark: Consider $BC([0, 1])$, $f_n : [0, 1] \rightarrow \mathbb{R} \quad x \mapsto \cos(nx)$, where we see that (f_n) does not converge pointwise and it has no convergent subsequence, but $\|f_n\| = 1$. That is, a bounded sequence of functions in the space $BC([0, 1])$ does not necessarily have a convergent subsequence. While in finite dimensional space \mathbb{R}^n , any bounded sequence has convergent subsequence according to the Bolzano–Weierstrass Theorem.

Equicontinuity

Definition 16.0.0.0.1

Let (A, d) be a metric space, let $\mathcal{F} \subseteq C(A)$ be a family of continuous functions. \mathcal{F} is said to be equicontinuous provided that for all $\epsilon > 0$, given $x \in A$, there exists $\delta = \delta(\epsilon, x)$ such that if we have $d(x, y) < \delta$ for $y \in A$, then $|f(x) - f(y)| < \epsilon$ for all $f \in \mathcal{F}$.

Definition 16.0.0.0.2

Let (A, d) be a metric space, let $\mathcal{F} \subseteq C(A)$ be a family of continuous functions. \mathcal{F} is said to be uniformly equicontinuous provided that for all $\epsilon > 0$, there exists $\delta = \delta(\epsilon)$ such that if we have $d(x, y) < \delta$ for $x, y \in A$, then $|f(x) - f(y)| < \epsilon$ for all $f \in \mathcal{F}$.

Theorem 16.1

Let (A, d) be a compact metric space and let $\mathcal{F} \subseteq C(A)$ be a equicontinuous family, then \mathcal{F} is uniformly equicontinuous.

Proof. First we can use the equicontinuity of \mathcal{F} to get a finite cover of A . Let $\epsilon > 0$ be given, for all $x \in A$, there exists $\delta_x > 0$ such that for all $y \in \mathcal{F}$ and all $y \in A$, if $d(x, y) < \delta_x$, then $|f(x) - f(y)| < \epsilon$. Now we can cover A by the following:

$$A \subseteq \bigcup_{x \in A} B_{\delta_x/2}(x)$$

and by the compactness of A , we can write the following for some $J \in \mathbb{N}$:

$$A \subseteq \bigcup_{j=1}^J B_{\delta_{x_j}/2}(x_j)$$

Here we define:

$$\delta := \min_{1 \leq j \leq J} \delta_{x_j}/2$$

Then for all $x \in A$, there exists $j \in \mathbb{N}$ with $j \leq J$ such that $x \in B_{\delta_{x_j}/2}(x_j)$. Take $y \in A \cap B_\delta(x)$, we see that we can write:

$$d(y, x_j) \leq d(x, x_j) + d(x, y) \leq \frac{\delta_{x_j}}{2} + \delta < \frac{\delta_{x_j}}{2} + \frac{\delta_{x_j}}{2} \leq \delta_{x_j}$$

now we see that $y \in B_{\delta_{x_j}}(x_j)$, and so we can write:

$$|f(y) - f(x_j)| < \epsilon$$

Note also we have $|f(x) - f(x_j)| < \epsilon$, then here we conclude that, for all $x \in A$, for all $y \in A \cap B_\delta(x)$, we can write:

$$|f(y) - f(x)| \leq |f(y) - f(x_j)| + |f(x) - f(x_j)| < 2\epsilon$$

since ϵ is arbitrary, here we get the uniform equicontinuity of \mathcal{F} . □

Theorem 16.2 (Arzela-Ascoli Theorem)

Let (A, d) be a compact metric space, let $\mathcal{F} \subseteq C(A)$ be a family that is equicontinuous and for all $x \in A$, $\sup_{f \in \mathcal{F}} |f(x)| < \infty$.¹ Then \mathcal{F} is compact.

Proof. Since A is compact, then $BC(A) = C(A)$, then by Theorem 15.1, \mathcal{F} is complete, hence it suffices to show that \mathcal{F} is totally bounded. Let $\epsilon > 0$ be given, by Theorem 16.1, we have:

$$\exists \delta = \delta(\epsilon) \text{ s.t. } \forall x, y \in A, d(x, y) < \delta, \text{ then } |f(x) - f(y)| < \epsilon, \forall f \in \mathcal{F} \quad (4.2)$$

that is \mathcal{F} is uniformly equicontinuous. Note that we can write:

$$A \subseteq \bigcup_{x \in A} B_\delta(x)$$

and by compactness of A , we can write:

$$A \subseteq \bigcup_{j=1}^J B_\delta(x_j) \quad (4.3)$$

Now we will construct a cover of \mathcal{F} , first define the following:

$$M := \max_{1 \leq j \leq J} \sup_{f \in \mathcal{F}} |f(x_j)|$$

Since $[-M, M]$ is compact, one can choose $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{R}$ such that we have:

$$[-M, M] \subseteq \bigcup_{j=1}^m (\alpha_j - \epsilon, \alpha_j + \epsilon) \quad (4.4)$$

We can define:

$$\Gamma := (\text{the set of constant maps between } \{x_j \mid 1 \leq j \leq J\} \text{ and } \{\alpha_j \mid 1 \leq j \leq J\})$$

Note that Γ contains m^J maps. For any $\gamma \in \Gamma$, we can define:

$$\mathcal{F}_\gamma := \{f \in \mathcal{F} \mid f(x_j) \in (\gamma(x_j) - \epsilon, \gamma(x_j) + \epsilon), 1 \leq j \leq J\} \quad (4.5)$$

and here we claim that $\mathcal{F} = \bigcup_{\gamma \in \Gamma} \mathcal{F}_\gamma$. Note that it is clear that we have $\bigcup_{\gamma \in \Gamma} \mathcal{F}_\gamma \subseteq \mathcal{F}$. Here it suffices to show that $\mathcal{F} \subseteq \bigcup_{\gamma \in \Gamma} \mathcal{F}_\gamma$. For all $f \in \mathcal{F}$, and for any $1 \leq j \leq J$, there exists $l = l(x_j) \in \{1, 2, \dots, m\}$ such that $f(x_j) \in (\alpha_l - \epsilon, \alpha_l + \epsilon)$. Then, with $\gamma \in \Gamma$ that satisfies $\gamma(x_j) = \alpha_{l(x_j)}$, we see that $f \in \mathcal{F}_\gamma$. Hence $\mathcal{F} = \bigcup_{\gamma \in \Gamma} \mathcal{F}_\gamma$. We claim that $\text{diam}(\mathcal{F}_\gamma) > 4\epsilon$, in which case we have \mathcal{F} covered by m^J balls of radius 2ϵ , and hence we see that \mathcal{F} is totally bounded. For given γ , note that for all $x \in A$, by (4.3), we know that $\exists j \in \{1, 2, \dots, J\}$ such that $x \in B_\delta(x_j)$, take $f, g \in \mathcal{F}_\gamma$, we can write:

$$\begin{aligned} |f(x) - g(x)| &\leq |f(x) - f(x_j)| + |g(x) - g(x_j)| + |g(x_j) - \gamma(x_j)| + |f(x_j) - \gamma(x_j)| \\ &< \epsilon + \epsilon + \epsilon + \epsilon = 4\epsilon \end{aligned}$$

The result of this theorem follows. □

¹Here we say \mathcal{F} is equicontinuous and pointwise bounded

Lipchitz Functions

Definition 17.0.0.0.1

Let (X, d) be a metric space, a function $f : X \rightarrow \mathbb{C}$ is said to be Lipschitz continuous provided that there exists $C > 0$ such that $|f(x) - f(y)| \leq C \cdot d(x, y)$ for all $x, y \in X$.

Definition 17.0.0.0.2

Let (X, d) be a metric space, and let $f : X \rightarrow \mathbb{C}$ be a Lipschitz continuous function, we define the Lipchitz constant of f to be the following:

$$\begin{aligned} \text{Lip}(f) &:= \inf\{C \mid |f(x) - f(y)| \leq C \cdot d(x, y), \forall x, y \in X\} \\ &= \sup\left\{\frac{|f(x) - f(y)|}{d(x, y)} \mid x, y \in X, \text{ with } x \neq y\right\} \end{aligned}$$

Lemma 17.0.1

Any Lipchitz continuous function is uniformly continuous.

Proof. “Just think about it, it is not that difficult.” □

Lemma 17.0.2

Let (A, d) be a compact metric space. For $M > 0$, we define:

$$\mathcal{F}_M := \{f \mid f \text{ is Lipschitz on } A, \text{ Lip}(f) \leq M\} \quad (4.6)$$

Then \mathcal{F}_M is uniformly equicontinuous and closed.

Proof. Let $\epsilon > 0$ be given, and let $f \in \mathcal{F}_M$, then for $\delta = \epsilon/M$, and $x, y \in X$ that satisfies $d(x, y) < \delta$, we can write:

$$|f(x) - f(y)| \leq \text{Lip}(f) \cdot d(x, y) < M \cdot \frac{\epsilon}{M} = \epsilon$$

This shows that \mathcal{F}_M is uniformly equicontinuous. Now we will show that \mathcal{F}_M is closed. Consider a Cauchy sequence of functions (f_n) in \mathcal{F}_M . Note that we have $(f_n) \rightarrow f$ for some $f \in BC(A)$ because $BC(A)$ is a Banach space and $C(A) = BC(A)$ for compact A . Now we can write, for $x, y \in X$:

$$\frac{|f(x) - f(y)|}{d(x, y)} = \frac{|\lim_{n \rightarrow \infty} (f_n(x) - f_n(y))|}{d(x, y)} = \lim_{n \rightarrow \infty} \frac{|f_n(x) - f_n(y)|}{d(x, y)} \leq M$$

and this shows that $f \in \mathcal{F}_M$, which completes the proof. □

By Arzela-Ascoli Theorem, any closed and bounded set in \mathcal{F}_M defined by (4.6) is compact. Now one can pick $x_0 \in A$ and define:

$$\mathcal{B}_M := \{f \in \mathcal{F}_M \mid f(x_0) = 0\}$$

for $f \in \mathcal{B}_M$, we see that we have the following for all $x \in A$:

$$|f(x)| = |f(x) - f(x_0)| \leq M \cdot d(x, x_0) < \infty$$

in which case we see that $\|f\| < \infty$, and hence \mathcal{B}_M is bounded. Then \mathcal{B}_M is compact by the Arzela-Ascoli Theorem. Here we note that \mathcal{B}_M is also closed, because for convergent sequence (f_n) in \mathcal{B}_M , we know that (f_n) converges to some $f \in \mathcal{F}_M$ by Lemma 17.0.2, and we see that $f(x_0) = 0$, hence $f \in \mathcal{B}_M$.

Definition 17.0.2.0.1

Let (X, d) be a metric space, a function $f : X \rightarrow X$ is said to be a contraction provided that there exists $K \in (0, 1)$ such that we have $d(f(x), f(y)) \leq K \cdot d(x, y)$ for all $x, y \in X$.

Note: Here we see that contraction maps are uniformly continuous maps.

Theorem 17.1 (Banach Fixed Point Theorem ²)

Let $(X, \|\cdot\|)$ be a Banach space, and let $f : X \rightarrow X$ be a contraction with $k \in (0, 1)$ such that $\|f(x) - f(y)\| \leq k \cdot \|x - y\|$. Then we have the followings hold:

1. The function f has a unique fixed point $x \in X$.
2. For all $x_0 \in X$, and sequence $(x_n)_{n \geq 1}$ defined by $x_n = f(x_{n-1})$, the sequence (x_n) converges to the fixed point x of f .
3. The rate of convergence for (x_n) in (2) defined by $C = \|f(x_0) - x_0\|/(1 - k)$ satisfies the following:

$$\|x - x_n\| \leq C \cdot k^n$$

Proof. First we consider a sequence (x_n) defined by (2) in X , we can write the following:

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq k \cdot \|x_n - x_{n-1}\| \leq \cdots \leq k^n \|x_1 - x_0\|$$

Now we take $m, n \in \mathbb{N}$ with $m > n > 0$, we can write:

$$\begin{aligned} \|x_m - x_n\| &= \sum_{j=n}^{m-1} \|x_{j+1} - x_j\| \\ &\leq \sum_{j=n}^{m-1} k^j \|x_1 - x_0\| \\ &= k^n \|x_1 - x_0\| \sum_{j=0}^{m-n-1} k^j \\ &\leq k^n \|x_1 - x_0\| \sum_{j=0}^{\infty} k^j \\ &= k^n \|x_1 - x_0\| \frac{1}{1 - k} \end{aligned}$$

Hence we can write:

$$\|x_m - x_n\| \leq \frac{k^n \|x_1 - x_0\|}{1 - k} \quad (4.7)$$

where we see that (x_n) is Cauchy in the Banach space X , hence (x_n) converges to some $x \in X$. Now by the continuity of f , we see that $(f(x_n))$ also converges to $f(x)$, and hence $f(x) = x$. This shows that f has a fixed point in X . Now we will show that the fixed point of f is unique. Suppose $x = f(x)$ and $x' = f(x')$ for some $x, x' \in X$, then we can write the followings:

$$\|x - x'\| = \|f(x) - f(x')\| \leq k \|x - x'\|$$

rearranging we have:

$$(1 - k) \|x - x'\| \leq 0 \quad \Rightarrow \quad \|x - x'\| = 0 \quad \Rightarrow \quad x = x'$$

²Banach Fixed Point Theorem is also called the Contraction Mapping Theorem

Finally, we take m to be large enough, and through (4.7) we get the following:

$$\|x - x_n\| = \lim_{m \rightarrow \infty} \|x_m - x_n\| \leq \frac{k^n \|x_1 - x_0\|}{1 - k}$$

The result of the theorem follows. □

5 | Bounded Linear Operators

The Bounded Operators

Definition 18.0.0.1

Let X and Y be vector spaces over the field \mathbb{K} . A function $\Lambda : X \rightarrow Y$ is called a linear operator from X to Y provided that it satisfies $\Lambda(c_1x_1 + c_2x_2) = c_1\Lambda(x_1) + c_2\Lambda(x_2)$ for all $x_1, x_2 \in X$ and all $c_1, c_2 \in K$.

Theorem 18.1

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be normed spaces, let $\Lambda : X \rightarrow Y$ be a linear operator, then the followings are equivalent:

1. Λ is continuous on X .
2. Λ is continuous at $\vec{0} \in X$.
3. There exists $c > 0$ such that $\|\Lambda x\|_y < C\|x\|_x$ for all $x \in X$
4. If $U \subseteq X$ is bounded, then $\Lambda(U) \subseteq Y$ is also bounded

If one of the conditions holds for Λ , we say that Λ is bounded.

Proof. It is immediate that we have (1) implies (2). Now we will show that (2) implies (3). Let $x \in X$ with $\|x\| \neq 0$, we can write:

$$\|\Lambda x\|_y = \|x\|_x \cdot \left\| \Lambda \frac{x}{\|x\|_x} \right\|_y \quad (5.1)$$

Pick $\epsilon = 1$, then there exists $\delta > 0$ such that we have:

$$\|\Lambda x\|_y \leq 1 \quad \text{for all } x \in \overline{B_\delta(0)}$$

write $\xi = x/\delta \in \overline{B_1(0)}$ for all $x \in \overline{B_\delta(0)}$, then we have:

$$\|\Lambda \xi\|_y = \left\| \Lambda \frac{x}{\delta} \right\|_y = \frac{1}{\delta} \|\Lambda x\|_y \leq \frac{1}{\delta}$$

Hence now (5.1) reads the following:

$$\|\Lambda x\|_y = \frac{1}{\delta} \cdot \|x\|_x$$

Now we will show that (3) implies (4). Take U be a bounded set, then we must have $U \subseteq B_R(0)$ for some $R \geq 0$. Then we can take $x \in U$ and write the following:

$$\|\Lambda x\|_y \leq C\|x\|_x \leq RC$$

Finally, we will show that (4) implies (1). Here we take $U = \overline{B_1(0)} \subseteq X$, then by assumption there exists $R > 0$ such that we have $\Lambda(U) \subseteq B(0, R)$. Consider $x_0, x \in X$ such that $x_0 \neq x$, then we have:

$$\|\Lambda x - \Lambda x_0\|_y = \|\Lambda(x - x_0)\|_y = \|x - x_0\|_x \left\| \Lambda \frac{x - x_0}{\|x - x_0\|_x} \right\|_y \leq R \|x - x_0\|_x$$

so here we see that Λ is a Lipschitz function, which implies that Λ is a continuous function, and hence complete the proof. \square

Definition 18.1.0.0.1

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be normed spaces.

$\mathcal{B}(X, Y)$ is the set of bounded linear operators from X to Y .

Theorem 18.2

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be normed spaces. $\mathcal{B}(X, Y)$ is a normed space when equipped with the norm:

$$\|\cdot\|_{\mathcal{B}} : \mathcal{B}(X, Y) \rightarrow \mathbb{R} \quad \Lambda \mapsto \sup \left\{ \frac{\|\Lambda x\|_y}{\|x\|_x} \mid x \in X, x \neq 0 \right\}$$

Here $(\mathcal{B}(X, Y), \|\cdot\|_{\mathcal{B}})$ is a Banach space when Y is complete.¹

Proof. Here we note that we have:

$$\sup \left\{ \frac{\|\Lambda x\|_y}{\|x\|_x} \mid x \in X, x \neq 0 \right\} = \sup \{ \|\Lambda x\|_y \mid x \in X, \|x\|_x \leq 1 \}$$

Here we will show the triangle inequality under $\|\cdot\|_{\mathcal{B}}$. For $\Lambda_1, \Lambda_2 \in \mathcal{B}(X, Y)$, for $\|x\|_x = 1$, we can write the following:

$$\|(\Lambda_1 + \Lambda_2)x\|_y \leq \|\Lambda_1 x\|_y + \|\Lambda_2 x\|_y = \|\Lambda_1\|_{\mathcal{B}} + \|\Lambda_2\|_{\mathcal{B}}$$

taking the supremum of both sides and the result follows. Here we see that the space $(\mathcal{B}(X, Y), \|\cdot\|_{\mathcal{B}})$ is indeed a normed space. Now suppose $(Y, \|\cdot\|_y)$ is a complete space. Let (Λ_n) be a Cauchy sequence in $(\mathcal{B}(X, Y), \|\cdot\|_{\mathcal{B}})$. For all $x \in X$, then we can write the following for $n, m \geq 1$:

$$\|\Lambda_n x - \Lambda_m x\|_y = \|(\Lambda_n - \Lambda_m)x\|_y \leq \|\Lambda_n - \Lambda_m\|_{\mathcal{B}} \cdot \|x\|_x$$

where we see that the term $\|\Lambda_n - \Lambda_m\|_{\mathcal{B}}$ can be made arbitrarily small by the fact that (Λ_n) is a Cauchy sequence in $\mathcal{B}(X, Y)$. Hence we know that $(\Lambda_n x)_{n \geq 1}$ is a Cauchy sequence in Y , which must converges. Here we define:

$$\Lambda : X \rightarrow Y \quad x \mapsto \lim_{n \rightarrow \infty} \Lambda_n x$$

Here we will check that Λ is an element in $\mathcal{B}(X, Y)$, where we can write the following for all $\alpha_1, \alpha_2 \in \mathbb{K}$ and $x_1, x_2 \in X$:

$$\Lambda(\alpha_1 x_1 + \alpha_2 x_2) = \lim_{n \rightarrow \infty} \Lambda_n(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \lim_{n \rightarrow \infty} \Lambda_n x_1 + \alpha_2 \lim_{n \rightarrow \infty} \Lambda_n x_2 = \alpha_1 \Lambda x_1 + \alpha_2 \Lambda x_2$$

¹For notation, we denote the norm $\|\cdot\|_{\mathcal{B}}$ as $\|\cdot\|$ in obvious context.

Now it suffices to show that (Λ_n) converges to Λ . Here we can write, for all $\epsilon > 0$, there exists $N > 0$ such that for all $n, m > N$, we have $\|\Lambda_n - \Lambda_m\|_{\mathcal{B}} \leq \epsilon$, then for all $x \in X$, we can write the following:

$$\|\Lambda_n x - \Lambda_m x\|_y \leq \|\Lambda_n - \Lambda_m\|_{\mathcal{B}} \cdot \|x\|_x \leq \epsilon \|x\|_x \quad \text{for all } n, m > N$$

taking the large limit of m , we get the following:

$$\|\Lambda_n x - \Lambda x\|_y \leq \epsilon \|x\|_x \quad \text{for all } n > N$$

and hence we have:

$$\|\Lambda_n - \Lambda\|_{\mathcal{B}} \leq \epsilon \quad \text{for all } n > N$$

that is, we conclude that $\lim_{n \rightarrow \infty} \Lambda_n = \Lambda$, which completes the proof. \square

Theorem 18.3 (Banach-Schauder Open Mapping Theorem)

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach normed spaces.

If $\Lambda \in \mathcal{B}(X, Y)$ is surjective, then $\Lambda(U) \subseteq Y$ is open for all open set $U \subseteq X$.

Proof. “Just think about it, it is not that difficult.” \square

Corollary 18.3.1 (Bounded Inverse Theorem)

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach normed spaces.

If $\Lambda \in \mathcal{B}(X, Y)$ is bijective, then $\Lambda^{-1} \in \mathcal{B}(Y, X)$.

Proof. It is trivial to check that Λ^{-1} is a linear operator. We will show that Λ^{-1} is continuous at $0 \in Y$. Let $\epsilon > 0$ be given, consider $U = B_\epsilon(0) \subseteq Y$. Since Λ is bijective, then by Banach-Schauder Open Mapping Theorem, we know that $\Lambda(U) \subseteq X$ is open, and we know that $0 \in \Lambda(U)$, so there exists $\delta > 0$ such that, by the bijectivity of Λ , we have the following holds:

$$\Lambda(U) \subseteq B_\delta(0) \subseteq \Lambda(U)$$

then we can write the following:

$$\Lambda^{-1}(B_\delta(0)) \subseteq U = B_\epsilon(x)$$

which shows that Λ^{-1} is continuous at 0. \square

Definition 18.3.1.0.1

Let X, Y be vector spaces, and let $\Lambda : X \rightarrow Y$ be a linear operator.

$\ker(\Lambda) := \{x \in X \mid \Lambda x = 0\}$ is a subspace of X , and $\dim(\ker(\Lambda)) :=$ nullity of Λ .

$\text{im}(\Lambda) := \{y \in Y \mid \exists x \in X, \text{ s.t. } \Lambda x = y\}$ is a subspace of Y .

Note: For vector spaces X, Y , if $\Lambda : X \rightarrow Y$ is bounded and linear, then $\ker(\Lambda)$ is closed.

Remark: For finite dimensional vector space X, Y , and linear operator $\Lambda : X \rightarrow Y$, if we have $\ker(\Lambda) = \{0\}$, then $\text{im}(\Lambda) = Y$, but for infinite dimensional space, this result is not true.

Example: Consider S to be the right shift operator defined on the l^∞ space, where $S((x_1, x_2, x_3, \dots)) = (0, x_1, x_2, x_3, \dots)$. The kernel of S is $\{0\}$, but the image of S is not l^∞ , and here l^∞ is a space of infinite dimension.

Theorem 18.4

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach normed spaces, and let $\Lambda \in \mathcal{B}(X, Y)$.

The following statements are equivalent:

1. $\exists C > 0$ such that $\|\Lambda x\|_y \geq C \cdot \|x\|_x$ for all $x \in X$.
2. Λ has closed image and $\ker(\Lambda) = \{0\}$.

Proof. First we will show that (1) implies (2). Let $x \in \ker(\Lambda)$, then we know that $\Lambda x = 0$, here we can also write:

$$\|x\| \leq \frac{1}{C} \cdot \|\Lambda x\|_y = 0 \quad \Rightarrow \quad x = 0$$

To show that the image of Λ is closed, consider a sequence (y_n) in the image of Λ that converges to some $y \in Y$. We will show that y belongs to the image of Λ . For $n \in \mathbb{N}$, since y_n belongs to $\text{im}(\Lambda)$, then there exists $x_n \in X$ such that $\Lambda x_n = y_n$, then we get a sequence (x_n) in X . First we will show that (x_n) is Cauchy. Note that for $n, m \in \mathbb{N}$, we have the following holds:

$$\|x_m - x_n\|_X \leq \frac{1}{C} \cdot \|\Lambda x_m - \Lambda x_n\|_y = \frac{1}{C} \cdot \|y_m - y_n\|_y$$

from here we see that (x_n) is Cauchy as (y_n) is Cauchy, then we know that (x_n) converges to some $x \in X$ because X is a Banach space. Now we can write:

$$y = \lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} \Lambda x_n = \Lambda \left(\lim_{n \rightarrow \infty} x_n \right) = \Lambda x$$

this completes the proof of this direction. Now we will show (2) implies (1). Here we let $V = \text{im}(\Lambda)$, where V is closed in Y , and hence $(V, \|\cdot\|_y)$ constitutes a Banach space, and $\Lambda : X \rightarrow V$ is a surjection. Since the kernel of Λ is trivial, then we must have $\Lambda : X \rightarrow V$ being a bijection. By Corollary 18.3.1, we know that $\Lambda^{-1} : V \rightarrow X$ belongs to $\mathcal{B}(V, X)$. Let $K := \|\Lambda^{-1}\|_{\mathcal{B}}$. Take $x \in X$, then there exists a unique $y \in V$ such that $\Lambda x = y$, and hence $x = \Lambda^{-1}y$, then we can write the following:

$$\|x\|_x = \|\Lambda^{-1}y\|_x \leq \|\Lambda^{-1}\|_{\mathcal{B}} \cdot \|y\|_y = K \cdot \|y\|_y = K \cdot \|\Lambda x\|_y$$

the result of this theorem follows. □

Uniform Boundedness and Convergence

Theorem 19.1 (Banach-Steinhaus Uniform Bounded Principal)

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach spaces and let $\mathcal{F} \subseteq \mathcal{B}(X, Y)$.

Then there are two possibilities:

1. \mathcal{F} is uniformly bounded and that we have $\sup\{\|\Lambda\| \mid \Lambda \in \mathcal{F}\} < \infty$.
2. \exists dense set $S \subseteq X$ such that $\sup\{\|\Lambda x\|_y \mid \Lambda \in \mathcal{F}\} = \infty$ for all $x \in S$.

Proof. For all $m \in \mathbb{N}$, we first define:

$$S_m := \{x \in X \mid \|\Lambda x\|_y > m, \text{ for some } \Lambda \in \mathcal{F}\}$$

Here we will show that S_n is open for all $n \in \mathbb{N}$. Let $x \in S_n$, we see that there exists $\Lambda \in \mathcal{F}$ such that $\|\Lambda x\|_y = n + \delta > n$. Take $\epsilon \in (0, \delta/\|\Lambda\|_{\mathcal{B}})$, then for all $t \in X$ with $\|t\|_x < \epsilon$, writing $\xi = x + t \in B_\epsilon(x)$, we get the following:

$$\|\Lambda \xi\|_y = \|\Lambda x + \Lambda t\|_y \geq \|\Lambda x\|_y - \|\Lambda t\|_y = n + \delta - \|\Lambda t\|_y > n + \delta - \|\Lambda\|_{\mathcal{B}} \|t\|_x = n$$

this shows that $\xi \in S_n$. Now we have two cases:

- A. There exists $N \in \mathbb{N}$ such that S_N is not dense in X , then there exists at least one $\xi \in X$ with some $r > 0$ such that $B_r(\xi) \cap S_N = \emptyset$. Now for all $x \in B_r(\xi)$, we have $x \notin S_N$, that is $\|\Lambda x\|_y \leq N$ for all $\Lambda \in \mathcal{F}$. Take $t \in X$ with the property that $\|t\|_x < r$, then for all $\Lambda \in \mathcal{F}$, we have:

$$\|\Lambda t\|_y = \|\Lambda(\xi + t) - \Lambda \xi\|_y \leq \|\Lambda(\xi + t)\|_y + \|\Lambda \xi\|_y \leq 2N \quad (5.2)$$

Now take $x \in X$, for all $\Lambda \in \mathcal{F}$, via (5.2) we can write the following:

$$\|\Lambda x\|_y = \left\| \Lambda \left(\frac{xr}{\|x\|_x} \right) \right\|_y \cdot \frac{\|x\|_x}{r} \leq \frac{2N}{r} \cdot \|x\|_x$$

that is, we have $\|\Lambda\|_{\mathcal{B}} \leq 2N/r$, for all $\Lambda \in \mathcal{F}$, this shows that case (A) corresponds to statement (1) in the Theorem.

- B. In this case, we suppose S_n is dense in X for all n , then we write $\overline{S_n} = X$ for all $n \in \mathbb{N}$. Hence we set a sequence of open dense set S_n , and hence $S = \bigcap_{n=1}^{\infty} S_n$ is also dense in X . Then for all $x \in S$, and all $n \in \mathbb{N}$, there exists some $\Lambda \in \mathcal{F}$ such that $\|\Lambda x\|_y > n$, so for all $x \in S$, we have:

$$\sup\{\|\Lambda x\|_y \mid \Lambda \in \mathcal{F}\} = \infty$$

This shows that (B) corresponds to statement (2) in the theorem.

The result of the theorem follows. □

Definition 19.1.0.0.1

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach spaces. $\mathcal{F} \subseteq \mathcal{B}(X, Y)$ is said to be pointwise bounded on $U \subseteq X$ provided that for all $x \in U$, we have $\sup\{\|\Lambda x\|_y \mid \Lambda \in \mathcal{F}\} < \infty$.

Corollary 19.1.1

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach spaces, if $\mathcal{F} \subseteq \mathcal{B}(X, Y)$ is pointwise bounded on $B_1(0) \subseteq X$, then \mathcal{F} is uniformly bounded.

Proof. Case (2) in Theorem 19.1 cannot hold under the assumption of this Corollary, the result follows. □

Theorem 19.2 (The continuity of pointwise limit for bounded linear map)

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be Banach spaces, and let (Λ_n) be a sequence in $\mathcal{B}(X, Y)$ that converges pointwise, then the function Λ defined by the following belongs to $\mathcal{B}(X, Y)$:

$$\Lambda : X \rightarrow Y \quad x \mapsto \lim_{n \rightarrow \infty} \Lambda_n x$$

Proof. First we will show that (Λ_n) is uniformly bounded. We see that $\mathcal{F} = \{\Lambda_n \mid n \in \mathbb{N}\}$ converges pointwise in $B_1(0) \subseteq X$, then we have (y_n) , where $y_n := \Lambda_n x$ for some $x \in B_1(0)$, being bounded in Y , so by Corollary 19.1.1, we see that \mathcal{F} is uniformly bounded, that is, there exists $C > 0$ such that $\|\Lambda_n\|_{\mathcal{B}} \leq C$ for all $n \in \mathbb{N}$. Now we will show that Λ is a linear map, let $x, x' \in X$ and let α, α' be scalars, here we can write:

$$\Lambda(\alpha x + \alpha' x') = \lim_{n \rightarrow \infty} \Lambda_n(\alpha x + \alpha' x') = \lim_{n \rightarrow \infty} (\alpha \Lambda_n x + \alpha' \Lambda_n x') = \alpha \Lambda x + \alpha' \Lambda x'$$

which shows that Λ is linear. Now, for all $x \in X$, we see that we have the following holds:

$$\|\Lambda x\|_y = \left\| \lim_{n \rightarrow \infty} \Lambda_n x \right\|_y = \lim_{n \rightarrow \infty} \|\Lambda_n x\|_y \leq C \|x\|_x$$

hence C is an upper bound on $\{\|\Lambda x\|_y / \|x\|_x \mid x \in X, x \neq 0\}$, this completes the proof of the theorem. \square

Definition 19.2.0.0.1

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be normed spaces, a sequence (Λ_n) of functions in $\mathcal{B}(X, Y)$ is said to converge strongly provided that (Λ_n) converges pointwise in X .

Theorem 19.3

Let $(X, \|\cdot\|_x)$ and $(Y, \|\cdot\|_y)$ be normed spaces, let (Λ_n) be a sequence of functions in $\mathcal{B}(X, Y)$ that converges uniformly to Λ , then (Λ_n) converges strongly.

Proof. Here we can write the following for all $x \in X$:

$$\|\Lambda_n x - \Lambda x\|_y \leq \|\Lambda_n - \Lambda\|_{\mathcal{B}} \|x\|_x$$

taking the limit of large n , the result follows. \square

Example: The converse of Theorem 19.3 is not true.

Here we let $X = C[0, 1]$, $Y = \mathbb{R}$, and $\Lambda_n : X \rightarrow Y$ be functions defined by:

$$\Lambda_n : X \rightarrow Y \quad f \mapsto \int_0^1 f(x) \sin(nx\pi) dx$$

For all polynomial $p \in C[0, 1]$, we can see that we have:

$$\Lambda_n p = \int_0^1 p(x) \sin(n\pi x) dx = \frac{p(0) - p(1)(-1)^n}{n\pi} + \frac{1}{n\pi} \int_0^1 p'(x) \cos(n\pi x) dx$$

where we see that:

$$\lim_{n \rightarrow \infty} \frac{p(0) - p(1)(-1)^n}{n\pi} = 0 \quad \lim_{n \rightarrow \infty} \frac{1}{n\pi} \int_0^1 p'(x) \cos(n\pi x) dx = 0$$

That is, we see that $(\Lambda_n p)$ converges to 0 for all polynomials p . Now for all $f \in X$, by Weierstrass Approximation Theorem, for all $n \in \mathbb{N}$, there exists a polynomial p such that we can write:

$$\|f - p\|_x = \sup_{t \in [0,1]} |f(t) - p(t)| < \frac{1}{n}$$

then we can write, for some constant $K > 0$:

$$|\Lambda_n f| \leq |\Lambda_n(f - p)| + |\Lambda_n p| \leq \|\Lambda_n\|_{\mathcal{B}} \cdot \|f - p\|_x + \frac{K}{n}$$

note also we have:

$$\|\Lambda_n\|_{\mathcal{B}} = \sup_{f \neq 0} \frac{|\Lambda_n f|}{\|f\|_x} \leq 1$$

because we have:

$$|\Lambda_n f| = \left| \int_0^1 f(x) \sin(n\pi x) dx \right| \leq \int_0^1 |f(x)| \cdot |\sin(n\pi x)| dx \leq \|f\|_x$$

and hence combining we get:

$$|\Lambda_n f| \leq \frac{1}{n} + \frac{K}{n}$$

This shows that (Λ_n) converges pointwise to 0 in $X = C[0, 1]$. Moreover, consider $f_n \in X$ defined by $f_n(x) = \sin(n\pi x)$, we can write:

$$\|\Lambda_n\|_{\mathcal{B}} = \sup_{\phi \in X} \left\{ \frac{|\Lambda_n \phi|}{\|\phi\|_x} \right\} \geq \frac{|\Lambda_n f_n|}{\|f_n\|_X} = \frac{1}{2}$$

as we have $\|f_n\|_x = 1$ and that the following holds:

$$|\Lambda_n f| = \int_0^1 \sin^2(n\pi x) dx = \frac{1}{2}$$

Hence we conclude that $(\|\Lambda_n\|_{\mathcal{B}})_{n \geq 1}$ does not converges to 0. This shows that (Λ_n) converges strongly to 0, but it does not converge to 0.

5.19.1 Application in Numerical Quadrature

Let $\omega \in L^1[0, 1]$ and we want to approximate the following integral for all $f \in C[0, 1]$:

$$\int_0^1 f(x) \omega(x) dx$$

Let $X = BC[0, 1]$, equipped with the norm $\|f\|_X = \sup_{x \in [0,1]} |f(x)|$. From previous result, we know that X is a Banach space. Here we define a linear map:

$$l : X \rightarrow \mathbb{R} \quad f \mapsto \int_0^1 f(x) \omega(x) dx$$

We claim that l is a bounded linear map. Note that we can write the following:

$$|l(f)| = \left| \int_0^1 f(x) \omega(x) dx \right| \leq \int_0^1 |f(x)| \cdot |\omega(x)| dx \leq \|f\|_X \cdot \int_0^1 |\omega(x)| dx \leq \|f\|_X \cdot \|\omega\|_{L^1}$$

which holds for all $f \in X$, hence we can write:

$$\|l\| = \sup_{f \neq 0} \frac{|l(f)|}{\|f\|_X} \leq \|\omega\|_{L^1}$$

which shows that l is a bounded linear map.

The Quadrature Rule

Consider the interval $[a, b]$, we introduce a set of nodes: $\{x_j^{(n)} \mid j \in \mathbb{N} \cup \{0\}, 0 \leq j \leq n\}$ of distinct points in $[a, b]$ that satisfies $a \leq x_0^{(n)} < x_1^{(n)} < x_2^{(n)} < \dots < x_n^{(n)} \leq b$. Also define the corresponding weights $\{\omega_j^{(n)} \mid j \in \mathbb{N} \cup \{0\}, 0 \leq j \leq n\}$. Now define the function l_n as the following:

$$l_n : X \rightarrow \mathbb{R} \quad l_n(f) = \sum_{j=0}^n \omega_j^{(n)} f(x_j^{(n)})$$

here we see that $l_n \in \mathcal{B}(X, \mathbb{R})$, and we have $\|l_n\| = \sum_{j=0}^n |\omega_j^{(n)}|$:

$$|l_n(f)| = \left| \sum_{j=0}^n \omega_j^{(n)} f(x_j^{(n)}) \right| \leq \sum_{j=0}^n |\omega_j^{(n)}| \cdot |f(x_j^{(n)})| \leq \sum_{j=0}^n |\omega_j^{(n)}| \cdot \|f\|_X$$

hence it follows that:

$$\|l_n\| \leq \sum_{j=0}^n |\omega_j^{(n)}|$$

Note that, if one has $f(x_j^{(n)}) = \text{sign}(\omega_j^{(n)})$. Then we have $l_n(f) = \sum_{j=0}^n |\omega_j^{(n)}|$.

Lemma 19.3.1 (Newton-Cotes)

Pick the nodes $\{x_j^{(n)} \mid j \in \mathbb{N} \cup \{0\}, 0 \leq j \leq n\}$, for $f \in C[a, b]$, interpolate $(f(x_j^{(n)}))_{j=0}^n$ by the Lagrange polynomials, we can write:

$$f(x) \approx \sum_{j=0}^n f(x_j^{(n)}) p_j(x)$$

where we have p_j being the the Lagrange polynomial:

$$p_j : \mathbb{R} \rightarrow \mathbb{R} \quad x \mapsto \prod_{q=0, q \neq j}^n \frac{x - x_q^{(n)}}{x_j^{(n)} - x_q^{(n)}}$$

Note that we have $p_j(x_q^{(n)}) = \delta_{jq}$. Then we have:

$$l_n(f) = \sum_{j=0}^n f(x_j^{(n)}) \omega_j^{(n)} \quad \omega_j^{(n)} = \int_0^1 p_j(x) dx$$

Note that the quadrature is exact for f being polynomial of degree less than n . We also see that $\{p_0, p_1, \dots, p_n\}$ is a linear independent set of polynomials, hence it is a basis of the space of polynomials of degree less than or equal to n .

Lemma 19.3.2 (Gauss-Jacobi)

Pick both nodes and weights such that we have exact quadrature for polynomial f of degree less than equal to $2n + 1$. One can also derive an approximation of quadrature for arbitrary $f \in C[a, b]$.

Theorem 19.4 (Polya Theorem)

Given $\omega \in L^1[0, 1]$, let (l_n) be a sequence of quadrature rules such that we have:

$$\lim_{n \rightarrow \infty} l_n(p) = l(p)$$

for all p in the set of polynomial. Then we have:

$$\lim_{n \rightarrow \infty} l_n(f) = l(f) \quad \forall f \in C[0, 1]$$

if and only if we have:

$$\sup_{n \geq 0} \sum_{j=0}^n |\omega_j(n)| < \infty$$

Proof. First we suppose $(l_n(f))$ converges to $l(f)$ for all $f \in X$, that is, (l_n) converges pointwise to l . Then we know that $l_n(f)$ is bounded for all $f \in X$. In particular, for $\|f\|_X \leq 1$, by Corollary 19.1.1, we can write the following:

$$\sup_{n \geq 0} \sum_{j=0}^n |\omega_j^{(n)}| = \sup_{n \geq 0} \|l_n\| = M < \infty$$

For the converse, we suppose now $M = \sup_{n \geq 0} \|l_n\| = \sup_{n \geq 0} \sum_{j=0}^n |\omega_j^{(n)}| < \infty$ and $l_n(p)$ approaches $l(p)$ for all polynomial p . For all $f \in X = C[0, 1]$, we have:

$$\begin{aligned} |l_n(f) - l(f)| &= |l_n(f - p) + l_n(p) - l(p) - l(f - p)| \\ &\leq |l_n(f - p)| + |l_n(p) - l(p)| + |l(f - p)| \\ &\leq (\|l_n\| + \|l\|) \cdot \|f - p\|_X + |l_n(p) - l(p)| \\ &\leq (M + \|l\|) \cdot \|f - p\|_X + |l_n(p) - l(p)| \end{aligned}$$

for some $M > 0$, and for some polynomial p . Given $\epsilon > 0$, by Weierstrass Approximation Theorem, there exists some polynomial p such that $\|f - p\|_X < \epsilon$. For such p , by assumption on (l_n) , we can write the following:

$$|l_n(p) - l(p)| < \epsilon \quad \Rightarrow \quad |l_n(f) - l(f)| < (M + \|l\|) \cdot \epsilon + \epsilon \quad \forall n \in \mathbb{N}$$

□

Approximation Schemes

Here we consider two problems:

1. Suppose we want to solve the problem $Au = f$, where $A : X \rightarrow Y$ is an invertible linear operator and $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ are Banach space.
2. Consider the nearby problem $A_\epsilon u_\epsilon = f_\epsilon$ where $A_\epsilon : X \rightarrow Y$ has bounded inverse, that is A_ϵ^{-1} is a bounded linear map. Here A_ϵ is defined in attempting to have $\lim_{\epsilon \rightarrow 0} u_\epsilon = u$ when $\lim_{\epsilon \rightarrow 0} A_\epsilon u_\epsilon = \lim_{\epsilon \rightarrow 0} f_\epsilon = f$.

For example, suppose we have a mesh, or nodes, in $[0, 1]$, with spacing $h = 1/n$ for some given n . Then for $x \in [x_j, x_{j+1}]$, we can approximate the derivative of differentiable function $u \in C[0, 1]$ at x by the following:

$$u'(x) \approx \frac{u(x_{j+1}) - u(x_j)}{h} \quad (5.3)$$

In such case, if one defines A_ϵ to be the function of finding the approximated derivative of u as in (5.3), then as $\epsilon = h$ approaches zero, A_ϵ approaches A where A is the differential operator.

The family (2) of equation associated to some $\epsilon > 0$ is called an approximation scheme for (1). In such sense we define the following:

Definition 20.0.0.0.1

Approximation scheme is consistent if we have $\lim_{\epsilon \rightarrow 0} A_\epsilon v = Av$ for all $v \in X$.

Definition 20.0.0.0.2

Approximation scheme is stable if there exists a constant M , independent of ϵ , such that we have $\|A_\epsilon^{-1}\| \leq M$ for all ϵ .

Theorem 20.1 (Lax Equivalence Theorem)

Approximation scheme is convergent, that is $\lim_{\epsilon \rightarrow 0} \|u_\epsilon - u\|_X = 0$ for $\lim_{\epsilon \rightarrow 0} \|f_\epsilon - f\|_Y$ if and only if the approximation scheme is consistent and stable.

Proof. For the \Leftarrow direction, the approximation scheme is stable and consistent. Suppose we have $Au = f$ and $A_\epsilon u_\epsilon = f_\epsilon$. Then we can write the following:

$$u - u_\epsilon = A_\epsilon^{-1} (A_\epsilon u - Au + f - f_\epsilon)$$

note that we can write:

$$\|u - u_\epsilon\| = \|A_\epsilon^{-1}\| (\|A_\epsilon u - Au\|_Y + \|f - f_\epsilon\|_Y) \quad (5.4)$$

here we see that $\|u - u_\epsilon\|$ in (5.4) can be bounded, because, for small ϵ , $\|A_\epsilon^{-1}\|$ is bounded by stability, $\|A_\epsilon u - Au\|_Y$ approaches zero by consistency, and $\|f - f_\epsilon\|_Y$ approaches zero by assumption. Now for \Rightarrow direction, suppose there exists f_ϵ such that we have $\|f_\epsilon - f\|_Y$ approaches zero when ϵ approaches zero, and yet $\|u_\epsilon - u\|_X$ does not approaches zero when ϵ approaches zero. Similarly, (5.4) suggests that the approximation scheme is either, or both, not consistent or not stable. \square

Operator Power Series

Let $(X, \|\cdot\|)$ be a Banach space, and let $A \in \mathcal{B}(X, X) = \mathcal{B}(X)$. Here we define:

$$A^2 : X \rightarrow X \quad x \mapsto A(A(x))$$

Lemma 21.0.1

Let $(X, \|\cdot\|)$ be a Banach space, and let $A \in \mathcal{B}(X)$, then $A^2 \in \mathcal{B}(X)$

Proof. Note that we can write the following for all $x \in X$:

$$\|A^2x\| = \|A(Ax)\| \leq \|A\|_{\mathcal{B}} \cdot \|Ax\| \leq \|A\|_{\mathcal{B}}^2 \|x\|$$

hence it follows that we have:

$$\|A^2\|_{\mathcal{B}} \leq \|A\|_{\mathcal{B}}^2$$

The result follows. □

Definition 21.0.1.0.1

Let $(X, \|\cdot\|)$ be a Banach space, and let $A \in \mathcal{B}(X)$, we define, for $n \in \mathbb{N}$:

$$A^n : X \rightarrow X \quad x \mapsto \underbrace{AAA \cdots A}_n x$$

Note: A^n defined in Definition 21.0.1.0.1 is a bounded operator.

Theorem 21.1

Let $(X, \|\cdot\|)$ be a Banach space, and let $\sum_{n=0}^{\infty} \alpha_n z^n$ be a power series in \mathbb{C} , that is $\alpha_n \in \mathbb{R}$ and $z \in \mathbb{C}$, with radius of convergence ρ . Let $A \in \mathcal{B}(X)$, with $\|A\|_{\mathcal{B}} < \rho$, then we have $\sum_{n=0}^{\infty} \alpha_n A^n \in \mathcal{B}(X)$.

Proof. First note that $\mathcal{B}(X)$ is a Banach space. Let $z := \|A\|_{\mathcal{B}} < \rho$, then we can write:

$$S_N := \sum_{n=0}^N |\alpha_n| \cdot \|A^n\|_{\mathcal{B}} \leq \sum_{n=0}^N |\alpha_n| \cdot \|A\|_{\mathcal{B}}^n = \sum_{n=0}^N |\alpha_n| \cdot |z|^n \quad (5.5)$$

Hence, as N approaches infinity, the RHS of (5.5) converges, so $(S_N)_{N \geq 1}$ is an increasing sequence in \mathbb{R} that is bounded, then we can write:

$$\lim_{N \rightarrow \infty} S_N = \sum_{n=1}^{\infty} |\alpha_n| \cdot \|A^n\|_{\mathcal{B}} < \infty$$

Since $\mathcal{B}(X)$ is Banach, and $\sum_{n=0}^{\infty} A^n$ is absolutely convergent, then we can write:

$$L = \sum_{n=0}^{\infty} \alpha_n A^n \in \mathcal{B}(X)$$

□

Definition 21.1.0.0.1

Let $(X, \|\cdot\|)$ be a Banach space, and let $A \in \mathcal{B}(X)$, we define:

$$e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

Example: Consider we have a system:

$$\frac{\partial}{\partial t} F(t, x) = AF(t, x)$$

where $t, x \in [a, b]$ with some reasonable function F, A , and suppose further we have $F(0, x) = f(x) \in X = L^2[a, b]$, then we can write:

$$F(t, x) = e^{tA} f(x)$$

Theorem 21.2 (Neumann series)

Let $(X, \|\cdot\|)$ be a Banach space, and let $A \in \mathcal{B}(X)$ with $\|A\|_{\mathcal{B}} < 1$. Then the function:

$$I - A : X \rightarrow X \quad x \mapsto x - Ax$$

defines a bijective continuous linear operator with inverse $(I - A)^{-1} = \sum_{n=0}^{\infty} A^n$.

Proof. By Theorem 21.1, since we have $\|A\|_{\mathcal{B}} < 1$ and the radius of convergence of the complex series $\sum_{n=0}^{\infty} z^n$ is 1, then we know that:

$$L := \sum_{n=0}^{\infty} A^n \in \mathcal{B}(X) \quad \|L\|_{\mathcal{B}} \leq \sum_{n=0}^{\infty} \|A\|_{\mathcal{B}}^n = \frac{1}{1 - \|A\|_{\mathcal{B}}}$$

Now we define the sequence (L_n) in $\mathcal{B}(X)$ by the following:

$$L_n := \sum_{j=0}^n A^j$$

where we see that we have $(L_n) \rightarrow L$, and $(\|L_n - L\|_{\mathcal{B}}) \rightarrow 0$, and we can write:

$$L_{n+1} = \sum_{j=0}^{n+1} A^j = I + \sum_{j=1}^{n+1} A^j = I + A \sum_{j=0}^n A^j = I + AL_n$$

taking the large limit of n , we obtain the following:

$$L = I + AL \quad \Rightarrow \quad (I - A)L = I$$

on the other hand, we have:

$$L_{n+1} = \sum_{j=0}^{n+1} A^j = I + \left(\sum_{j=1}^n A^j \right) A$$

hence we have, taking the large limit of n :

$$L(I - A) = I$$

and it follows that $L = (I - A)^{-1}$. This completes the proof. \square

Example: Let $(X, \|\cdot\|)$ be a Banach space, and let $A \in \mathcal{B}(X)$, consider the system $Ax = b$, with $A = M + E = (I + EM^{-1})M$, where M is invertible with $M^{-1} \in \mathcal{B}(X)$, and $\|EM^{-1}\|_{\mathcal{B}} < 1$. Then we can write:

$$(I + EM^{-1})Mx = b$$

where we write $Mx = y$, and hence we have:

$$y = (I + EM^{-1})^{-1}b = \sum_{j=0}^{\infty} (-EM^{-1})^j b = Mx$$

6 | Hilbert Spaces

Inner Product Spaces

Definition 22.0.0.0.1

Let X be a vector space over a field \mathbb{F} . The inner product on X is a bilinear function $(\cdot, \cdot) : X \times X \rightarrow \mathbb{C}$ that satisfies the followings:

1. $(x, y) = \overline{(y, x)}$ for all $x, y \in X$. If the field $\mathbb{F} = \mathbb{R}$, then such property is called the symmetric property, if instead the field $\mathbb{F} = \mathbb{C}$, then such property is called the Hermitian property.
2. For all $x, y, z \in X$, and all $\alpha, \beta \in \mathbb{F}$, we have $(x, \alpha y + \beta z) = \alpha(x, y) + \beta(x, z)$. That is the function (\cdot, \cdot) is linear in its second argument. If $\mathbb{F} = \mathbb{R}$, then the inner product is linear in both arguments, if $\mathbb{F} = \mathbb{C}$ instead, then the inner product is sesquilinear.
3. For all $x \in X$, $(x, x) \geq 0$ and $(x, x) = 0$ if and only if $x = 0$.

Note: Note that the second property in Definition 22.0.0.0.1 implies the following:

$$(\alpha y + \beta z, x) = \overline{(x, \alpha y + \beta z)} = \overline{\alpha(x, y) + \beta(x, z)} = \overline{\alpha} \overline{(x, y)} + \overline{\beta} \overline{(x, z)} = \overline{\alpha} (y, x) + \overline{\beta} (z, x)$$

Definition 22.0.0.0.2

Let X be a vector space over a field \mathbb{F} , let (\cdot, \cdot) be an inner product on X , then $(X, (\cdot, \cdot))$ is called an inner product space.

Theorem 22.1

Let $(X, (\cdot, \cdot))$ be an inner product space. The function $\|\cdot\| : X \rightarrow \mathbb{R} \quad x \mapsto \sqrt{(x, x)}$ defines a norm on X . $(X, \|\cdot\|)$ is a normed space and we have, for all $x, y \in X$, the followings hold:

1. $|(x, y)| \leq \|x\| \cdot \|y\|$. Such property is called the Cauchy Schwartz Inequality
2. $\|x + y\| \leq \|x\| + \|y\|$. Such property is called the Minkovsky Inequality.

Proof. It is clear that we have $\|x\| \geq 0$, $\|x\| = 0$ if and only if $x = 0$ for all $x \in X$ from the definition of inner product. For all $\alpha \in \mathbb{F}$, we can write the following:

$$\|\alpha x\|^2 = (\alpha x, \alpha x) = \alpha(\alpha x, x) = \alpha \overline{\alpha} (x, x) = |\alpha|^2 \|x\|^2$$

Hence it follows that $\|\cdot\|$ defines a norm on X . Now we will show the Cauchy Schwartz Inequality, here we note that for all $y \in X$, we have $0 = (0, y) = (0+0, y) = (0, y) + (0, y)$, and hence we have:

$$0 = |(0, y)| \leq \|0\| \cdot \|y\| = 0$$

here we will suppose $x, y \in X$ with $x \neq 0 \neq y$, denote $a = \|x\|^2$, $b = (x, y)$, and $c = \|y\|^2$, then for all $\alpha \in \mathbb{F}$, we can write the following:

$$0 \leq (x + \alpha y, x + \alpha y) = (x, x) + \alpha(x, y) + \bar{\alpha}(y, x) + |\alpha|^2(y, y) = a + \alpha b + \bar{\alpha} \bar{b} + |\alpha|^2 c$$

here we set $\alpha = -\bar{b}/c$, then we get:

$$0 \leq a - \frac{\bar{b}b}{c} - \frac{b\bar{b}}{c} + \frac{|b|^2}{c^2}c = a - \frac{|b|^2}{c} \quad \Rightarrow \quad |b|^2 \leq ac$$

the result of Cauchy Inequality follows. For the Minkovsky Inequality, let $x, y \in X$, then we can write the following:

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

rearranging we get the following:

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2$$

□

Lemma 22.1.1

Inner product functions are continuous functions.

Proof. Consider the inner product (\cdot, \cdot) defined on a vector space X over a field \mathbb{F} . Then we can write the following for $x, y, x', y' \in X$:

$$\begin{aligned} |(x, y) - (x', y')| &= |(x - x', y) + (x', y - y')| \\ &\leq |(x - x', y)| + |(x', y - y')| \\ &\leq \|y\| \cdot \|x - x'\| + \|x'\| \cdot \|y - y'\| \\ &\leq \|y\| \cdot \|x - x'\| + (\|x\| + \|x - x'\|) \cdot \|y - y'\| \\ &\leq \|y\| \cdot \|x - x'\| + \|x\| \cdot \|y - y'\| + \|x - x'\| \cdot \|y - y'\| \end{aligned}$$

Here it is clear that (\cdot, \cdot) is continuous. □

Definition 22.1.1.0.1

A Hilbert space is an inner product space $(X, (\cdot, \cdot))$ that is complete with respect to the norm induced by the inner product.

Note: Hilbert spaces are Banach spaces.

Example: For $X = \mathbb{C}^n$, inner product (\cdot, \cdot) defined by $(x, y) = \sum_{j=1}^N \bar{x}_j y_j$. Then the induced norm is given by:

$$\|x\| = \sqrt{\sum_{j=1}^N |x_j|^2}$$

the space $(\mathbb{C}^n, (\cdot, \cdot))$ constitutes a Hilbert space.

Example: For $X = \mathbb{C}^{N \times M}$, inner product (\cdot, \cdot) defined by the following:

$$(A, B) = \sum_{i=1}^N \sum_{j=1}^M \bar{A}_{ij} B_{ij} = \text{Trace}(A^* B)$$

The induced norm is given by:

$$\|A\| = \sqrt{\sum_{i=1}^N \sum_{j=1}^M |A_{ij}|^2}$$

which is a Frobenius norm, and that $(\mathbb{C}^{n \times M}, (\cdot, \cdot))$ constitutes a Hilbert space.

Example: For $X = l^2$, inner product (\cdot, \cdot) defined by $(x, y) = \sum_{j=1}^{\infty} \bar{x}_j y_j$, where $x = (x_j)$ and $y = (y_j)$. The induced norm is given by the following:

$$\|x\| = \sum_{j=1}^{\infty} |x_j|^2 < \infty$$

and here $(l^2, (\cdot, \cdot))$ constitutes a Hilbert space.

Orthogonality

Definition 23.0.0.0.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space. For $x, y \in H$, x and y are said to be orthogonal, denoted as $x \perp y$, provided that $(x, y) = 0$. For $A, B \subseteq H$, A and B are said to be orthogonal sets, denoted as $A \perp B$, provided that $(x, y) = 0$ for all $x \in A$ and all $y \in B$.

Definition 23.0.0.0.2

Let $(H, (\cdot, \cdot))$ be a Hilbert space. For $A \subseteq H$, the orthogonal complement of A , denoted as A^\perp , is defined as the set $\{x \in H \mid x \perp a, \forall a \in A\}$.

Theorem 23.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space. For all $A \subseteq H$, A^\perp is a closed subspace of H .

Proof. For all $x, \xi \in A^\perp$, for all α, β in the scalar field, and for all $z \in A$, we can write the following:

$$(x, z) = (\xi, z) = 0$$

and hence we have:

$$(\alpha x + \beta \xi, z) = \bar{\alpha}(x, z) + \bar{\beta}(\xi, z) = 0$$

Then it is easy to see that A^\perp is a subspace. To show that A^\perp is closed, we consider a sequence (x_n) of points in A^\perp , and suppose (x_n) converges to some $x \in H$. Here for all $z \in A$, we see that $(x_n, z) = 0$ for all $n \in \mathbb{N}$, then we see that the sequence $((x_n, z))$ converges to 0, then we can write:

$$0 = \lim_{n \rightarrow \infty} (x_n, z) = \left(\lim_{n \rightarrow \infty} x_n, z \right) = (x, z)$$

hence this shows that $x \in A^\perp$, and that completes the proof. □

Definition 23.1.0.0.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space.

A projection from H to H is a linear map $P : H \rightarrow H$ such that $P^2 = P$.

Definition 23.1.0.0.2

Let $(H, (\cdot, \cdot))$ be a Hilbert space.

A projection P is said to be orthogonal provided that $(Px, y) = (x, Py)$ for all $x, y \in H$.

Definition 23.1.0.0.3

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let M, N be subspaces of H . The direct sum $M \oplus N$ is defined to be the following:

$$M \oplus N := \{z \in H \mid z \text{ can be decomposed uniquely as } z = x + y \text{ for some } x \in M, y \in N\}$$

If we have further $M \perp N$, then $M \oplus N$ is called the orthogonal direct sum.

Note: In the definition of 23.1.0.0.3, we see that M, N must satisfy $M \cap N = \{0\}$.

Theorem 23.2

Let $(H, (\cdot, \cdot))$ be a Hilbert space. Let $P : H \rightarrow H$ be a projection, then $H = \text{im}(P) \oplus \ker(P)$. Conversely, if $H = M \oplus N$ for some subspaces M, N of H , then there exists a projection $P : H \rightarrow H$ such that $M = \text{im}(P)$ and $N = \ker(P)$.

Proof. We first show that we have $x \in \text{im}(P)$ if and only if $x = Px$. Say $x \in \text{im}(P)$, we see that there exists $y \in H$ such that $x = Py$, then $Px = P^2y = Py = x$. Note also if $x \in \text{im}(P) \cap \ker(P)$, we have $x = Px = 0$, hence we have $\ker(P) \cap \text{im}(P) = \{0\}$. Now for $x \in H$, we see that $x = Px + (x - Px)$, where $Px \in \text{im}(P)$, and on the other hand $P(x - Px) = Px - P^2x = 0$ so $x - Px \in \ker(P)$. This shows that $H = \text{im}(P) \oplus \ker(P)$. For the converse, suppose $H = M \oplus N$, then for all $x \in H$, we can write $x = \xi + \eta$ for $\xi \in M$ and $\eta \in N$. We define $P : H \rightarrow H$ $\xi + \eta \mapsto \xi$. First we notice that $P^2x = P(Px) = P(\xi) = \xi = Px$. Now for $x = \xi + \eta$ and $x' = \xi' + \eta'$, we can write the following with scalars α, α' :

$$P(\alpha x + \alpha' x') = P(\alpha \xi + \alpha' \xi' + \alpha \eta + \alpha' \eta') = \alpha \xi + \alpha' \xi' = \alpha Px + \alpha' P(x')$$

The image of P is the set $\{x \in H \mid Px = x\} = M$, and the kernel of P is the set N as one can check easily. This completes the proof. \square

Theorem 23.3 (Projection Theorem)

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let V be a closed subspace of H , then we have:

1. $H = V \oplus V^\perp$
2. For all $x \in H$, there exists unique $y \in V$ such that $P_V(x) = y$ is closest to x , that is we have y satisfies $\|y - x\| = \min_{u \in V} \|u - x\| = d(x, V)$.
3. For all $x \in V^\perp$, there exists $z = P_{V^\perp}(x)$ that is the unique closest point to x .
4. P_V and P_{V^\perp} characterized in (2) and (3) are orthogonal projections, and they are continuous with norm less than or equal to 1.

Proof. First we will show that, for all $x \in H$, there exists unique $y \in V$ that is closest to x . By definition, we can write $d(x, V) = \inf_{u \in V} \|u - x\|$. By characterization of infimum, for all $n \in \mathbb{N}$, we can find $y_n \in V$ such that $d(x, V) \leq \|y_n - x\| \leq d(x, V) + \frac{1}{n}$. Hence we obtain (y_n) of points in V such that the sequence $(\|y_n - x\|)$ converges to $d(x, V)$. We claim that (y_n) is Cauchy, in which case we know that (y_n) converges to some $y \in V$ by the completeness of H and the closeness of V . We will show that (y_n) is Cauchy, note that we have the following:

$$\|y_n - y_m\|^2 = \|(y_n - x) - (y_m - x)\|^2 = \|y_n - x\|^2 + \|y_m - x\|^2 - 2\Re(y_n - x, y_m - x) \quad (*)$$

and on the other hand:

$$\begin{aligned} 4 \left\| \frac{y_n + y_m}{2} - x \right\|^2 &= \|(y_n - x) + (y_m - x)\|^2 \\ &= \|y_n - x\|^2 + \|y_m - x\|^2 + 2\Re(y_n - x, y_m - x) \end{aligned} \quad (**)$$

Combining we get the following:

$$\|y_n - y_m\|^2 = 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4 \left\| \frac{y_n + y_m}{2} - x \right\|^2$$

Notice here $(y_n + y_m)/2 \in V$:

$$\left\| \frac{y_n + y_m}{2} - x \right\| \geq d(x, V)$$

and hence we have:

$$\begin{aligned} \|y_n - y_m\|^2 &= 4 \left(d^2(x, V) - \left\| \frac{y_n + y_m}{2} - x \right\|^2 \right) + 2(\|y_n - x\|^2 - d^2(x, V)) + 2(\|y_m - x\|^2 - d^2(x, V)) \\ &\leq 2(\|y_n - x\|^2 - d^2(x, V)) + 2(\|y_m - x\|^2 - d^2(x, V)) \end{aligned}$$

Now it is easy to see that (y_n) is Cauchy. Now we can write the following for $y \in V$:

$$d(x, V) = \lim_{n \rightarrow \infty} \|y_n - x\| = \left\| \lim_{n \rightarrow \infty} y_n - x \right\| = \|y - x\|$$

Now suppose both $y, y' \in V$ satisfy $\|x - y\| = \|x - y'\| = d(x, V)$, we can use (*) and (**) to write the followings:

$$\|y - y'\|^2 + 4 \left\| \frac{y + y'}{2} - x \right\|^2 = 2\|y - x\|^2 + 2\|y' - x\|^2 = 4d^2(x, V)$$

that is we have:

$$0 \leq \|y - y'\|^2 = 4 \left(d^2(x, V) - \left\| \frac{y + y'}{2} - x \right\|^2 \right) \leq 0$$

hence we must have $y = y'$. Now for $x \in H$, we can define $P_V(x) = y$ and such that $\|y - x\| = d(x, V)$. We will show that, for all $x \in H$, we have $x - P_V(x) \perp V$. Take $v \in V$, and $\lambda \in \mathbb{R}$ such that $\Lambda(\lambda) := \|x - (y + \lambda v)\|^2$, note that Λ has a minimum at $\lambda = 0$, then we can write the following:

$$\left. \frac{d\Lambda}{d\lambda} \right|_{\lambda=0} = \frac{d}{d\lambda} (\|x - y\|^2 + \lambda^2\|v\|^2 - 2\lambda\Re(x - y, v)) \big|_{\lambda=0} = 2\Re(x - y, v) = 0$$

if one repeats the argument for iv , we can get $(x - y, v) = 0$ for all $v \in V$. Note that no other $y' \in V$ gives $x - y' \perp V$, otherwise we can write the following:

$$\|y - y'\|^2 = (y - y', y - x - (y' - x)) = (y - y', y - x) - (y - y', y' - x) = 0$$

which implies $y = y'$. Now we will show that P_V is a linear map. Let $x_1, x_2 \in H$ and let α_1, α_2 be scalars, then we can write $y_1 = P_V(x_1)$ being unique in V such that $x_1 - y_1 \in V^\perp$, and we have $y_2 = P_V(x_2)$ being unique in V such that $x_2 - y_2 \in V^\perp$. Hence we see that we have:

$$\alpha_1 x_1 + \alpha_2 x_2 - (\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1(x_1 - y_1) + \alpha_2(x_2 - y_2) \in V^\perp$$

hence by previous argument, we see here $(\alpha_1 y_1 + \alpha_2 y_2) = P_V(\alpha_1 x_1 + \alpha_2 x_2)$, hence showing that P_V is a linear map. Now we will show that P_V is a projection. Take $x \in H$, we see that $P_V(x) = y$ is closest to $x \in V$. Moreover, we see that $P_V^2(x) = P_V(y) = y$ being the closest point to y in V . Hence we see that P_V is a projection. Now we will show that P_V is an orthogonal projection. For all $x, \xi \in H$, we can write

$(P_V(x), \xi) = (P_V(x), \xi - P_V(\xi)) + (P_V(x), P_V(\xi))$, where $P_V(x) \in V$ and $\xi - P_V(\xi) \in V^\perp$, hence we can write the following:

$$(P_V(x), \xi) = (P_V(x), P_V(\xi)) = (x, P_V(\xi)) - (x - P_V(x), P_V(\xi)) = (x, P_V(\xi))$$

this shows that P_V is an orthogonal projection. Similar result can be proved for P_{V^\perp} , the steps are similar, and the proof for P_{V^\perp} is left for the reader, and moreover, one can show that $P_{V^\perp} = I - P_V$. Now we will show that P_V, P_{V^\perp} are continuous. Let $x \in H$ with $x \neq 0$, here we write the followings:

$$\frac{\|P_V(x)\|^2}{\|x\|^2} \leq \frac{\|P_V(x)\|^2 + \|(I - P_V)(x)\|^2}{\|x\|^2} = \frac{\|x\|^2}{\|x\|^2} = 1$$

where we used the fact that:

$$\begin{aligned} (x, x) &= (P_V(x) + (I - P_V)(x), P_V(x) + (I - P_V)(x)) \\ &= \|P_V(x)\|^2 + \|(I - P_V)(x)\|^2 + 2\Re(P_V(x), (I - P_V)(x)) \\ &= \|P_V(x)\|^2 + \|(I - P_V)(x)\|^2 \end{aligned}$$

Hence it follows that we have $\|P_V\| \leq 1$, so P_V is continuous. Similar results can be proved for P_{V^\perp} . Note that, if $V \neq \{0\}$, take $x \in V$, we have $P_V(x) = x$, hence in such case $\|P_V\| = 1$. Lastly, for all $x \in H$, we see that $x = P_V(x) + (I - P_V)(x)$, where $P_V(x) \in V$ and $(I - P_V)(x) \in V^\perp$ and hence concluding that we have $H = V \oplus V^\perp$. This completes the proof of the theorem. \square

Theorem 23.4

Let $(H, (\cdot, \cdot))$ be a Hilbert space. The followings hold:

1. If $P : H \rightarrow H$ is an orthogonal projection, then the image of P is closed and we have $H = \text{im}(P) \oplus \ker(P)$, with $\ker(P)$ being orthogonal to $\text{im}(P)$.
2. If V is a closed subspace of H , then there exists an orthogonal projection $P_V : H \rightarrow H$ that satisfies $V = \text{im}(P)$ and $V^\perp = \ker(P)$.

Proof. Here (2) follows from Theorem 23.3, we will prove (1) here. Note that it suffices to show $\text{im}(P) = (\ker(P))^\perp$. We first want to show $\text{im}(P) \subseteq (\ker(P))^\perp$. Note that $x \in \text{im}(P)$ if and only if $x = Px$ as we have proved in Theorem 23.2. For $z \in \ker(P)$, we want to show that $(x, z) = 0$ where $x \in \text{im}(P)$. Here we can write the following:

$$(x, z) = (P(x), z) = (x, P(z)) = (x, 0) = 0$$

this shows that $\text{im}(P) \subseteq (\ker(P))^\perp$. Now we will show that $(\ker(P))^\perp \subseteq \text{im}(P)$. For $x \in (\ker(P))^\perp$, and $z \in \ker(P)$, we can write the following:

$$0 = (x, z) = (x, z) - (x, P(z)) = (x, z) - (P(x), z) = (x - P(x), z)$$

where we see that $x - P(x) \in \ker(P)$, here we can take $z = x - P(x)$, so we have $(x - P(x)) = 0$, and this shows that $x \in \text{im}(P)$. This shows that $\text{im}(P) = (\ker(P))^\perp$, completing the proof. \square

6.23.1 Linear Least Squares

Let $A \in \mathbb{R}^{m \times n}$ be a linear operator, and let $c \in \mathbb{R}^m$. Consider a function defined by $\mathcal{O}(\xi) = \|A\xi - c\|^2$ for $\xi \in \mathbb{R}^n$, minimizing \mathcal{O} is a least squares problem, and this is of particular interest when we have $c \notin \text{im}(A)$.

Let $V = \text{im}(A)$, which is a subspace of \mathbb{R}^m , which is closed because it is finite dimensional. Now by Theorem 23.3, there exists a unique $y \in V$ such that $\|c - y\| = d(c, V)$. Now since $A\xi - y \in V$, and $y - c \in V^\perp$, then we can write the following:

$$\|(A\xi - y) + (y - c)\|^2 = \|A\xi - y\|^2 + \|y - c\|^2$$

For $y \in V$, there exists $\xi \in \mathbb{R}^n$ such that $A\xi = y$, then we see that $A\xi - c = y - c \perp V$, hence for all $x \in \mathbb{R}^n$, we see that $(A\xi - c, Ax) = \bar{x}^T A^*(A\xi - c) = 0$. Take $x = \vec{e}_j$, then we get the following:

$$A^*A\xi = A^*c \quad (\text{Normal Equation})$$

where A^* is the Hermitian conjugate of A . Solving the normal equation gives us the result for ξ which minimize the distance.

Orthonormal Sets

Definition 24.0.0.0.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space, let $S \subseteq H$ be a subset.

S is said to be orthogonal provided that for all $x, y \in S$ with $x \neq y$, we have $(x, y) = 0$.

S is said to be orthonormal provided that S is orthogonal, and all $x \in S$ satisfies $\|x\| = 1$.

Example: A function that is the sum of finitely many periodic functions is called quasiperiodic. Let X be the space of quasiperiodic function defined on \mathbb{R} with codomain \mathbb{C} . Here for $f \in X$, we can write the following for all $t \in \mathbb{R}$:

$$f(t) = \sum_{k=1}^n a_k e^{i\omega_k t} \quad n \in \mathbb{N}, \quad a_k \in \mathbb{C}, \quad \omega_k \in \mathbb{R}$$

Here we can define an inner product on X :

$$(\cdot, \cdot) : X \times X \rightarrow \mathbb{R} \quad (f, g) \mapsto \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \overline{f(t)} g(t) dt$$

We claim that $S = \{f_\omega : \mathbb{R} \rightarrow \mathbb{C} \mid t \mapsto e^{i\omega t} \mid \omega \in \mathbb{R}\}$ defines a uncountable orthonormal set. Here for ω, ω' , we can write the following:

$$(f_\omega, f_{\omega'}) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{i(\omega' - \omega)t} dt = \lim_{T \rightarrow \infty} \frac{\sin((\omega - \omega')T)}{(\omega - \omega')T} = \begin{cases} 0 & \omega \neq \omega' \\ 1 & \omega = \omega' \end{cases}$$

Here \bar{X} equipped with the norm (\cdot, \cdot) defined above is called the Hilbert space of L^2 almost periodic functions. Here \bar{X} consists of functions f that satisfies the following for all $t \in \mathbb{R}$:

$$f(t) = \sum_{k=1}^{\infty} a_k e^{i\omega_k t} \quad \text{such that} \quad \sum_{k=1}^{\infty} |a_k|^2 < \infty$$

where $a_k \in \mathbb{C}$.

6.24.1 Unordered Sums

Let $S = \{x_\alpha \in H \mid \alpha \in \mathcal{J}\}$ where \mathcal{J} is an index set, can be uncountable, and $(H, (\cdot, \cdot))$ is a Hilbert space. For any finite $J \subseteq \mathcal{J}$, we write $s_J = \sum_{\alpha \in J} x_\alpha$.

Definition 24.0.0.0.2

Let \mathcal{J} be an index set, can be uncountable, let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \in H \mid \alpha \in \mathcal{J}\}$. An unordered sum $\sum_{\alpha \in \mathcal{J}} x_\alpha$ converges unconditionally to some $x \in H$, denoted as $x = \sum_{\alpha \in \mathcal{J}} x_\alpha$, provided that for all $\epsilon > 0$, there exists finite $J_\epsilon \subseteq \mathcal{J}$ such that $\|x - s_J\| < \epsilon$ for all finite $J \subseteq \mathcal{J}$ such that $J_\epsilon \subseteq J$.

Definition 24.0.0.0.3

Let \mathcal{J} be an index set, can be uncountable, let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \in H \mid \alpha \in \mathcal{J}\}$. $\sum_{\alpha \in \mathcal{J}} x_\alpha$ converges absolutely provided that $\sum_{\alpha \in \mathcal{J}} \|x_\alpha\|$ converges unconditionally in \mathbb{R} .

Note: Absolutely convergence implies unconditionally convergence, and converse is true in finite dimensions.

Example: Consider a Hilbert space $(H, (\cdot, \cdot))$ of infinite dimension.

Take a set $\{x_n \mid n \in \mathbb{N}\}$ that is orthonormal. Consider the following sum:

$$\sum_{n=1}^{\infty} \frac{x_n}{n} \quad (6.1)$$

(6.1) is not an absolutely converging sum because we have:

$$\sum_{n=1}^{\infty} \left\| \frac{x_n}{n} \right\| = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

Definition 24.0.0.4

An unordered sum $\sum_{\alpha \in \mathcal{J}} x_\alpha$ is Cauchy provided that for all $\epsilon > 0$, there exists finite $J_\epsilon \subseteq \mathcal{J}$ such that for all finite $K \subseteq \mathcal{J} \setminus J_\epsilon$, we have $\|\sum_{\alpha \in K} x_\alpha\| < \epsilon$.

Example: Consider the series defined in (6.1). For all $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that we have the following holds:

$$\sum_{n > N_\epsilon} \frac{1}{n^2} < \epsilon$$

For all finite $J \subseteq \mathbb{N}$ such that $J_\epsilon = \{1, 2, \dots, N_\epsilon\} \subseteq J$, we can write the following:

$$\left\| \sum_{n \in J \setminus J_\epsilon} \frac{x_n}{n} \right\|^2 \leq \sum_{n \in J \setminus J_\epsilon} \left\| \frac{x_n}{n} \right\|^2 = \sum_{n \in J \setminus J_\epsilon} \frac{1}{n^2} < \sum_{n > N_\epsilon} \frac{1}{n^2} < \epsilon$$

this shows that (6.1) is Cauchy, and hence converges unconditionally but not absolutely.

Proposition 24.0.1

A Cauchy unordered sum can only have countably many nonzero term.

Proof. Take any Cauchy unordered sum $\sum_{\alpha \in \mathcal{J}} x_\alpha$. For some $n \in \mathbb{N}$, we write $\epsilon = \frac{1}{n}$, then there exists finite $J_n \subseteq \mathcal{J}$ such that for all finite $K \subseteq \mathcal{J} \setminus J_n$, we have $\|\sum_{\alpha \in K} x_\alpha\| < 1/n$. Here we define $L = \bigcup_{n=1}^{\infty} J_n$, which is countable. For all $\alpha \in \mathcal{J} \setminus L$, let $K = \{\alpha\}$, we see that we have $\|x_\alpha\| < \frac{1}{n}$ for all n , then $x_\alpha = 0$, which completes the proof. \square

Proposition 24.0.2

An unordered sum in a Hilbert space H is convergent if and only if it is Cauchy.

Proof. Here we will first show that convergent unordered sum is Cauchy, here we denote $x = \sum_{\alpha \in \mathcal{J}} x_\alpha$ for the convergent unordered sum $\sum_{\alpha \in \mathcal{J}} x_\alpha$. From definition, for all $\epsilon > 0$, there exists finite $J_\epsilon \subseteq \mathcal{J}$ such that for all finite $J \subseteq \mathcal{J}$ such that $J_\epsilon \subseteq J$, we have $\|x - \sum_{\alpha \in J} x_\alpha\| < \epsilon$. For all finite $K \subseteq \mathcal{J} \setminus J_\epsilon$, we define $J = K \cup J_\epsilon$. Then we can write the following:

$$\left\| \sum_{\alpha \in K} x_\alpha \right\| = \left\| \sum_{\alpha \in J} x_\alpha - \sum_{\alpha \in J_\epsilon} x_\alpha \right\| \leq \left\| x - \sum_{\alpha \in J} x_\alpha \right\| + \left\| x - \sum_{\alpha \in J_\epsilon} x_\alpha \right\| < 2\epsilon$$

this completes the \Rightarrow direction. Now for the \Leftarrow direction, for some $n \in \mathbb{N}$, we write $\epsilon = \frac{1}{n}$, there exists finite $J_n \subseteq \mathcal{J}$ such that for all finite $K \subseteq \mathcal{J} \setminus J_n$, $\|\sum_{\alpha \in K} x_\alpha\| < 1/n$. Here we can modify J_n to obtain the sequence (J_n) such that $J_n \subseteq J_{n+1}$. For all $m \geq n$, we define $K = J_m \setminus J_n$, we can write:

$$\left\| \sum_{\alpha \in K} x_\alpha \right\| = \left\| \sum_{\alpha \in J_m} x_\alpha - \sum_{\alpha \in J_n} x_\alpha \right\| < \frac{1}{n}$$

hence the sequence (S_n) of points in H defined by $S_n = \sum_{\alpha \in J_n} x_\alpha$ is Cauchy, and so (S_n) converges to some $x \in H$. That is for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $\|S_n - x\| < \epsilon$ for all $n \geq N$. Here we can define $N_\epsilon = \max\{N, \lfloor 1/\epsilon \rfloor\}$, then for all finite $J \subseteq \mathcal{J}$ such that $J_{N_\epsilon} \subseteq J$, we can write the following with $s_J := \sum_{\alpha \in J} x_\alpha$:

$$\|s_J - x\| \leq \|s_J - S_{N_\epsilon}\| + \|S_{N_\epsilon} - x\| \leq \left\| \sum_{\alpha \in J \setminus J_{N_\epsilon}} x_\alpha \right\| + \epsilon < \frac{1}{N_\epsilon} + \epsilon < 2\epsilon$$

this shows that $x = \sum_{\alpha \in \mathcal{J}} x_\alpha$, which completes the proof. \square

Proposition 24.0.3

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \mid \alpha \in \mathcal{J}\}$ be any orthogonal set in H , then we have the followings:

1. The unordered sum $\sum_{\alpha \in \mathcal{J}} x_\alpha$ converges unconditionally if and only if we have the unbounded sum $\sum_{\alpha \in \mathcal{J}} \|x_\alpha\|^2$ converges unconditionally.
2. If $\sum_{\alpha \in \mathcal{J}} x_\alpha$ converges, then we have:

$$\left\| \sum_{\alpha \in \mathcal{J}} x_\alpha \right\|^2 = \sum_{\alpha \in \mathcal{J}} \|x_\alpha\|^2$$

Proof. For all finite $K \subseteq \mathcal{J}$, we have $\|\sum_{\alpha \in K} x_\alpha\|^2 = \sum_{\alpha \in K} \|x_\alpha\|^2$, so we have $\sum_{\alpha \in \mathcal{J}} x_\alpha$ is Cauchy if and only if $\sum_{\alpha \in K} \|x_\alpha\|^2$ is Cauchy, then we see here (1) follows from Proposition 24.0.2. Now suppose we have $x = \sum_{\alpha \in \mathcal{J}} x_\alpha \in H$, as in the proof of Proposition 24.0.2, we know that there exists a sequence (J_n) of subsets of \mathcal{J} such that the sequence (S_n) defined by $S_n = \sum_{\alpha \in J_n} x_\alpha$ converges to x . Now we can write the following:

$$\|x\|^2 = \left\| \lim_{n \rightarrow \infty} S_n \right\|^2 = \lim_{n \rightarrow \infty} \|S_n\|^2 = \lim_{n \rightarrow \infty} \sum_{\alpha \in J_n} \|x_\alpha\|^2 = \sum_{\alpha \in \mathcal{J}} \|x_\alpha\|^2$$

which completes the proof. \square

Theorem 24.1 (Bessel's Inequality)

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \mid \alpha \in \mathcal{J}\}$ be an orthonormal set in H . Then for all $x \in H$, we have the followings:

1.

$$\sum_{\alpha \in \mathcal{J}} |(x_\alpha, x)|^2 \leq \|x\|^2$$

2.

$$x_S := \sum_{\alpha \in \mathcal{J}} (x_\alpha, x) x_\alpha \text{ is convergent}$$

3.

$$x - x_S \perp S$$

Proof. First we will prove (1). Note that for all $J \subseteq \mathcal{J}$, we can write the following:

$$\begin{aligned} \left\| x - \sum_{\alpha \in J} (x_\alpha, x) x_\alpha \right\|^2 &= \left(x - \sum_{\alpha \in J} (x_\alpha, x) x_\alpha, x - \sum_{\alpha \in J} (x_\alpha, x) x_\alpha \right) \\ &= \|x\|^2 - \sum_{\alpha \in J} \overline{(x_\alpha, x)} (x_\alpha, x) - \sum_{\alpha' \in J} (x_{\alpha'}, x) (x, x_{\alpha'}) + \sum_{\alpha \in J} \sum_{\alpha' \in J} \overline{(x_\alpha, x)} (x_{\alpha'}, x) (x_\alpha, x_{\alpha'}) \\ &= \|x\|^2 - 2 \sum_{\alpha \in J} |(x_\alpha, x)|^2 + \sum_{\alpha \in J} |(x_\alpha, x)|^2 \\ &= \|x\|^2 - \sum_{\alpha \in J} |(x_\alpha, x)|^2 \geq 0 \end{aligned}$$

here the last inequality comes from the fact that:

$$\left\| x - \sum_{\alpha \in J} (x_\alpha, x) x_\alpha \right\|^2 \geq 0$$

Now we define $F = \{J \subseteq \mathcal{J} \mid J \text{ is finite}\}$, and we define:

$$M = \sup_{J \in F} \sum_{\alpha \in J} |(x_\alpha, x)|^2 \leq \|x\|^2$$

For all $\epsilon > 0$, we note that there exists $J_\epsilon \in F$ such that we have:

$$M - \epsilon < \sum_{\alpha \in J_\epsilon} |(x_\alpha, x)|^2 \leq M$$

If one takes any finite $J \in \mathcal{J}$ such that $J_\epsilon \subseteq J$, then we know that we have:

$$M - \epsilon < \sum_{\alpha \in J_\epsilon} |(x_\alpha, x)|^2 \leq \sum_{\alpha \in J} |(x_\alpha, x)|^2 \leq M$$

Then we see here for all $\epsilon > 0$, there exists finite $J_\epsilon \in F$ such that for all finite $J \supseteq J_\epsilon$, we have the following holds:

$$\left| M - \sum_{\alpha \in J} |(x_\alpha, x)|^2 \right| < \epsilon \quad \Rightarrow \quad M = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, x)|^2$$

this completes the proof of (1). Note here (2) follows from Proposition 24.0.3 and by using the following fact from (1):

$$\sum_{\alpha \in \mathcal{J}} \|(x_\alpha, x)x_\alpha\|^2 = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, x)|^2 = M \leq \|x\|^2$$

For (3), consider $\gamma \in \mathcal{J}$, then we can write:

$$\begin{aligned} (x - x_S, x_\gamma) &= (x, x_\gamma) - \left(\sum_{\alpha \in \mathcal{J}} (x_\alpha, x)x_\alpha, x_\gamma \right) \\ &= (x, x_\gamma) - \sum_{\alpha \in \mathcal{J}} \overline{(x_\alpha, x)}(x_\alpha, x_\gamma) \\ &= 0 \end{aligned}$$

Here the result follows. □

Orthonormal Bases

Definition 25.0.0.0.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \mid \alpha \in \mathcal{J}\}$ be an orthonormal set in H . The closed linear span of S is defined by the following:

$$[S] := \left\{ \sum_{\alpha \in \mathcal{J}} c_\alpha x_\alpha \mid c_\alpha \in \mathbb{C} \text{ and } \sum_{\alpha \in \mathcal{J}} |c_\alpha|^2 < \infty \right\}$$

Note: Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \mid \alpha \in \mathcal{J}\}$ be an orthonormal set in H , for $J \subseteq \mathcal{J}$ that is finite, we can write:

$$\sum_{\alpha \in J} c_\alpha x_\alpha \in \text{span}(S)$$

where $c_\alpha \in \mathbb{C}$, hence it is easy to see here $[S] = \overline{\text{span}(S)}$.

Moreover, for all $x \in H$, from Theorem 24.1, we see that we have:

$$x_S = \sum_{\alpha \in \mathcal{J}} (x_\alpha, x) x_\alpha \in [S] \quad \quad x - x_S \perp S$$

By the Projection Theorem, we also see that $H = [S] \oplus [S]^\perp$, and x_S is the unique vector in $[S]$ closest to x , that is, we have:

$$\|x - x_S\| = \min_{y \in [S]} \|x - y\|$$

Definition 25.0.0.0.2

An orthonormal set S in a Hilbert space H is said to be complete provided that it satisfies any of the criteria in Theorem 25.1 proposed in the following.

Definition 25.0.0.0.3

An orthonormal basis of H is any complete orthonormal set. ¹

Theorem 25.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \mid \alpha \in \mathcal{J}\}$ be a orthonormal set. The followings are equivalent:

1. $(x_\alpha, x) = 0$ for all $\alpha \in \mathcal{J}$ implies $x = 0$.
2. $x = x_S = \sum_{\alpha \in \mathcal{J}} (x_\alpha, x) x_\alpha$ for all $x \in H$.
3. $\|x\|^2 = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, x)|^2$ for all $x \in H$.
4. $[S] = H$.
5. S is a maximal orthonormal set in H .

Proof. Here (1) implies $S^\perp = \{0\}$, then by Theorem 24.1, we know that $x - x_S \in S^\perp$ hence $x = x_S$, this shows that (1) implies (2). Now we will show that (2) implies (3), here (2) implies $x = x_S = \sum_{\alpha \in \mathcal{J}} (x_\alpha, x) x_\alpha$, then by Proposition 24.0.3, if we have a convergent sum in form of x_S , we can write:

$$\|x\|^2 = \sum_{\alpha \in \mathcal{J}} \|(x_\alpha, x) x_\alpha\|^2 = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, x)|^2$$

¹A proof on Page 140 on *Hunter and Nachtergaele* shows any Hilbert space has an orthonormal basis

Now we will show (3) implies (4). From (3), we know that we have $\|x\|^2 = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, x)|^2$ for all $x \in H$. Here we can take any $\xi \in S^\perp$, we know that $(x_\alpha, \xi) = 0$ for all $\alpha \in \mathcal{J}$, then here we get $\|\xi\|^2 = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, \xi)|^2 = 0$ and hence $\xi = 0$, so we know that $S^\perp = \{0\}$, then by Theorem 24.1, we know that $x = x_S \in [S]$ for all $x \in H$. On the other hand, we have $[S] \subseteq H$, concluding we get $H = [S]$. Now we will show that (4) implies (5). If S were not maximal orthonormal set in H , then there exists $u \in H$ such that we have $u \perp S$ and $\|u\| = 1$, but then we will have $u \in H = [S]$ so $u = u_S = \sum_{\alpha \in \mathcal{J}} (x_\alpha, u)x_\alpha = 0$, which is a contradiction. Finally (5) implies (1) is trivial, the result follows. \square

Theorem 25.2 (Parseval's Theorem)

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let $S = \{x_\alpha \mid \alpha \in \mathcal{J}\}$ be an orthonormal basis. For $x = \sum_{\alpha \in \mathcal{J}} a_\alpha x_\alpha$, and $y = \sum_{\alpha \in \mathcal{J}} b_\alpha x_\alpha$, we have:

$$(x, y) = \sum_{\alpha \in \mathcal{J}} \overline{a_\alpha} b_\alpha$$

Proof. From Theorem 25.1 part (3), we can write the following:

$$\|x + y\|^2 = \sum_{\alpha \in \mathcal{J}} |(x_\alpha, x + y)|^2 = \sum_{\alpha \in \mathcal{J}} |a_\alpha + b_\alpha|^2 \quad \|x - y\|^2 = \sum_{\alpha \in \mathcal{J}} |a_\alpha - b_\alpha|^2$$

Hence combining we get the following:

$$\|x + y\|^2 - \|x - y\|^2 = \sum_{\alpha \in \mathcal{J}} |a_\alpha + b_\alpha|^2 - \sum_{\alpha \in \mathcal{J}} |a_\alpha - b_\alpha|^2 \quad (6.2)$$

expanding the LHS of (6.2) we get:

$$\|x + y\|^2 - \|x - y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re(x, y) - (\|x\|^2 + \|y\|^2 - 2\Re(x, y)) = 4\Re(x, y)$$

Combing with the RHS of (6.2) we get:

$$\|x + y\|^2 - \|x - y\|^2 = 4\Re(x, y) = 4\Re \sum_{\alpha \in \mathcal{J}} \overline{a_\alpha} b_\alpha$$

For the imaginary part, proceed similarly with $\|x + iy\|$ and $\|x - iy\|$ to get the similar result. \square

6.25.1 The Gram-Schmidt Orthogonalization

Let H be a separable Hilbert space, let $V = \{v_n \mid n \in \mathbb{N}\}$ be a set of linearly independent vectors, let $[V]$ denote the smallest subspace of H that is closed and contains V . We can construct orthonormal basis $U = \{u_n \mid n \geq 1\}$ of $[V]$ from V as the following:

$$u_1 = \frac{v_1}{\|v_1\|} \quad u_{n+1} = c_{n+1} \left(u_{n+1} - \sum_{j=1}^n (u_j, v_{n+1}) u_j \right)$$

where $n \geq 1$, and c_{n+1} is defined such that $\|u_{n+1}\| = 1$.

Example: For $w \in C(-1, 1)$ with $w(x) > 0$ for $x \in (-1, 1)$. First we define the following:

$$C_w[-1, 1] := \left\{ f \in C[-1, 1] \mid \int_{-1}^1 w(x) |f(x)|^2 < \infty \right\}$$

Here we can define an inner product on the space $C_w[-1, 1]$:

$$(f, g)_w := \int_{-1}^1 w(x) \bar{f}(x) g(x) dx$$

and the norm:

$$\|f\|_w := \sqrt{\int_{-1}^1 w(x) |f(x)|^2 dx}$$

here the completion of $C_w[-1, 1]$ is $L_w^2[-1, 1]$. For the set of monomials $V = \{x^n \mid n \geq 0\}$ defined on $[-1, 1]$, we see that $[V] = L_w^2[-1, 1]$. If $w(x) = 1$ for all $x \in (-1, 1)$, then the Gram-Schmidt Orthogonalization process gives Legendre polynomial defined on $[-1, 1]$. If $w(x) = \sqrt{1-x^2}$ for $x \in (-1, 1)$, then the Gram-Schmidt Orthogonalization process gives Chebyshev polynomials defined on $[-1, 1]$. If $w(x) = e^{-x^2}$ for $x \in (-1, 1)$, then the Gram-Schmidt Orthogonalization process gives Hermite polynomials.

7 | Bounded Linear Operators on Hilbert Spaces

Duality of Hilbert Spaces

Definition 26.0.0.0.1

A linear functional on a Hilbert space H is a linear map $\psi : H \rightarrow \mathbb{C}$.

Note: For Hilbert space H , $\mathcal{B}(H, \mathbb{C})$ equipped with norm:

$$\|\cdot\| : \mathcal{B}(H, \mathbb{C}) \rightarrow \mathbb{R} \quad \psi \mapsto \sup_{x \in H, x \neq 0} \frac{|\psi(x)|}{\|x\|}$$

defines a Banach space.

Example: Let $(H, (\cdot, \cdot))$ be a Hilbert space. For all $y \in H$, we can associate to it a linear functional $\psi(x) := (y, x)$, and here it is easy to check that we have $\psi \in \mathcal{B}(H, \mathbb{C})$, with $\|\psi\| = \|y\|$. To see that, we write the following:

$$\frac{|\psi(x)|}{\|x\|} = \frac{|(y, x)|}{\|x\|} \leq \frac{\|y\| \cdot \|x\|}{\|x\|} = \|y\| \quad \Rightarrow \quad \|\psi\| \leq \|y\|$$

On the other hand, we can write the following:

$$\psi(y) = (y, y) = \|y\|^2 \quad \Rightarrow \quad \frac{|\psi(y)|}{\|y\|} = \|y\| \quad \Rightarrow \quad \|\psi\| = \|y\|$$

Theorem 26.1 (Riesz Representation Theorem)

Let $(H, (\cdot, \cdot))$ be a Hilbert space and let $\psi \in \mathcal{B}(H, \mathbb{C})$. There exists a unique $y \in H$ such that $\psi(x) = (y, x)$ for all $x \in H$. Moreover, we have $\|y\| = \|\psi\|$.

Proof. There is a trivial case where $\psi = 0$, in which case setting $y = 0$ gets the result. Here we suppose $\psi(x) \neq 0$ for some $x \in H$. Let $V = \ker(\psi)$. Note that V is closed as we have shown in previous chapter. From the Projection Theorem, we know that we have $H = V \oplus V^\perp$, where V^\perp is a closed subspace, and in particular, we have $V^\perp \neq \{0\}$ because $\psi \neq 0$. Then there exists $\eta \in V^\perp - \{0\}$, one can normalize η such that $\psi(\eta) = 1$. Here we define:

$$P : H \rightarrow H \quad x \mapsto \psi(x)\eta$$

Note here P is a bounded linear map as we can write the following for all $x \in H$:

$$\|Px\| = \sup_{x \neq 0} \frac{\|\psi(x)\eta\|}{\|x\|} = \sup_{x \neq 0} \frac{|\psi(x)| \cdot \|\eta\|}{\|x\|} = \|\psi\| \cdot \|\eta\|$$

also, we see that P satisfies the following for all $x \in H$:

$$P^2x = PPx = \psi(Px)\eta = \psi(\psi(x)\eta)\eta = \psi(x)\psi(\eta)\eta = \psi(x)\eta = Px$$

this shows that ψ is a projection. From Theorem 23.4, we can write:

$$H = \text{im}(P) \oplus \ker(P)$$

Note that the image of P is given by $\{x \in H \mid x \in Px = \psi(x)\eta\}$, hence we see that the image of P is $\text{span}(\{\eta\})$, which is an 1-dimensional space, and hence is closed. On the other hand, $\ker(P)$ is $\{x \in H \mid Px = \psi(x)\eta = 0\} = \ker(\psi) = V$. That is, we must have $\text{im}(P) = V^\perp$. Now we can write, for all $x \in H$, we have $x = \alpha\eta + \xi$ for some scalar α and some $\xi \in V$. Here we define:

$$y = \frac{\eta}{\|\eta\|^2}$$

here we see that we have:

$$(y, x) = \left(\frac{\eta}{\|\eta\|^2}, \alpha\eta + \xi \right) = \alpha$$

On the other hand, we can write:

$$\psi(x) = \psi(\alpha\eta + \xi) = \alpha\psi(\eta) + \psi(\xi) = \alpha$$

Here we will show that y is unique, if one has y, y' satisfy $\psi(x) = (y, x) = (y', x)$, then we must have $(y - y', x) = 0$ for all $x \in H$, which implies $y = y'$ if one compute $(y - y', y - y')$. Note that $\|y\| = \|\psi\|$ follows from previous argument. Here we have completed the proof. \square

Definition 26.1.0.0.1

Let H be a Hilbert space, $\mathcal{B}(H, \mathbb{C})$ is called the topological dual of H , denoted as H^* .

Note: Let H be a Hilbert space, we can define:

$$J : H \rightarrow H^* \quad y \mapsto (y, \cdot)$$

where it is easy to see that J defines a bijection. Also note that we have:

$$\|J(y)\| = \|y\|$$

hence we know that J is an isometry, defines an isomorphism between H and H^* .

Weak Convergence

Definition 27.0.0.0.1

A sequence (x_n) of points in a Hilbert space H converges weakly to $x \in H$, denoted as $(x_n) \rightharpoonup x$ provided that we have the following holds for all $y \in H$:

$$\lim_{n \rightarrow \infty} (y, x_n) = (y, x)$$

Note: From Definition 27.0.0.0.1, by Riesz Representation Theorem, we have $(x_n) \rightharpoonup x$ if we have $\phi(x_n) \rightarrow \phi(x)$ for all $\phi \in H^*$.

Note: Weak convergence is unique. Suppose we have $(x_n) \rightharpoonup x$ and $(x_n) \rightharpoonup x'$, then we can write $(y, x) = (y, x')$ for all $y \in H$. Putting $y = x - x'$, we must have $\|x - x'\| = 0$ and hence $x = x'$.

Theorem 27.1

Let $(H, (\cdot, \cdot))$ be a Hilbert space, and let D be a dense set in H . A sequence (x_n) of points in H converges weakly to $x \in H$ if and only if we have the followings hold:

1. There exists $M > 0$ such that $\|x_n\| < M$.
2. The sequence $((y, x_n))$ converges to (y, x) for all $y \in D$.

Proof. For the \Rightarrow direction, suppose we have (x_n) converges weakly to x . By definition, we have $(y, x_n) \rightarrow (y, x)$ for all $y \in H$, and in particular, for $y \in D$, hence it suffices to show that (x_n) is bounded. For all x_n , we can associate $\psi_n := (x_n, \cdot) : H \rightarrow \mathbb{R} \quad x \mapsto (x_n, x)$, here we see that $\psi_n \in H^*$. Take $z \in H$, we can write the following:

$$\psi_n(z) = (x_n, z) \rightarrow (x, z)$$

hence we see here the sequence (ψ_n) of point in H^* converges pointwise, then by the Banach-Steinhaus Theorem, there exists $M > 0$ such that $\sup_{n \geq 1} \|\psi_n\| = M$, and hence we can write $\|\psi_n\| = \|x_n\| \leq M$ for all n . For the \Leftarrow direction, suppose we have (1) and (2) holds. Let $\epsilon > 0$ be given, and let $z \in H$, we want to show that we have $|(z, x) - (z, x_n)| < \epsilon$ for large enough n . Since D is dense, then there exists $y \in D$ such that $\|z - y\| < \epsilon$. We also know that there exists $N \in \mathbb{N}$ such that we have $|(x_n, y) - (x, y)| < \epsilon$ for all $n \geq N$. Here we can write the following for large enough n :

$$\begin{aligned} |(x_n - x, z)| &\leq |(x_n - x, y)| + |(x_n - x, z - y)| \\ &< \epsilon + \|x_n - x\| \cdot \|z - y\| \\ &< \epsilon + (\|x_n\| + \|x\|)\epsilon \\ &< \epsilon(1 + \|x\| + M) \end{aligned}$$

which proves the result. □

Theorem 27.2

Let (x_n) be a sequence of points in a Hilbert space $(H, (\cdot, \cdot))$ such that $(x_n) \rightarrow x$. Then we have $(x_n) \rightharpoonup x$.

Proof. For all $z \in H$, we can write the following:

$$|(x_n, z) - (x, z)| = |(x_n - x, z)| \leq \|x_n - x\| \cdot \|z\| \tag{7.1}$$

and the last term in (7.1) converges to 0 for large enough n . The result follows. □

Note: The converse of Theorem 27.2 is not true.

Theorem 27.3

Let $(H, (\cdot, \cdot))$ be a Hilbert space of infinite dimensional. Let (e_n) be orthonormal sequence in H , then we have $(e_n) \rightharpoonup 0$.

Proof. Let $y \in H$, by Bessel Inequality, we can write the following:

$$\sum_{n=1}^{\infty} |(e_n, y)|^2 \leq \|y\|^2$$

hence we must have $(e_n, y) \rightarrow 0$. This completes the proof. \square

Example: Consider the sequence of functions (f_n) in $H = L^2[0, 1]$ defined by $f_n(x) = \sin(n\pi x)$ for all $x \in [0, 1]$. Here we can write the following:

$$\|f_n\| = \left(\int_0^1 \sin^2(n\pi x) dx \right)^{1/2} = \frac{1}{\sqrt{2}}$$

hence (f_n) is a bounded sequence. We will show that (f_n) converges to 0. According to Theorem 27.1, it suffices to show that we have $((f_n, p))$ converges to 0 for all polynomial p , as the set of polynomials is dense in $L^2[0, 1]$. Here we can write the following:

$$\begin{aligned} (f_n, p) &= \int_0^1 \sin(n\pi x) p(x) dx = - \int_0^1 p(x) \left(\frac{\cos(n\pi x)}{n\pi} \right)' dx \\ &= - \frac{1}{n\pi} \cos(n\pi x) p(x) \Big|_0^1 + \frac{1}{n\pi} \int_0^1 \cos(n\pi x) p'(x) dx \\ &= \frac{p(0) - (-1)^n p(1)}{n\pi} + \frac{1}{n\pi} \int_0^1 \cos(n\pi x) p'(x) dx \end{aligned}$$

hence we can write:

$$|(f_n, p)| \leq \frac{|p(0)| + |p(1)|}{n\pi} + \frac{1}{n\pi} \int_0^1 |p'(x)| dx \quad (7.2)$$

and for large enough n , from (7.2), we see that we have:

$$\lim_{n \rightarrow \infty} |(f_n, p)| = 0$$

This shows that (f_n) is weak convergent.

Example: Consider we have a sequence of functions (f_n) where $f_n \in H = L^2[0, 1]$ is defined by the following for $x \in [0, 1]$:

$$f_n(x) = \begin{cases} \sqrt{n} & x \in [0, 1/n] \\ 0 & 1/n < x \leq 1 \end{cases}$$

Here we can write the following:

$$\|f_n\| = \left(\int_0^1 |f_n(x)|^2 dx \right)^{1/2} = \left(\int_0^{1/n} n dx \right)^{1/2} = 1$$

Continuous function is dense in $L^2[0, 1]$, that is, $D = C[0, 1]$ is dense in $L^2[0, 1]$. For $g \in C[0, 1]$, we can write:

$$|(f_n, g)| = \left| \int_0^1 f_n(g)g(x) dx \right| = \left| \int_0^{1/n} \sqrt{n}g(x) dx \right| \leq \sqrt{n} \int_0^{1/n} |g| dx \leq \frac{C}{\sqrt{n}}$$

where we see that the last term converges to 0 for large n , hence we see here $(f_n) \rightarrow 0$.

Example: Consider a sequence (f_n) of functions in $H = L^2(\mathbb{R})$ with f_n being defined by the following for $x \in \mathbb{R}$:

$$f_n(x) = \begin{cases} 1 & x \in (n, n + 1/2) \\ 0 & \text{otherwise} \end{cases}$$

Here we can write the following:

$$\|f_n\| = \left(\int_n^{n+1/2} 1 dx \right)^{1/2} = \frac{1}{\sqrt{2}}$$

Now we consider the set of functions of compact support, denoted as D . It is not hard to show that D is dense in $L^2(\mathbb{R})$, and for $g \in D$, we can write the following:

$$(f_n, g) = \int_n^{n+1/2} g(x) dx \quad (7.3)$$

and here it is easy to see that RHS of (7.2) converges to 0 for large enough n as the support of g is compact. This shows that $(f_n) \rightarrow 0$.

Theorem 27.4

Let (u_n) be a bounded sequence in a Hilbert space $(H, (\cdot, \cdot))$, there exists a weakly convergent subsequence.

Proof. Assume the case where $\dim(H) = \infty$. For finite dimension of H , the result follows from the Heine-Borel Theorem. We also assume that H is separable, the result also holds for non-separable H , but the proof is rather complicated. Let (e_n) be an orthonormal basis of H . By assumption, there exists $M > 0$ such that $\|u_n\| \leq M$ for all $n \in \mathbb{N}$. Here we note that $((u_n, e_1))$ is a sequence of complex number in \mathbb{C} , and it satisfies the following:

$$|(u_n, e_1)| \leq \|u_n\| \cdot \|e_1\| \leq M$$

Hence by Bolzano–Weierstrass Theorem, we know that there exists a subsequence $(u_m^{(1)})$ such that we have $(u_n^{(1)}, e_1)$ converges to $\xi_1 \in \mathbb{C}$. Similarly, by considering the sequence $((u_n^{(1)}, e_2))$, we get a subsequence $(u_n^{(2)})$ such that $((u_n^{(2)}, e_2))$ converges to $\xi_2 \in \mathbb{C}$. Concluding, we get a subsequence $(u_n^{(j)})_{n \geq 1}$ such that we have $((u_n^{(j)}, e_j))_{n \geq 1}$ converges to $\xi_j \in \mathbb{C}$. Here we consider the sequence (v_n) defined by $v_n = u_n^{(n)}$. It is not hard to see here we have $((v_n, e_j))_{n \geq 1}$ converges to ξ_j because $(v_n)_{n \geq j}$ is simply a subsequence of $(u_n^{(j)})_{n \geq 1}$ which has the property that $(u_n^{(j)}, e_j)_{n \geq 1}$ converges to ξ_j . Now we will show

that (v_n) converges weakly, that is, there exists some $v \in H$ such that $((v_n, a))_{n \geq 1}$ converges to (v, a) for all $a \in H$. For all $a \in H$, we can write the following by Theorem 25.1:

$$a = \sum_{j=1}^{\infty} (e_j, a) e_j$$

hence for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that we have the following for all $n \geq N$:

$$\left\| a - \sum_{j=1}^n (e_j, a) e_j \right\| < \epsilon$$

Here we will show that $((v_n, a))$ is a Cauchy sequence in \mathbb{C} . Define a sequence $(\beta_n^N)_{n \geq 1}$ in \mathbb{C} by the following:

$$\beta_n^N = \left(v_n, \sum_{j=1}^N (e_j, a) e_j \right) = \sum_{j=1}^N (e_j, a) (v_n, e_j)$$

hence we see here:

$$\lim_{n \rightarrow \infty} \beta_n^N = \sum_{j=1}^N (e_j, a) \xi_j$$

Hence we can write the following for large enough $m, n \in \mathbb{N}$:

$$\begin{aligned} & |(v_n, a) - (v_m, a)| \\ &= \left| \left(v_n, \sum_{j=1}^N (e_j, a) e_j + \sum_{j=N+1}^{\infty} (e_j, a) e_j \right) - \left(v_m, \sum_{j=1}^N (e_j, a) e_j + \sum_{j=N+1}^{\infty} (e_j, a) e_j \right) \right| \\ &= \left| \beta_n^N + \left(v_n, \sum_{j \geq N+1} (e_j, a) e_j \right) - \beta_m^N - \left(v_m, \sum_{j \geq N+1} (e_j, a) e_j \right) \right| \\ &\leq |\beta_n^N - \beta_m^N| + \left| \left(v_n, \sum_{j \geq N+1} (e_j, a) e_j \right) \right| + \left| \left(v_m, \sum_{j \geq N+1} (e_j, a) e_j \right) \right| \\ &\leq |\beta_n^N - \beta_m^N| + (\|v_n\| + \|v_m\|) \left\| \sum_{j \geq N+1} (e_j, a) e_j \right\| \end{aligned}$$

Now it is easy to see here $|(v_n, a) - (v_m, a)|$ is small as we know that the term $\|v_n\| + \|v_m\|$ is bounded by $2\|M\|$, and the term $\left| \sum_{j \geq N+1} (e_j, a) e_j \right|$ is bounded by ϵ . This shows that $((v_n, a))$ is a Cauchy sequence and hence it converges. Here we define ϕ as the following:

$$\phi : H \rightarrow \mathbb{C} \quad a \mapsto \lim_{n \rightarrow \infty} (v_n, a)$$

First we note that ϕ is linear, for all $a, a' \in H$, and all $\alpha, \alpha' \in \mathbb{C}$, we see here we have:

$$\begin{aligned} \phi(\alpha a + \alpha' a') &= \lim_{n \rightarrow \infty} (v_n, \alpha a + \alpha' a') \\ &= \alpha \lim_{n \rightarrow \infty} (v_n, a) + \alpha' \lim_{n \rightarrow \infty} (v_n, a') \\ &= \alpha \phi(a) + \alpha' \phi(a') \end{aligned}$$

On the other hand, we can write the following:

$$|\phi(a)| = \left| \lim_{n \rightarrow \infty} \phi(a) \right| = \left| \lim_{n \rightarrow \infty} (v_n, a) \right| \leq M \|a\| \quad \Rightarrow \quad \|\phi\| \leq M$$

Hence we see that ϕ is a bounded linear operator, and hence by the Riesz Representation Theorem, there exists $v \in H$ such that $\forall a \in H$, we have $\phi(a) = (v, a)$. Here the result of the theorem follows. \square

Adjoint Operator

Theorem 28.1

Let $(X, (\cdot, \cdot)_X)$ and $(Y, (\cdot, \cdot)_Y)$ be Hilbert spaces, let $A \in \mathcal{B}(X, Y)$.

We have the followings hold:

1. There exists a unique operator $A^* \in \mathcal{B}(Y, X)$, called the adjoint operator of A , that satisfies the following for all $x \in X$ and all $y \in Y$:

$$(Ax, y)_Y = (x, A^*y)_X$$

2. The adjoint operator A^* of A satisfies $\|A^*\|_{\mathcal{B}(Y, X)} = \|A\|_{\mathcal{B}(X, Y)}$
3. We have $(\text{im}(A))^\perp = \ker(A^*)$ and $(\text{im}(A^*))^\perp = \ker(A)$.

Note: Here (3) implies that we can write the following by the Projection Theorem:

$$Y = \ker(A^*) \oplus \overline{\text{im}(A)} \quad X = \ker(A) \oplus \overline{\text{im}(A^*)}$$

Remark: If A were unbounded, determining the adjoint of A is more complicated.

Proof of Theorem 28.1. Take arbitrary $y \in Y$, we define the following:

$$l_y : X \rightarrow \mathbb{C} \quad x \mapsto (y, Ax)_Y$$

Here we claim that $l_y \in X^*$, as we can easily check that l_y is linear and we have the following holds for all $x \in X$:

$$|l_y(x)| \leq \|y\|_Y \cdot \|Ax\|_Y \leq \|y\|_Y \cdot \|A\|_{\mathcal{B}(X, Y)} \cdot \|x\|_X$$

which implies that we have:

$$\frac{|l_y(x)|}{\|x\|_X} \leq \|A\|_{\mathcal{B}(X, Y)} \cdot \|y\|_Y$$

and we see here l_y is bounded, so $l_y \in X^*$. By the Riesz Representation Theorem, there exists a unique $\xi_y \in X$ such that we have:

$$l_y(x) = (y, Ax)_Y = (\xi_y, x)_X$$

Now we define $A^* : Y \rightarrow X$ as the following:

$$A^* : Y \rightarrow X \quad y \mapsto \xi_y \tag{7.4}$$

where ξ_y is defined above uniquely for each $y \in Y$. Hence we can write the following:

$$l_y(x) = (y, Ax)_Y = (\xi_y, x)_X = (A^*y, x)_X$$

Now suppose we have A^* and \tilde{A}^* that satisfy (7.4), then we can write the following for all $x \in X$ and all $y \in Y$:

$$(y, Ax)_Y = (A^*y, x)_X = (\tilde{A}^*y, x)_X$$

hence we have the following for all $x \in X$ and all $y \in Y$:

$$(A^*y - \tilde{A}^*y, x) = 0 \tag{7.5}$$

now we can choose $x = A^*y - \tilde{A}^*y$, in which case (7.5) implies that we have:

$$A^*y = \tilde{A}^*y \quad \forall y \in Y$$

hence we see here $A^* = \tilde{A}^*$. To show that A^* is linear, we take any $y, y' \in Y$ and $\alpha, \alpha' \in \mathbb{C}$, we can write the following for all $x \in X$:

$$\begin{aligned} (x, A^*(\alpha y + \alpha' y'))_X &= (Ax, \alpha y + \alpha' y')_Y = \alpha(Ax, y)_Y + \alpha'(Ax, y')_Y \\ &= \alpha(x, A^*y)_X + \alpha'(x, A^*y')_X = (x, \alpha A^*y + \alpha' A^*y')_X \end{aligned}$$

hence it is immediate that we have:

$$(x, A^*(\alpha y + \alpha' y') - \alpha A^*y - \alpha' A^*y')_X = 0$$

choosing $x = A^*(\alpha y + \alpha' y') - \alpha A^*y - \alpha' A^*y'$, we obtain:

$$A^*(\alpha y + \alpha' y') = \alpha A^*y + \alpha' A^*y'$$

Now we will check the boundedness of A^* , here we can write the following for all $y \in Y$:

$$\|A^*y\|_X^2 = (A^*y, A^*y)_X = (y, AA^*y)_Y \leq \|y\|_Y \cdot \|AA^*y\|_Y \leq \|y\|_Y \cdot \|A\|_{\mathcal{B}(X,Y)} \cdot \|A^*y\|_X$$

hence for $y \notin \ker(A^*)$, we can write:

$$\|A^*y\|_X \leq \|y\|_Y \cdot \|A\|_{\mathcal{B}(X,Y)} \quad \Rightarrow \quad \frac{\|A^*y\|_X}{\|y\|_Y} \leq \|A\|_{\mathcal{B}(X,Y)}$$

which shows that A^* is a bounded linear operator. Moreover, we have checked that $\|A^*\|_{\mathcal{B}(Y,X)} \leq \|A\|_{\mathcal{B}(X,Y)}$. Note here we can also write the following for all $x \in X$:

$$\|Ax\|_Y^2 = (Ax, Ax)_Y = (x, A^*Ax)_X \leq \|x\|_X \cdot \|A^*Ax\|_X \leq \|x\|_X \cdot \|A^*\|_{\mathcal{B}(Y,X)} \cdot \|Ax\|_Y$$

and it is easy to see here, for $x \notin \ker(A)$, we have:

$$\frac{\|Ax\|_Y}{\|x\|_X} \leq \|A^*\|_{\mathcal{B}(Y,X)}$$

So far we have proved (1) and (2). (3) is just straightforward computation as we will show next. Note that we can write the following:

$$\begin{aligned} (\operatorname{im}(A))^\perp &= \{y \in Y \mid (z, y)_Y = 0 \ \forall z \in \operatorname{im}(A)\} \\ &= \{y \in Y \mid (Ax, y)_Y = 0 \ \forall x \in X\} \\ &= \{y \in Y \mid (x, A^*y)_X = 0 \ \forall x \in X\} \end{aligned}$$

hence if $y \in \ker(A^*)$, we see here $y \in (\operatorname{im}(A))^\perp$. On the other hand, for $y \in (\operatorname{im}(A))^\perp$, we can set $x = A^*y$ in which case we get $\|A^*y\|_X = 0$ and hence $y \in \ker(A^*)$, so we can write $(\operatorname{im}(A))^\perp = \ker(A^*)$. The proof for $(\operatorname{im}(A^*))^\perp = \ker(A)$ employs similar argument and is left for the reader. This completes the proof of the theorem. \square

Example: Consider $(X, (\cdot, \cdot)_X)$ and $(Y, (\cdot, \cdot)_Y)$ being Hilbert spaces, and let $A : \in \mathcal{B}(X, Y)$. We are interested in the system $Ax = b$ for $b \in Y$ and we want to know when A has a unique solution $x \in X$. Immediately we see here x exists if $b \in \operatorname{im}(A)$. If $\operatorname{im}(A)$ is closed, then we can check where $b \perp \ker(A^*)$ to see if b exists.

The Fredholm Theory

In this section, we will consider that we have $X = Y = H$ being a Hilbert space, and we denote $\mathcal{B}(H, H) = \mathcal{B}(H)$.

Definition 29.0.0.0.1

For $A \in \mathcal{B}(H)$, we say that A satisfies the Fredholm alternative provided we have one of the followings holds:

1. The systems $Ax = 0$ and $A^*x = 0$ have only solution $x = 0$, and the systems $Ax = b$, $A^*x = b$ have unique solutions $\forall b \in H$.
2. $\ker(A)$ and $\ker(A^*)$ are nontrivial finite dimensional of some dimension, and $Ax = b$ has solution if $b \perp \ker(A^*)$, $A^*x = b$ has solution if $b \perp \ker(A)$.

Example: Consider the case we have $\dim(H) = n$.

Any $A \in \mathcal{B}(H)$ can be represented by a matrix $\mathbb{A} \in \mathbb{C}^{n \times n}$. Here we notice that $\text{im}(A)$ and $\text{im}(A^*)$ are closed because they are subspaces of finite dimensional space H . Moreover, $\ker(A)$ and $\ker(A^*)$ are null spaces of \mathbb{A} and $\mathbb{A}^* = \overline{\mathbb{A}}^T$. Hence it is easy to see here A is Fredholm from elementary results from linear algebra.

Example: Consider the case where $H = L^2[0, 1]$, and $A : H \rightarrow H$ being an operator defined by $Af(x) = xf(x)$ for all $x \in [0, 1]$. First we will check that A is bounded:

$$\|Af\| = \left(\int_0^1 x^2 |f(x)|^2 dx \right)^{1/2} \leq \|f\| \quad \Rightarrow \quad \|A\| \leq 1$$

We will show here $A^* = A$. Note that for all $f, g \in H$, we can write:

$$(g, Af) = \int_0^1 \overline{g(x)} (xf(x)) dx = \int_0^1 \overline{(xg(x))} f(x) dx = (Ag, f)$$

On the other hand, we see here:

$$\ker(A) = \{f \in H \mid xf(x) = 0 \text{ almost everywhere}\}$$

the image of A is not closed, here we can take b being the identity transformation, and the system $Af = b$ has no solution in H because the function characterized by $f(x) = 1/x$ does not belong to H .

Example: Consider the case where we have $H = l^2$. let $A : H \rightarrow H$ $\xi = (\xi_n)$ defined by, for $\xi = (\xi_1, \xi_2, \dots) \in l^2$, $A\xi = (0, \xi_1, \xi_2, \xi_3, \dots)$. Let $\eta = (\eta_n) \in H$, we can write:

$$(\eta, A\xi) = \sum_{n=1}^{\infty} \bar{\eta}_n (A\xi)_n$$

where $(A\xi)_1 = 0$ and $(A\xi)_i = \xi_{i-1}$ for $i > 1$. Hence we can write:

$$(\eta, A\xi) = \sum_{n=2}^{\infty} \bar{\eta}_n \xi_{n-1} = \sum_{j=1}^{\infty} \bar{\eta}_{j+1} \xi_j = (A^* \eta, \xi)$$

that is, we see here $A^*\eta = (\eta_2, \eta_3, \eta_4, \dots)$.

On the other hand, we see here $\ker(A) = \{0\}$, and we have the following:

$$\ker(A^*) = \text{span}\{(1, 0, 0, \dots)\}$$

This shows that A is not Fredholm. Consider the system $A\xi = \beta = (\eta_n)$, we can write:

$$A\xi = (0, \xi_1, \xi_2, \dots) \quad \Rightarrow \quad \beta_1 = 0 \quad \Rightarrow \quad \beta \in (\ker(A^*))^\perp = \overline{\text{im}(A)}$$

On the other hand, we can consider the system $A^*\xi = \beta = (\xi_2, \xi_3, \dots)$, where we see that $\text{im}(A^*) = H$.

8 | Spectrum of Operators

Operators on Finite Dimensional Spaces

Consider linear operator $T : X \rightarrow X$ where X is a normed vector space. Here we begin with the case where X is finite dimensional, we note that T can be represented by matrices and we can apply the spectral theory on T from elementary linear algebra. Here we denote $\dim(X) = n$, then T can be represented by some $n \times n$ matrix. Let $\{b_1, b_2, \dots, b_n\}$ be a basis of X , then for all $x \in X$, we can write:

$$x = \sum_{j=1}^n \xi_j b_j \quad \text{for some } \xi_j \in \mathbb{C}$$

Here we can also write:

$$Tx = \sum_{j=1}^n \xi_j T b_j = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} \xi_j \right) b_i$$

where we can define a matrix A with entries a_{ij} , and get the following:

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = A \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = A \vec{\xi}$$

here $\vec{\xi}$ and \vec{y} are coordinates of x and Tx , respectively, and here a_{ij} is the i -th component of Tb_j in basis $\{b_1, b_2, \dots, b_n\}$.

Definition 30.0.0.0.1

Let X be an n -dimensional normed space and let $T : X \rightarrow X$ be a linear operator. $\lambda \in \mathbb{C}$ is called an eigenvalue of T provided that $\ker(T - \lambda I) \neq \{0\}$. That is, λ is an eigenvalue of T provided that there exists $x \in X \setminus \{0\}$ such that $Tx = \lambda x$, in which case x is called an eigenvector of T corresponds to λ .

Definition 30.0.0.0.2

Let X be an n -dimensional normed space and let $T : X \rightarrow X$ be a linear operator.

$$\sigma(T) := \{\lambda \in \mathbb{C} \mid \lambda \text{ is an engenvalue of } T\} := \text{the spectrum of } T$$

$$\rho(T) := \mathbb{C} \setminus \sigma(T) := \text{the resolvent set of } T$$

Note: In the Definition 30.0.0.0.2, if $\lambda \in \rho(T)$, then $T - \lambda I$ is invertible.

Let X be a normed space and let $T : X \rightarrow X$ be a linear operator. Since T is represented by a matrix $A \in \mathbb{C}^{n \times n}$ for a given basis $\{b_1, b_2, \dots, b_n\}$ of the normed space X . We can use linear algebra to compute the result. A here has n eigenvalues, which are the roots of the characteristic polynomial $\det(A - \lambda I) = 0$. If $A\vec{\xi} = \lambda\vec{\xi}$ for some $\lambda \in \mathbb{C}$, where $\vec{\xi}$ is a coordinate, then $\vec{\xi}$ is an eigenvector of A . Here we can also write:

$$x = \sum_{j=1}^n \xi_j b_j \quad \text{satisfies} \quad Tx = \sum_{j=1}^n \xi_j T b_j = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} \xi_j \right) b_i = \lambda x$$

thus $\lambda \in \sigma(T)$ and x is the eigenvector of T with coordinate $\vec{\xi}$.

Note: The matrix representation of T is not unique, but it is unique for a given basis $\{b_1, b_2, \dots, b_n\}$ of X .

Theorem 30.1

Let X be an n -dimensional normed space and let $T : X \rightarrow X$ be a linear operator. All matrices representing the linear operator T have the same eigenvalues.

Proof. Consider two bases $\mathbb{B}_1 = \{b_1, b_2, \dots, b_n\}$ and $\mathbb{B}_2 = \{\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_n\}$ for the space X . and let A be the matrix representation of T with respect to \mathbb{B}_1 , and \tilde{A} be the matrix representation of T with respect to \mathbb{B}_2 . We claim that A and \tilde{A} have the same eigenvalues. Here we can write the following for $j \in \{1, 2, \dots, n\}$:

$$\tilde{b}_j = \sum_{i=1}^n \beta_{ij} b_i$$

with \tilde{A} being defined by the following:

$$T\tilde{b}_j = \sum_{i=1}^n \tilde{a}_{ij} \tilde{b}_i$$

On the other hand, we can write:

$$T\tilde{b}_j = T \sum_{l=1}^n \beta_{lj} b_l = \sum_{l=1}^n \beta_{lj} T b_l = \sum_{l=1}^n \beta_{lj} \sum_{r=1}^n a_{rl} b_r = \sum_{i=1}^n \tilde{a}_{ij} \sum_{r=1}^n \beta_{ri} b_r = \sum_{i=1}^n \tilde{a}_{ij} \tilde{b}_i$$

From linear independency, we get the following:

$$\sum_{l=1}^n a_{rl} \beta_{lj} = \sum_{i=1}^n \beta_{ri} \tilde{a}_{ij} \quad \text{for all } r, j \in \{1, 2, \dots, n\}$$

Here we can define a matrix B with entries β_{lj} . Then it follows that we have $AB = B\tilde{A}$, and here B is invertible because any pair of bases can be written in terms of one another. Hence $B\tilde{A} = AB$ implies $\tilde{A} = B^{-1}AB$. Finally, we get the following:

$$\det(\tilde{A} - \lambda I) = \det(B^{-1}(A - \lambda I)B) = \det(A - \lambda I)$$

in which case we see that the result follows. \square

Operators on Infinite Dimensional Spaces

Consider a bounded linear operator $T : H \rightarrow H$ where $\dim(H) = \infty$.

Definition 31.0.0.0.1

Let H be a Hilbert space, and let $T : H \rightarrow H$ be a bounded linear operator. The resolver set of T is defined to be:

$$\rho(T) := \{\lambda \in \mathbb{C} \mid \text{the function } T_\lambda : H \rightarrow H \quad v \mapsto (T - \lambda I)v \text{ is invertible}\}$$

For $\lambda \in \rho(T)$, $R_\lambda := T_\lambda^{-1}$ is called the resolvent.

Note: In Definition 31.0.0.0.2, according to the Open Mapping Theorem, R_λ is a linear bounded operator.

Definition 31.0.0.0.2

Let H be a Hilbert space, and let $T : H \rightarrow H$ be a bounded linear operator. The spectrum of T is defined by $\sigma(T) := \mathbb{C} \setminus \rho(T)$.

Definition 31.0.0.0.3

Let H be a Hilbert space, and let $T : H \rightarrow H$ be a bounded linear operator. The point spectrum of T is defined by $\sigma_p(T) := \{\lambda \in \mathbb{C} \mid T_\lambda \text{ is not injective}\}$.

Note: Here we see from Definition 31.0.0.3 that, if $\ker(T_\lambda) \neq \{0\}$, then for all $x \in \ker(T_\lambda)$, we have $T_\lambda x = Tx - \lambda x = 0$ for $x \neq 0$, hence x is an eigenvector of T and λ is the corresponding eigenvalue.

Definition 31.0.0.0.4

Let H be a Hilbert space, and let $T : H \rightarrow H$ be a bounded linear operator. The continuous spectrum of T is defined by:

$$\sigma_c(T) = \{\lambda \in \mathbb{C} \mid T_\lambda \text{ is injective but not surjective, with } \overline{\text{im}(T_\lambda)} = H\}$$

The residual spectrum of T is defined by:

$$\sigma_r(T) := \{\lambda \in \mathbb{C} \mid T_\lambda \text{ is injective but } T_\lambda(H) \text{ is not dense in } H\}$$

Note: From here, we see clearly that we have $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$ for bounded linear operator T defined on a Hilbert space H . In particular, if H is of finite dimensional, then we have $\sigma(T) = \sigma_p(T)$ with $\sigma_c(T) = \sigma_r(T) = \emptyset$.

Example: Consider $H = L^2[0, 1]$, and $T : H \rightarrow H$ defined by $T(f(x)) = x \cdot f(x)$. Here $T \in \mathcal{B}(H)$, and $\|T\| \leq 1$, T is self-adjoint. Note here $\sigma_p(T) = \emptyset$, that is T_λ is injective for all $\lambda \in \mathbb{C}$. If λ were eigenvalue of T , then there exists $f \neq 0$ such that we can write:

$$T_\lambda f = Tf - \lambda f = xf - \lambda f = 0 \quad \forall x \in [0, 1]$$

Clearly no $f \neq 0$ can satisfy such condition. We will here show that $\rho(T) = \mathbb{C} \setminus [0, 1]$. For all $\lambda \notin [0, 1]$, and for all $g \in H$, we see that $f = g/(x - \lambda) \in H$, for which satisfies:

$$T_\lambda f = Tf - \lambda f = xf - \lambda f = \frac{xg - \lambda g}{x - \lambda} = g$$

and here we can write:

$$\int_0^1 |f|^2 dx = \int_0^1 \frac{|g|^2}{|x - \lambda|^2} dx < \infty$$

so we see here f is bounded and we have $T_\lambda^{-1}g = f \in \mathcal{B}(H)$, where λ is arbitrary in $\mathbb{C} \setminus [0, 1]$, then this shows that we have $\mathbb{C} \setminus [0, 1] \subseteq \rho(T)$. On the other hand, consider $\lambda \in [0, 1]$, we see here the constant function $g = 1 \notin \text{im}(T_\lambda)$ because $f = 1/(x - \lambda) \notin H$ and hence $\lambda \notin \rho(T)$. This shows that $\sigma(T) = [0, 1]$. Consider $\lambda \in [0, 1]$ and $g \in H$, we can define the following sequence:

$$g_n(x) := \begin{cases} g(x) & |x - \lambda| \geq 1/n \\ 0 & \text{otherwise} \end{cases}$$

Here we write $T_\lambda f_n = g_n$, where we define:

$$f_n(x) = \frac{g_n(x)}{x - \lambda} = \frac{g(x) (1 - \mathbb{I}_{|x - \lambda| < 1/n}(x))}{x - \lambda}$$

where we also see that we have:

$$\int_0^1 |f_n|^2 dx = \int_0^1 \frac{|g(x)|^2 (1 - \mathbb{I}_{|x - \lambda| < 1/n}(x))^2}{(x - \lambda)^2} dx < \infty$$

Hence (g_n) is a sequence of points in $\text{im}(T_\lambda)$. Moreover, we can write the following:

$$\|g - g_n\|^2 = \int_0^1 |g(x)|^2 \mathbb{I}_{|x - \lambda| < 1/n}(x) dx \quad (8.1)$$

here the RHS of (8.1) approaches 0 when n is large enough, hence we see that $(g_n) \rightarrow g$, in which case we conclude that $\overline{\text{im}(T_\lambda)} = H$, hence $\sigma(T) = \sigma_c(T)$.

Theorem 31.1

Let H be a Hilbert space, and let $T \in \mathcal{B}(H)$. The resolvent set $\rho(T)$ is nonempty and open, and hence $\sigma(T)$ is closed.

Proof. First one wants to show that $\rho(T)$ is nonempty. Let $\lambda \in \mathbb{C}$ that satisfies $|\lambda| > \|T\|_{\mathcal{B}}$. Denote $T_\lambda = T - \lambda I = -\lambda(I - \lambda^{-1}T)$, here we define $A = \lambda^{-1}T$, where we see that we have:

$$\|A\| = \|T/\lambda\| = |\lambda|^{-1}\|T\| < 1$$

Hence according to the Neumann series, we can write the following:

$$T_\lambda^{-1} = -\lambda^{-1} \sum_{j=1}^{\infty} (-\lambda^{-1}T)^j$$

This shows that $\rho(T)$ is nonempty. In particular, we see here $\mathbb{C} \setminus B_{\|T\|}(0)$ is contained in $\rho(T)$. Now we will show that $\rho(T)$ is open. Take $\lambda_0 \in \rho(T)$, and let $\lambda \in \mathbb{C}$ that is close enough to λ_0 . We can write the following:

$$T_\lambda = T - \lambda I = T - \lambda_0 I - (\lambda - \lambda_0)I = T_{\lambda_0}(I - (\lambda - \lambda_0)R_{\lambda_0})$$

If $|\lambda - \lambda_0| < 1/\|R_{\lambda_0}\|$, then we see here T_λ is invertible by the Neumann series. Hence for all $\lambda_0 \in \rho(T)$, we see here $B_{1/\|R_{\lambda_0}\|}(\lambda_0)$ is contained in $\rho(T)$. This shows that $\rho(T)$ is open, which polishes off the results. \square

Theorem 31.2

Let H be a Hilbert space, and let $T \in \mathcal{B}(H)$.
 $\sigma(T)$ is compact and is contained in $\overline{B_{\|T\|}}(0)$.

Proof. Proposition 9.8 on p. 211 on *Hunter and Nachtergaele* states that $\sigma(T) \neq \emptyset$ if we have $T \in \mathcal{B}(H)$. We also know that $\sigma(T)$ is closed by Theorem 31.1. Moreover, in the proof of Theorem 31.1, we know that $\mathbb{C} - B_{\|T\|}(0) \subseteq \rho(T)$. That is, if $|\lambda| > \|T\|$, then $\lambda \in \rho(T)$, hence we know that $\sigma(T) \subseteq \overline{B_{\|T\|}}(0)$. Hence by Heine-Borel, we see here $\sigma(T)$ is compact. This completes the proof. \square

Spectrum of Compact Operators

Definition 32.0.0.0.1

Let H be a Hilbert space. A linear operator $T : H \rightarrow H$ is said to be compact provided that for all $S \subseteq H$, $T(S)$ is precompact.

Note: Compact linear operators are bounded linear operators.

Lemma 32.0.1

Let H be a Hilbert space. A linear operator $T : H \rightarrow H$ is compact if and only if for all bounded sequence (x_n) of points in H , there exists a subsequence $(x_{n_k})_{k \geq 1}$ of (x_n) such that $(Tx_{n_k})_{k \geq 1}$ is convergent in H .

Proof. The proof of this lemma is elementary, here we omit the proof. □

Theorem 32.1

Let H be a Hilbert space that is of infinite dimensional, and let $T : H \rightarrow H$ be a compact linear operator. We have the followings hold:

1. $0 \in \sigma(T)$.
2. $\sigma(T) = \sigma_p(T) \cup \{0\}$.
3. $\lambda \in \sigma_p(T) \setminus \{0\}$ has finite geometric multiplicity.¹
4. $\sigma_p(T)$ is either finite or countably infinite with only accumulation point at 0.

Proof. First we will prove (1), we proceed by contradiction, here we suppose that $0 \in \rho(T)$, then we see here $R_0(T) = T^{-1}$ exists and is bounded. Then for all bounded $S \subseteq H$, $T^{-1}(S)$ is bounded. Hence we know that $T(T^{-1}(S))$ is precompact, that is, the identity operator $TT^{-1} = I$ is a compact operator, but I is not compact because the closed unit ball is not compact in infinite dimensional space H , here we reach a contradiction, so we know that $0 \in \sigma(T)$. Now we will show that $\sigma(T) = \sigma_p(T) \cup \{0\}$. Consider $\lambda \in \sigma(T) \setminus \{0\}$, suppose $\lambda \notin \sigma_p(T)$, then we know that $\ker(T - \lambda I) = \{0\}$. By Theorem 2 in Chapter 9 on the Math 556 Lecture Note, we know that $T - \lambda I$ is Fredholm, so $\text{im}(T - \lambda I) = H$, hence $\lambda \in \rho(T)$, where we reach a contradiction. Next we will prove (3), let $\lambda \in \sigma_p(T) \setminus \{0\}$, first we condier that the multiplicity of λ is infinity, that is $\dim(V) = \infty$ where V is the space for eigenvector corresponds to λ . Take (x_n) of points in V , with the property that $(x_n) \rightarrow x$ for some $x \in V$, hence we can write $Tx_n = \lambda x_n$, and so we get the following:

$$Tx = \lim_{n \rightarrow \infty} Tx_n = T \left(\lim_{n \rightarrow \infty} x_n \right) = \lambda x$$

This shows that V is a closed space, and hence V is a Hilbert space as $V \subseteq H$ and is closed. The restriction of T on V can now be denoted as $T_V := T|_V$, and we see here for all $x \in V$, we have $T_V x = Tx = \lambda x$, and hence $\lambda^{-1}T_V$ is the identity operator on V . However, T_V is compact as it is a restriction of a compact operator on a Hilbert space V , but $\lambda^{-1}T_V$ is not compact, and hence we reach a contradiction, so $\lambda \in \sigma_p(T) \setminus \{0\}$ has finite geometric multiplicity. Lastly, we will prove (4), first we will show that 0 is the only accumulation point in $\sigma_p(T)$ when $\sigma_p(T)$ is an infinite set. Let (λ_n) be a sequence of points in $\sigma_p(T)$ such that $(\lambda_n) \rightarrow \lambda$. Since $\sigma_p(T)$ is infinite, we can assume that $\lambda_n \neq \lambda_m$ for $n \neq m$. Let (x_n) be eigenvector for λ_n , that is $Tx_n = \lambda_n x_n$. We

¹The geometric multiplicity of an eigenvalue is the dimension of the eigenspace of the eigenvalue λ

claim that $\{x_1, x_2, \dots\}$ is a set of linearly independent vectors. We first suppose that $\alpha_1 x_1 + \alpha_2 x_2 = 0$, then we can write the following:

$$T(\alpha_1 x_1 + \alpha_2 x_2) = 0 = \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 = 0$$

combining here with $\alpha_1 x_1 + \alpha_2 x_2 = 0$ we get:

$$\lambda_1(\alpha_1 x_1 + \alpha_2 x_2) - \alpha_1 \lambda_1 x_1 - \alpha_2 \lambda_2 x_2 = 0 \Rightarrow \alpha_2(\lambda_1 - \lambda_2)x_2 = 0 \Rightarrow \alpha_2 = \alpha_1 = 0$$

Now suppose $N \in \mathbb{N}$ is the maximal index j such that $\{x_1, x_2, \dots, x_j\}$ is a linearly independent set. Then there exist $\alpha_1, \alpha_2, \dots, \alpha_N$, not all zero, such that we have:

$$x_{N+1} = \sum_{j=1}^N \alpha_j x_j$$

and hence we can write the following:

$$Tx_{N+1} = \lambda_{N+1}x_{N+1} = \lambda_{N+1} \sum_{j=1}^N \alpha_j x_j = \sum_{j=1}^N \alpha_j \lambda_j x_j$$

rearranging we can write the following:

$$\alpha_j(\lambda_j - \lambda_{N+1}) = 0 \quad \forall j \in \{1, 2, \dots, N\}$$

and hence $\alpha_j = 0$ for all $j \in \{1, 2, \dots, N\}$, reaching a contradiction. This shows that $\{x_1, x_2, \dots\}$ is a set of linearly independent vectors. Let $X_n := \text{span}\{x_1, x_2, \dots, x_n\}$. Note that we have $X_{n-1} \subseteq X_n$, and we can also write $(T - \lambda_n I)(X_n) \subseteq X_{n-1}$. We can choose orthonormal sequence (e_n) such that $e_n \in X_n \cap X_{n-1}^\perp$ and $\|e_n\| = 1$. Hence we know that (e_n) converges weakly to 0 by Theorem 27.3. We will show later that $(Te_n) \rightarrow 0$ as T is a compact operator. Take any m, n such that $1 \leq m < n$, we can write the following:

$$\|Te_n - Te_m\| = \|(T - \lambda_n I)e_n - (T - \lambda_m I)e_m + \lambda_n e_n - \lambda_m e_m\|$$

note that we have $\lambda_n e_n \in X_n \cap X_{n-1}^\perp$, $(T - \lambda_n I)e_n \in X_{n-1}$, $(T - \lambda_m I)e_m \in X_{m-1} \subseteq X_{n-1}$, and $\lambda_m e_m \in X_m \subseteq X_{n-1}$, and hence we have:

$$\|Te_n - Te_m\| = \sqrt{\|\lambda_n e_n\|^2 + \|(T - \lambda_n I)e_n - (T - \lambda_m I)e_m - \lambda_m e_m\|^2} \geq |\lambda_n|$$

let n be sufficiently large, since we have $(\lambda_n) \rightarrow \lambda$, and $(Te_n) \rightarrow 0$, we see here:

$$\|Te_m\| \geq \lim_{n \rightarrow \infty} |\lambda_n| = |\lambda|$$

and hence for sufficiently large m , we see that $0 \geq |\lambda|$, that is $\lambda = 0$. Hence 0 is the only accumulation point of $\sigma_p(T)$. We are left to show that $\sigma_p(T)$ is countable, here we consider $S_k := \{\lambda \in \sigma_p(T) \mid |\lambda| \geq k\}$, for $k \in \mathbb{Q}$. We see here $\bigcup_k S_k = \sigma_p(T)$ is a countable union of sets, so it suffices to show that S_k is countable. Since $\sigma(T)$ is compact by Theorem 31.2, then S_k is compact as $\sigma_p(T) = \sigma(T) \setminus \{0\}$ is compact, if S_k were infinite, then all sequence (λ_n) of distinct points in S_k induces accumulation point λ in S_k , with $|\lambda| \geq k$, but this contradicts the result that we just proved, where $\lambda = 0$ is the only accumulation point in $\sigma_p(T)$, hence S_k can have only finite many points. \square

Lemma 32.1.1

Let H be a Hilbert space, let $T : H \rightarrow H$ be a compact operator, and let (x_n) be a sequence of points in H that converges weakly to x . We have $(Tx_n) \rightarrow Tx$.

Proof. Here we consider (y_n) defined by $y_n := Tx_n$, and $y = Tx$. We will show that $(\|y_n - y\|) \rightarrow 0$. First we want to show that (y_n) converges weakly to y , take $g \in H^*$, we define $l : H \rightarrow \mathbb{C}$ such that $l(x) = g(Tx)$ for all $x \in H$. We claim that $l \in H^*$, we clearly see that l is linear as it is a composition of linear functions. Moreover, we have $|l(x)| = |g(Tx)| \leq \|g\| \cdot \|Tx\| \leq \|g\| \cdot \|T\| \cdot \|x\|$, which shows that $\|l\| \leq \|g\| \cdot \|T\|$, hence $l \in H^*$. Since (x_n) converges weakly to x , then we know that $(l(x_n)) \rightarrow l(x)$, and hence $(g(y_n)) \rightarrow g(y)$ as we know that $Tx_n = y_n$ and $Tx = y$, this shows that (y_n) converges weakly to y . To show that $(y_n) \rightarrow y$, we proceed by contradiction. Suppose we have (y_n) does not converge to y , then for all $\epsilon > 0$ there exists a subsequence $(y_{s(k)})_{k \geq 1}$, such that $\|y - y_{s(k)}\| \geq \epsilon$. Note here we have $y_{s(k)} = Tx_{s(k)}$, on the other hand, $(x_{s(k)})_{k \geq 1}$ is a sequence in H that is bounded because subsequence of a weakly convergent sequence is bounded. Since T is a compact operator, then we know that $(Tx_{s(k)})_{k \geq 1}$ has a convergent subsequence $(\tilde{y}_j)_{j \geq 1}$, that is $(\tilde{y}_j)_{j \geq 1} \rightarrow \tilde{y}$, and hence $(\tilde{y}_j)_{j \geq 1}$ also converges weakly to \tilde{y} . Weakly convergent limit is unique, that is we must have $y = \tilde{y}$, here we reach a contradiction and hence we must have $(\tilde{y}_j) \rightarrow y$ and $\|y - \tilde{y}_j\| \geq \epsilon$. \square

The Spectral Theorem

Definition 33.0.0.0.1

Let $(X, (\cdot, \cdot)_X)$ and $(Y, (\cdot, \cdot)_Y)$ be Hilbert spaces, let $A : \mathcal{B}(X, Y)$. A is said to be self-adjoint if $A = A^*$.

Theorem 33.1

Let H be a Hilbert space, and let $A : H \rightarrow H$ be a compact self-adjoint operator, then the eigenvalues of A are all real, and we have either $\lambda = \|A\|$ or $\lambda = -\|A\|$ being an eigenvalue of A .

Remark: Recall that $\sigma(A) \subseteq \overline{B_{\|A\|}(0)}$, the theorem states that there exists $\lambda \in \mathbb{R}$ being an eigenvalue of A that satisfies $|\lambda| = \|A\|$.

Proof. Suppose $\lambda \in \sigma_p(A)$ with corresponding eigenvector $x \in H$, we can write $Ax = \lambda x$, hence we have:

$$(x, Ax) = (x, \lambda x) = \lambda(x, x) = \lambda\|x\|^2$$

Since A is self-adjoint, we can write:

$$(A^*x, x) = (Ax, x) = (\lambda x, x) = \bar{\lambda}(x, x) = \bar{\lambda}\|x\|^2$$

Since $x \neq 0$, we must have $\lambda = \bar{\lambda}$. By Theorem 7 in Chapter 7 on the Math 556 Lecture Note, we can write the following:

$$\|A\| = \sup_{u \neq 0} \frac{|(u, Au)|}{\|u\|^2}$$

For all $n \in \mathbb{N}$, there exists $u_n \in H$ such that we have:

$$\|A\| - \frac{1}{n} < \frac{|(u_n, Au_n)|}{\|u_n\|^2} \leq \|A\|$$

hence we can construct a nonzero sequence (u_n) of points in H such that we have:

$$\|A\| = \lim_{n \rightarrow \infty} \left| \left(\frac{u_n}{\|u_n\|}, A \frac{u_n}{\|u_n\|} \right) \right| \quad (8.2)$$

Here we normalize the sequence and denote $x_n := u_n / \|u_n\|$. Here we claim that $(x_n, Ax_n) \in \mathbb{R}$, this can be easily seen from the following:

$$\overline{(Ax_n, x_n)} = (x_n, Ax_n) = (A^*x_n, x_n) = (Ax_n, x_n)$$

Note that we have $|(x_n, Ax_n)|_{n \geq 1} \rightarrow \|A\|$, we choose a subsequence of $(x_{s(k)})_{k \geq 1}$ such that $(x_{s(k)}, Ax_{s(k)})$ has the same sign for all $k \in \mathbb{N}$, and hence $((x_{s(k)}, Ax_{s(k)}))_{k \geq 1} \rightarrow \lambda$ for some $\lambda \in \mathbb{R}$ that has the property $|\lambda| = \|A\|$. We will show that $\lambda \in \sigma_p(A)$. Note that we can write the following:

$$\|Ax_{s(k)} - \lambda x_{s(k)}\|^2 = (Ax_{s(k)} - \lambda x_{s(k)}, Ax_{s(k)} - \lambda x_{s(k)}) \quad (8.3)$$

$$= \|Ax_{s(k)}\|^2 - 2\lambda(x_{s(k)}, Ax_{s(k)}) + \lambda^2\|x_{s(k)}\|^2 \quad (8.4)$$

$$\leq \|A\|^2\|x_{s(k)}\|^2 - 2\lambda(x_{s(k)}, Ax_{s(k)}) + \lambda^2 \quad (8.5)$$

$$= 2\lambda^2 - 2\lambda(x_{s(k)}, Ax_{s(k)}) \quad (8.6)$$

for large enough k , from (8.2), it follows that we have:

$$\lim_{k \rightarrow \infty} (2\lambda^2 - 2\lambda(x_{s(k)}, Ax_{s(k)})) = 0 \quad (8.7)$$

Since A is compact, then there exists a subsequence of $(x_{s(k)})$, denoted as $(\tilde{x}_j)_{j \geq 1}$, such that we have $(A\tilde{x}_j) \rightarrow \xi$ for some $\xi \in H$. From (8.3) and (8.4) we see that we have:

$$\lim_{j \rightarrow \infty} A\tilde{x}_j - \lambda\tilde{x}_j = 0$$

and since $(A\tilde{x}_j) \rightarrow \xi$, then we have $(\lambda\tilde{x}_j) \rightarrow \xi$. That is, we can write the following:

$$\xi = \lim_{j \rightarrow \infty} \lambda\tilde{x}_j = \lim_{j \rightarrow \infty} A\tilde{x}_j = A \lim_{j \rightarrow \infty} \tilde{x}_j = \frac{A\xi}{\lambda}$$

which implies that we have:

$$A\xi = \lambda\xi \quad \Rightarrow \quad \lambda \in \sigma_p(A)$$

and hence completes the proof. \square

Theorem 33.2

Let H be a Hilbert space, and let $A : H \rightarrow H$ be a self-adjoint compact operator. H has an orthonormal eigenbasis.

Proof. From Theorem 33.1, we see that there exists $x_1 \in H$ such that x_1 is the eigenvector of A for some eigenvalue λ_1 with the property $|\lambda_1| = \|A\|$. We can normalize x_1 such that $\|x_1\| = 1$. Here we define $H_1 := H$ and $V_1 = \text{span}\{x_1\}$. Note here we can write $H_1 = V_1^\perp \oplus V_1$ by the projection theorem. Here we write $H_2 := V_1^\perp$. Note that if $u \in H_2$, we claim that we have $Au \in H_2$. To prove the claim, we write the following by the self-adjoint property of A :

$$(Au, x_1) = (u, Ax_1) = \lambda_1(u, x_1) = \lambda \cdot 0 = 0$$

So $Au \in H_2$, and hence we can define $A_2 := A|_{H_2} : H_2 \rightarrow H_2$. Note that for all bounded $S \subseteq H_2$, $A_2(S) = A(S)$ is precompact in H , and hence $\overline{A_2(S)} \subseteq H_2$ as H_2 is closed, which implies A_2 is compact. On the other hand, for $u, v \in H_2$, we can write the following:

$$(A_2u, v) = (Au, v) = (u, Av) = (u, A_2v)$$

and hence A_2 is also self-adjoint. That is, A_2 is a self-adjoint compact operator, then by Theorem 33.1, there exists $x_2 \in H_2$ such that $A_2x_2 = Ax_2 = \lambda_2x_2$, where $|\lambda_2| = \|A_2\|$. We can also normalize x_2 such that $\|x_2\| = 1$. Also note that $(x_1, x_2) = 0$ as we know that $x_2 \in H_2 = V_1^\perp$ and $x_1 \in V_1$. Here we define $V_2 := \text{span}\{x_1, x_2\}$ and $H_3 := V_2^\perp$. Continue in this way, (1) until there exists $n \in \mathbb{N}$ such that we reach a termination at step n , or (2) continue for all $n \in \mathbb{N}$. For case (1), we terminate at step $n \in \mathbb{N}$ provided that we have either (1a) $\dim(H) = n$, in which case $V_n = \text{span}\{x_1, x_2, \dots, x_n\} = H$ and $V_n^\perp = \{0\}$, or (1b) we have $H_{n+1} = V_{n+1}^\perp = \ker(A)$. In either (1a) or (1b), we have $H = V_n \oplus H_{n+1}$, and hence the eigenbasis of H is given by:

$$\{x_1, x_2, \dots, x_n\} \cup (\text{orthonormal basis of } H_{n+1})$$

In case (2), we get orthonormal sequence (x_n) of eigenvectors of A , with corresponding $(\lambda_n)_{n \geq 1}$ with $\lambda_n \neq 0$ for all $n \in \mathbb{N}$, and note here $(\lambda_n) \rightarrow 0$ by the fact that 0 is the only accumulation point in $\sigma_p(T)$ and $\sigma_p(T)$ is compact according to the proof of Theorem 32.1. Note that we also have $|\lambda_n| = \|A_n\|$, in which case we can write the followings:

$$\begin{aligned} |\lambda_1| &= \|A_1\| = \sup_{\|u\|=1, u \in H_1} |(u, Au)| \\ |\lambda_2| &= \|A_2\| = \sup_{\|u\|=1, u \in H_2} |(u, Au)| \\ &\vdots \end{aligned}$$

Since $H_2 \subseteq H_1$, it follows that we have:

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots$$

On the other hand, we know that (x_n) converges weakly to 0 as it is a sequence of orthonormal vectors, and hence we have $(Ax_n) \rightarrow 0$. From here, we can write:

$$(|\lambda_n|) = (|(Ax_n, x_n)|)_{n \geq 1} \rightarrow 0$$

Now we consider the set:

$$V := \left\{ \sum_{n=1}^{\infty} \alpha_n x_n \mid (\alpha_n) \in l^2 \right\} \subseteq H$$

where we see here V is the closed linear span of $\{x_1, x_2, \dots\}$. That is we can write $H = V \oplus V^\perp$. If $V^\perp = \{0\}$, the result follows immediately. If $V^\perp \neq \{0\}$, then for all $u \in V^\perp$, then $u \perp V_{k-1} = \text{span}\{x_1, x_2, \dots, x_{k-1}\}$ for all $k \geq 2$, that is $u \in H_k = V_{k-1}^\perp$ for all $k \geq 2$, and hence we can write:

$$\|Au\| = \|A_k u\| \leq \|A_k\| \|u\| = |\lambda_k| \|u\| \quad \Rightarrow \quad \left\| A \frac{u}{\|u\|} \right\| \leq |\lambda_k| \quad \forall k \geq 2$$

Take large enough k , we must have:

$$\left\| A \frac{u}{\|u\|} \right\| = 0 \quad \Rightarrow \quad V^\perp = \ker(A)$$

Hence we can write:

$$\text{basis of } H = \underbrace{\{x_1, x_2, \dots\}}_{\text{basis of } V} \cup \left(\text{orthonormal basis of } V^\perp \right)$$

where we see that the result follows. □

9 | Linear Differential Operators

Adjoints of Unbounded Linear Operators

Linear differential operators are not bounded. Consider the following operator:

$$A : C^1[0, 1] \rightarrow C^1[0, 1] \quad g \mapsto g'$$

where g' denotes the derivative of the function g . For $f : [0, 1] \rightarrow \mathbb{R} \quad x \mapsto \sin(nx)$ for some $n \in \mathbb{N}$, we see here $\|f\| = 1$, and $Af(x) = n \cos(nx)$ for all $x \in [0, 1]$, that is, we see that $\|Af\| = n$, and hence A is not bounded.

Linear operators that are not bounded, such as A , is usually not well-defined on the whole Hilbert space H , but it can be made well-defined on a dense subset $D(A)$ of the Hilbert space. For example, the differential operator A is well-defined on $C^1[0, 1]$, which is a dense subset of the $L^2[0, 1]$ space, equipped with the L^2 -norm that gives:

$$\|f\|_{L^2} = \frac{1}{\sqrt{2}} \quad \|Af\|_{L^2} = \frac{n}{\sqrt{2}}$$

Here f is the $x \mapsto \sin(nx)$ function that has been defined previously. Clearly, A is well-defined on $C^1[0, 1]$, but not well-defined on the whole space $L^2[0, 1]$.

Example: Let $k \in \{1, 2\}$, we write $A_k f = f''$ with A_1, A_2 defined on the following sets:

$$D(A_1) = \{u \in C^2[0, 1] \mid u(0) = u(1) = 0\} \quad D(A_2) = C^2[0, 1]$$

Here we see that we have $A_1 u = f \in C[0, 1]$ for $u \in D(A_1)$. On the other hand, $D(A_2)$ is dense in $H = L^2[0, 1]$. Moreover, A_2 is said to be an extension of A_1 because we have $D(A_1) \subseteq D(A_2)$, and $A_2 u = A_1 u$ for all $u \in D(A_1)$. Henceforth, a linear operator A is densely defined on a Hilbert space H provided that $\overline{D(A)} = H$.

Definition 34.0.0.0.1

Let $A : D(A) \rightarrow H$ be a linear operator with $D(A)$ being a dense subset of a Hilbert space $(H, (\cdot, \cdot))$. The adjoint of A is a linear operator $A^* : D(A^*) \rightarrow H$ that satisfies $(Ax, y) = (x, A^*y)$ for all $y \in D(A^*)$ and all $x \in D(A)$.

In Definition 34.0.0.0.1, $D(A^*)$ should be the largest subspace in H such that we have $(Ax, y) = (x, A^*y)$ for all $y \in D(A^*)$ and all $x \in D(A)$. For a given $y \in H$, we can define the following for $x \in H$:

$$\varphi_y : D(A) \rightarrow \mathbb{C} \quad x \mapsto (y, Ax)$$

in which case we see that φ_y is a linear functional on H but not necessarily bounded. We can define $y \in D(A^*)$ provided that φ_y is bounded on $D(A)$. Then we can extend the domain of φ_y from $D(A)$ to the whole space H . The Bounded Linear Transformation Theorem, Theorem 5.19 on *Hunter and Nachtergaele*, states that there exists a unique bounded linear functional $\bar{\varphi}_y : H \rightarrow \mathbb{C}$ such that we have $\bar{\varphi}_y(x) = \varphi_y(x)$ for all $x \in D(A)$ and $\|\bar{\varphi}_y\| = \|\varphi_y\|$. Next, we can use the Riesz Representation Theorem for $\bar{\varphi}_y$, that is there exists a unique $z \in H$ such that $\bar{\varphi}_y(x) = (z, x)$ for all $x \in H$. Now we define $A^* : D(A^*) \rightarrow H$ accordingly such that we have $A^*y = z$, in which case we can write the following for all $x \in D(A)$:

$$\varphi_y(x) = (y, Ax) = \bar{\varphi}_y(x) = (z, x) = (A^*y, x)$$

Concluding the result, we make the following definition:

Definition 34.0.0.0.2

Let $A : D(A) \rightarrow H$ be a unbounded linear operator densely defined on a Hilbert space H , the adjoint of A is defined by $A^* : D(A^*) \rightarrow H$ such that we have:

$$D(A^*) := \{y \in H \mid \exists z \in H \text{ such that } (y, Ax) = (z, x) \forall x \in D(A)\}$$

and for $y \in D(A^*)$, $A^*y = z$ is unique.

9.34.1 Self-adjoint Unbounded Operators

Definition 34.0.0.0.3

Let $A : D(A) \rightarrow H$ be a unbounded linear operator densely defined on a Hilbert space H . A is said to be self-adjoint provided that $A = A^*$ and note specifically $D(A) = D(A^*)$, and we have $(y, Ax) = (Ay, x)$ for all $x, y \in D(A)$.

Definition 34.0.0.0.4

Let $A : D(A) \rightarrow H$ be a unbounded linear operator densely defined on a Hilbert space H . A is said to be symmetric provided that A^* is an extension of A , that is we have $D(A) \subseteq D(A^*)$, and $(y, Ax) = (A^*y, x)$ for all $x \in D(A)$ and $A^*y = Ay$ for all $y \in D(A)$. That is, A is said to be symmetric if $(y, Ax) = (Ay, x)$ for all $x, y \in D(A)$.

Definition 34.0.0.0.5

Let $A : D(A) \rightarrow H$ be a unbounded linear operator densely defined on a Hilbert space H . A is said to be closed provided that, if (x_n) is a sequence of points in $D(A)$ that satisfies $(x_n) \rightarrow x$ and $(Ax_n) \rightarrow y$, then $x \in D(A)$ and $Ax = y$.

Definition 34.0.0.0.6

Let $A : D(A) \rightarrow H$ be a unbounded linear operator densely defined on a Hilbert space H , A is said to be closable provided that, if (x_n) is a sequence of points in $D(A)$ such that $(x_n) \rightarrow 0$ and $(Ax_n) \rightarrow y$ for some $y \in H$, we have $y = 0$.

Definition 34.0.0.0.7

Let $A : D(A) \rightarrow H$ be a closable unbounded linear operator densely defined on a Hilbert space H . The closure of A is defined by the following:

$$\bar{A} : D(\bar{A}) \rightarrow H$$

where $D(\bar{A})$ is a dense subset of H defined by the following:

$$D(\bar{A}) := \{x \in H \mid \exists (x_n) \text{ of points in } D(A), \text{ and } y \in H, \text{ s.t. } (Ax_n) \rightarrow y\}$$

and we define $\bar{A}x = y$ for all $x \in D(\bar{A})$.

Definition 34.0.0.0.8

Let $A : D(A) \rightarrow H$ be a closable unbounded linear operator densely defined on a Hilbert space H . A is said to be essentially self-adjoint provided that A is symmetric and the closure of A is self-adjoint.

Example: Consider $A : D(A) \rightarrow H$ $u \mapsto u''$ where $H = L^2[0, 1]$, and we write:

$$D(A) := \{u \in C^2[0, 1] \mid u(0) = u(1) = 0\}$$

Here A is not closed. We will construct a sequence (u_n) in $D(A)$. For $n \in \mathbb{N}$, we define $u_n(x)$ for all $x \in [0, 1]$ by the following:

$$u_n(x) = \begin{cases} \frac{x^2}{2} - \left(\frac{1}{2} + \frac{1}{6n^2}\right)x + \frac{1}{6n^2} & x > 1/n \\ \frac{nx^3}{6} + \left(\frac{1}{2n} - \frac{1}{2} - \frac{1}{6n^2}\right)x & 0 \leq x \leq 1/n \end{cases}$$

It can be checked easily that $u_n \in D(A)$, and we have:

$$u_n''(x) = \begin{cases} 1 & x > 1/n \\ nx & 0 < x < 1/n \end{cases}$$

and hence we see that $(u_n''(x))$ converges to 1 if $x \in (0, 1]$, and converges to 0 if $x = 0$. Note that $(Au_n) = (u_n'')$ converges to $y \in H$, which is not continuous at $x = 0$, so u_n does not have limit in $D(A)$, this shows that A is not closed. Here A is also symmetric, as we can write the following for all $u, v \in D(A)$:

$$\begin{aligned} (Au, v) &= \int_0^1 u''(x)v(x) dx = \int_0^1 (u'(x))'v(x) dx \\ &= u'v|_0^1 - \int_0^1 u'v' \\ &= u'v|_0^1 - uv'|_0^1 + \int_0^1 uv'' \\ &= \int_0^1 u(x)v''(x) dx = (u, Av) \end{aligned}$$

One can show that A is closable and essentially self-adjoint with the closure of A defined by $\bar{A} : D(\bar{A}) \rightarrow H$, with the domain of \bar{A} :

$$D(\bar{A}) = \{u \in \mathbb{H}^2[0, 1] \mid u(0) = u(1) = 0\}$$

where $\mathbb{H}^2[0, 1]$ is the Sobolev space over the interval $[0, 1]$.

9.34.2 Adjoint of Linear Differential Operator**Definition 34.0.0.0.9**

A is a linear differential operator of order n provided that A is a linear map on space of n times continuously differentiable functions defined on (α, β) such that we have the following holds for $u \in C^n[\alpha, \beta]$:

$$Au = \sum_{j=0}^n a_j u^{(j)}$$

Here we assume that $u \in C^2[0, 1]$, by utilizing change of variable, u is equivalently defined on any closed interval $[\alpha, \beta]$. Now we write:

$$Au = au'' + bu' + cu \quad (9.1)$$

where a, b, c are sufficiently smooth. We also assume that $a(x) > 0$ for all $x \in [0, 1]$, and two conditions. Some common boundary conditions include the Dirichlet boundary condition which states $u(0) = u(1) = 0$, the Neumann boundary condition which states $u'(0) = u'(1) = 0$, and the periodic boundary condition which states $u(0) = u(1)$, $u'(0) = u'(1)$. We can also have initial conditions that state, for example, $u(0) = 0$, $u'(0) = 0$, or final condition that state, for example, $u(1) = 0$, $u'(1) = 0$. We denote the conditions in its compact form $Bu = 0$

Here we want to solve the system $Au = f$, $Bu = 0$, the domain of A is given by:

$$D(A) = \{u \in C^2[0, 1] \mid Bu = 0\}$$

In some cases, one has $Bu = \beta \neq 0$, for which case one can find $v \in C^2[0, 1]$ such that $Bv = \beta$, then define $w = u - v$, and write $Bw = B(u - v) = 0$, with $Aw = f - Av$, and solve the system $Aw = f - Av$, $Bw = 0$.

Theorem 34.1 (Green's Theorem)

Consider A given by (9.1), with $b \in C^1[0, 1]$, $c \in C[0, 1]$, and positive $a \in C^2[0, 1]$, and we utilize the inner product $(u, v) = \int_0^1 \bar{u}v \, dx$ for $u, v \in L^2[0, 1] = H$. Then we have:

$$(v, Au) - (A^*v, u) = [a(\bar{v}u' - \bar{v}'u) + (b - a)\bar{v}u]_0^1 \quad (9.2)$$

$$A^*v = (\bar{a}v)'' - (\bar{b}v)' + \bar{c}v \quad (9.3)$$

Proof. Here we can write the following:

$$\begin{aligned} (v, Au) &= \int_0^1 \bar{v}(au'' + bu' + cu) \, dx = [a\bar{v}u' + b\bar{v}u]_0^1 + \int_0^1 (-(a\bar{v})'u' - (b\bar{v})'u + c\bar{v}u) \, dx \\ &= [a\bar{v}u' + b\bar{v}u - (a\bar{v})'u]_0^1 + \int_0^1 ((a\bar{v})'' - (b\bar{v})' + c\bar{v})u \, dx \end{aligned}$$

rearranging we get:

$$(v, Au) - (A^*v, u) = [a\bar{v}u' + b\bar{v}u - (a\bar{v})'u]_0^1$$

This completes the proof. □

Theorem 34.1 gives us a criteria to find the adjoint of A , or check if A^* for a given operator is well-defined.

The boundary conditions for A^* are $B^*v = 0$ computed such that the boundary term in Green's Theorem (9.2) vanishes. Here we consider a simple example to ensure the existence of A^* under Dirichlet boundary condition $Bu = 0$. In such case, we have $u(0) = u(1) = 0$, and hence $B^*v = 0$ implies $v(0) = v(1) = 0$. Then we can use (9.3) to write the following to find A^* :

$$\begin{aligned} A^*v &= (\bar{a}v)'' - (\bar{b}v)' - \bar{c}v \\ &= \bar{a}v'' + (2\bar{a}' - \bar{b})v' + (\bar{a}'' - \bar{b}' + \bar{c})v \end{aligned}$$

applying the Dirichlet boundary condition to Theorem 34.1, here we want the following holds for self-adjoint operator A :

$$\bar{a}v'' + (2\bar{a}' - \bar{b})v' + (\bar{a}'' - \bar{b}' + \bar{c})v = av'' + bv' + cv$$

which reads the following:

$$\begin{aligned} a = \bar{a} &\Rightarrow a \text{ is real-valued} \\ b = 2\bar{a}' - b &\Rightarrow b + \bar{b} = 2\Re(b) = 2a' \\ \bar{a}'' - \bar{b}' + \bar{c} = c &\Rightarrow c - \bar{c} = 2i\Im(c) = \overline{a''} - \bar{b}' = a'' - \Re(b') + i\Im(b') = i\Im(b') \end{aligned}$$

combining we get:

$$a \text{ is real} \quad \Re(b) = a' \quad \Im(c) = \Im(b)/2$$

In the case that all coefficients a, b, c are all real, we can write the following:

$$a \text{ is arbitrary} \quad b = a' \quad c \text{ is arbitrary}$$

and hence we require A to be of the form:

$$Au = au'' + a'u' + cu = (au')' + cu \quad (\text{Sturm Liouville Operator } A)$$

Such operator is further studied in *Hunter and Nachtergaele*.

9.34.3 Green's Function

Consider the case where we have $Au = au'' + bu' + cu$, with $Bu = 0$ being the Dirichlet boundary condition, that we have discussed in section 9.34.2. For $f \in C[0, 1]$, we consider the problem $Au = f$, $Bu = 0$, with the domain of A :

$$D(A) = \{u \in C^2[0, 1] \mid Bu = 0\}$$

We look for solution of the form:

$$u(x) = \int_0^1 g(x, y) f(y) dy = Gf(x)$$

where $g(x, y)$ is called the Green's function.

The Dirac-delta function $\delta(x)$ is a distribution that has the following property:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \delta(x_0) = 0 \quad \text{for } x_0 \neq 0$$

From here we get:

$$\int_{-\infty}^{\infty} f(y)\delta(x, y) dy = f(x) \quad \int_{-\infty}^{\infty} \delta(y) dy = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

If we can understand the following problem:

$$Ag(x, y) = \delta(x - y) \quad g(0, y) = g(1, y) = 0 \quad Au = \int_{-\infty}^{\infty} Ag(x, y) f(y) dy = f(x)$$

in other words, we can write the following:

$$Ag(x, y) = \left(a \frac{\partial^2}{\partial x^2} + b \frac{\partial}{\partial x} + c \right) g(x, y)$$

then we can solve the following system along with its boundary condition:

$$Au = f \quad u(0) = u(1) = 0 \quad (9.4)$$

Definition 34.1.0.0.1

$g : [0, 1] \times [0, 1] \rightarrow \mathbb{C}$ is a Green's function provided that it satisfies the followings:

1. $g(x, y)$ is continuous on $0 \leq x, y \leq 1$.
2. $g(x, y)$ is twice continuously differentiable with respect to x on $0 \leq x \leq y \leq 1$, and $0 \leq y \leq x \leq 1$.
3. $g(x, y)$ satisfies $Ag(x, y) = 0$ for $0 < x < y < 1$ and $0 < y < x < 1$, the boundary condition $g(0, y) = g(1, y) = 0$, and we have:

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

where we write:

$$\lim_{x \rightarrow y^+} \frac{\partial}{\partial x}g(x, y) = \frac{\partial}{\partial x}g(y^+, y) \quad \lim_{x \rightarrow y^-} \frac{\partial}{\partial x}g(x, y) = \frac{\partial}{\partial x}g(y^-, y)$$

Proposition 34.1.1

From the settings in the previous discussion, we have a solution to (9.4) given by:

$$Gf = \int_0^1 g(x, y) f(y) dy = u(x)$$

Proof. Here we can write:

$$Au(x) = \left(a(x) \frac{\partial^2}{\partial x^2} + b(x) \frac{\partial}{\partial x} + c(x) \right) \left(\int_0^x g(x, y) f(y) dy + \int_x^1 g(x, y) f(y) dy \right)$$

where we have:

$$\begin{aligned} \frac{\partial}{\partial x} \int_0^x g(x, y) f(y) dy &= g(x, x) f(x) + \int_0^x \frac{\partial}{\partial x} g(x, y) f(y) dy \\ \frac{\partial}{\partial x} \int_x^1 g(x, y) f(y) dy &= -g(x, x) f(x) + \int_x^1 \frac{\partial}{\partial x} g(x, y) f(y) dy \\ b(x) \frac{\partial}{\partial x} (G) &= \int_0^1 b(x) \frac{\partial}{\partial x} g(x, y) f(y) dy \end{aligned}$$

then we get:

$$\begin{aligned} \frac{\partial^2}{\partial x^2} \left(\int_0^x g(x, y) f(y) dy + \int_x^1 g(x, y) f(y) dy \right) &= \frac{\partial}{\partial x} \left(\int_0^x \frac{\partial}{\partial x} g(x, y) f(y) dy + \int_x^1 \frac{\partial}{\partial x} g(x, y) f(y) dy \right) \\ &= \frac{\partial}{\partial x} g(x, x^-) f(x) - \frac{\partial}{\partial x} g(x, x^+) f(x) + \int_0^1 \frac{\partial^2}{\partial x^2} g(x, y) f(y) dy \\ &= f(x) \left(\frac{\partial}{\partial x} g(x^+, x) - \frac{\partial}{\partial x} g(x^-, x) \right) + \int_0^1 \frac{\partial^2}{\partial x^2} g(x, y) f(y) dy \end{aligned}$$

Combining we get:

$$\begin{aligned}
 Au(x) &= +a(x)f(x) \left(\frac{\partial}{\partial x}g(x^+, x) - \frac{\partial}{\partial x}g(x^-, x) \right) \\
 &\quad + \int_0^1 \left(a(x) \frac{\partial^2}{\partial x^2}g(x, y)f(y) + b(x) \frac{\partial}{\partial x}g(x, y)f(y) + c(x)g(x, y)f(y) \right) dy \\
 &= a(x)f(x) \frac{1}{a(x)} + \int_0^1 Ag(x, y)f(y) dy \\
 &= f(x)
 \end{aligned}$$

The result follows. \square

In the following we will discuss the construction of $g(x, y)$. Consider two solutions $Av_1 = 0$ and $Av_2 = 0$, with $v_1(0) = v_2(1) = 0$ and $v_1, v_2 \in C^2[0, 1]$. We seek for $g(x, y)$ that satisfies:

$$g(x, y) = \begin{cases} C(y)v_1(x)v_2(y) & 0 \leq x \leq y \\ C(y)v_1(y)v_2(x) & y \leq x \leq 1 \end{cases}$$

As a sanity check, we see that we have:

$$g(0, y) = 0 = g(1, y) \quad Ag = 0 \text{ for all } x \in (0, y) \cup (y, 1)$$

and we also have:

$$\frac{\partial}{\partial x}g(y^+, y) - \frac{\partial}{\partial x}g(y^-, y) = C(y) (v_1(y)v_2'(y) - v_1'(y)v_2(y)) = C(y)\mathcal{W}$$

where \mathcal{W} is the Wronskian.

If v_1, v_2 are linearly independent, then we have $\mathcal{W} \neq 0$, and hence we can define:

$$C(y) = \frac{1}{a(y) \cdot \mathcal{W}}$$

We will show that v_1 and v_2 are linearly independent if the system $Au = 0$ with $u(0) = u(1) = 0$ has a unique solution $u = 0$, that is $\ker(A) = \{0\}$. To see that, we proceed by contradiction, where suppose $v_2(x) = \alpha v_1(x)$, and in such case we can write the followings:

$$Av_1 = 0 \quad Av_2 = 0 \quad v_1(0) = v_1(1) = 0 \quad v_2(0) = v_2(1) = 0$$

In which case we see that $\ker(A) \neq \{0\}$, hence v_1 and v_2 are indeed linearly independent if $\ker(A) = \{0\}$. Concluding such result, if A has trivial kernel, then we can construct the Green's function g , and use Proposition 34.1.1 to obtain $u(x)$.